Development of Mining Sector Applications for Emerging Remote Sensing and Deep Learning Technologies
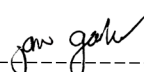

Jane Gallwey

Camborne School of Mines

College of Engineering, Mathematics and Physical Sciences

University of Exeter

Submitted by Jane Gallwey to the University of Exeter as a thesis for the degree of Doctor of Philosophy in Mining and Minerals Engineering, February 2021. This thesis is available for Library use on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement. I certify that all material in this thesis which is not my own work has been identified and that any material that has previously been submitted and approved for the award of a degree by this or any other University has been acknowledged.

Signature _____   23rd February 2021

Abstract

This thesis uses neural networks and deep learning to address practical, real-world problems in the mining sector. The main focus is on developing novel applications in the area of object detection from remotely sensed data. This area has many potential mining applications and is an important part of moving towards data driven strategic decision making across the mining sector. The scientific contributions of this research are twofold; firstly, each of the three case studies demonstrate new applications which couple remote sensing and neural network based technologies for improved data driven decision making. Secondly, the thesis presents a framework to guide implementation of these technologies in the mining sector, providing a guide for researchers and professionals undertaking further studies of this type.

The first case study builds a fully connected neural network method to locate supporting rock bolts from 3D laser scan data. This method combines input features from the remote sensing and mobile robotics research communities, generating accuracy scores up to 22% higher than those found using either feature set in isolation. The neural network approach also is compared to the widely used random forest classifier and is shown to outperform this classifier on the test datasets. Additionally, the algorithms' performance is enhanced by adding a confusion class to the training data and by grouping the output predictions using density based spatial clustering. The method is tested on two datasets, gathered using different laser scanners, in different types of underground mines which have different rock bolting patterns. In both cases the method is found to be highly capable of detecting the rock bolts with recall scores of 0.87-0.96.

The second case study investigates modern deep learning for LiDAR data. Here, multiple transfer learning strategies and LiDAR data representations are examined for the task of identifying historic mining remains. A transfer learning approach based on a Lunar crater detection model is used, due to the task similarities between both the underlying data structures and the geometries of the objects to be detected. The relationship between dataset resolution and detection accuracy is also examined, with the results showing that the approach is capable of detecting pits and shafts to a high degree of accuracy with

precision and recall scores between 0.80-0.92, provided the input data is of sufficient quality and resolution. Alongside resolution, different LiDAR data representations are explored, showing that the precision-recall balance varies depending on the input LiDAR data representation.

The third case study creates a deep convolutional neural network model to detect artisanal scale mining from multispectral satellite data. This model is trained from initialisation without transfer learning and demonstrates that accurate multispectral models can be built from a smaller training dataset when appropriate design and data augmentation strategies are adopted. Alongside the deep learning model, novel mosaicing algorithms are developed both to improve cloud cover penetration and to decrease noise in the final prediction maps. When applied to the study area, the results from this model provide valuable information about the expansion, migration and forest encroachment of artisanal scale mining in southwestern Ghana over the last four years.

Finally, this thesis presents an implementation framework for these neural network based object detection models, to generalise the findings from this research to new mining sector deep learning tasks. This framework can be used to identify applications which would benefit from neural network approaches; to build the models; and to apply these algorithms in a real world environment. The case study chapters confirm that the neural network models are capable of interpreting remotely sensed data to a high degree of accuracy on real world mining problems, while the framework guides the development of new models to solve a wide range of related challenges.

Acknowledgements

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AI	Artificial Intelligence
ASM	Artisanal Scale Mining
BGS	British Geological Survey
CCD	Charge Coupled Device
CMOS	Complementary Metal Oxide Semiconductor
CNN	Convolutional Neural Network
DBSCAN	Density Based Spatial Clustering of Applications with Noise
DEFRA	Department for Environment, Food and Rural Affairs
DEM	Digital Elevation Model
DSM	Digital Surface Model
DTM	Digital Terrain Model
EM	Electro-Magnetic
ESA	European Space Agency
FPFH	Fast Point Feature Histogram
GB	Gigabyte
GEOBIA	Geographic Object Based Image Analysis
GIS	Geographic Information System
GNSS	Global Navigation Satellite System
GPU	Graphics Processing Unit
GSD	Ground Sampling Distance
IMU	Inertial Measurement Unit
LBFGS	Limited-memory Broyden Fletcher Goldfarb Shanno
LiDAR	Light Detection and Ranging
LRM	Local Relief Model
LULC	Land Use Land Cover
MATLAB	Matrix Laboratory
MLP	Multi-Layer Perceptron
MODIS	Moderate Resolution Imaging Spectroradiometer
NASA	National Aeronautics and Space Administration
NDVI	Normalized Difference Vegetation Index
OSGB	Ordnance Survey Great Britain
PCA	Principal Component Analysis
ReLU	Rectified Linear Unit
RF	Random Forest
RGB	Red Green Blue
SAM	Spectral Angle Mapping
SCP	Simplified Classification Plugin
SDG	Sustainable Development Goals
SfM	Structure from Motion
SLAM	Simultaneous Location and Mapping
SLRM	Simplified Local Relief Model
SVF	Sky View Factor
SVM	Support Vector Machines
TLS	Terrestrial laser Scanning
UAV	Unmanned Aerial Vehicle
VHR	Very High Resolution

Declaration

I declare that the thesis has been composed by myself and that the work has not been submitted for any other degree or professional qualification. I confirm that the work submitted is my own, except where work which has formed part of jointly-authored publications has been included. This thesis includes three original papers published in peer reviewed journals and one conference paper published in peer reviewed proceedings. The theme of my research is applied deep learning for mining sector remote sensing applications. My contribution and those of the other authors to this work have been explicitly indicated at the start of each chapter. I confirm that appropriate credit has been given within this thesis where reference has been made to the work of others.

Jane Gallwey

23/02/2021

Introduction

1.1 Motivation

Mining is a critical global sector, as all objects used by societies are made from materials that are mined, extracted or grown. However, the world's easily accessible high-grade orebodies are becoming depleted, leaving only lower grade, deeper and more remotely situated deposits for future exploitation (Valenta et al., 2018). Unlike almost all other industries, mining cannot be moved; it must be located where the resource is regardless of security or topographic concerns. Other current challenges include market volatility, geopolitical risk, legal limits on natural resource use, shareholder activism and increased public scrutiny (Maennling and Toledano, 2019). To rise to these challenges, the mining sector must adapt to changing technologies for increased productivity and safety. Increasing digitisation can lead to better decision making based on more current, accurate data.

In recent years, remote sensing data capture systems have advanced rapidly and are now capable of generating ever greater quantities of data at unprecedented levels of detail. In 2020, human analysts cannot physically examine the volumes of data being generated from systems such as earth observation satellites, national Light Detection and Ranging (LiDAR) campaigns and autonomous vehicles' sensors. However, neural network based machine learning algorithms for Artificial Intelligence (AI) have been evolving even faster, fuelled by an exponential increase in computing power over the last decade. These advanced algorithms may offer a solution to manage and interpret these large volumes of remotely sensed data. This thesis will investigate how these dual technological advances could be adapted and leveraged by the mining industry.

1.2 Aims and objectives

The aim of this thesis is to investigate how different remote sensing data formats commonly used in mining can be processed using neural network pipelines to improve strategic analysis of large datasets.

To deliver this aim, the research objectives are:

1. To create a framework for identifying, developing and applying deep learning algorithms to mining sector remote sensing data.

2. Use this framework to design and implement algorithms for detection in:

    a. underground mine environments using 3D laser scan data

    b. regional landscapes using aerial LiDAR.

    c. national scale analysis from multispectral satellite imagery.

1.3 Methodology

To achieve these objectives this thesis first introduces the main concepts of remote sensing data capture systems, underlying data structures and neural network based algorithms. The literature review also investigates how these topics are currently utilised in the mining sector; together with how other industries are leveraging these technologies in similar circumstances. The literature review identifies that the most important factor for methodology design is the remote sensing data type. Therefore, the research plan was developed to demonstrate how neural networks can be applied to a wide range of data types for mining sector tasks, using case studies to develop methodologies for data analysis and associated interpretation.

Each case study, written as separate chapters, identifies a promising mining sector task for each data type, before designing and implementing a neural network based solution appropriate to the data type, scale and use case for the task. The individual methodologies developed for each task are described within the relevant chapters. The datasets used in Chapters 4 and 5 were generated from publicly available open source data, while the datasets used in Chapter 3 are from archive Camborne School of Mines data and an industry sponsored

data capture exercise. By applying the techniques to real world datasets, the challenges, benefits and accuracies of these technologies can be better understood.

1.4 Thesis structure

This thesis consists of seven chapters, including the introduction (Chapter 1) and the conclusion (Chapter 7). The main content chapters are as follows:

- Chapter 2 presents the literature review. This chapter begins with an overview of the topics of remote sensing and deep learning, alongside a review of integrations of these technologies.

- Chapter 3 describes a method for automatically detecting rock bolts from 3D laser scan data to generate as-built reports on installed bolting patterns.

- Chapter 4 details an application to detect historic mining pits from aerial LiDAR datasets. Next, the pit locations are used to infer mineralisation trends across historic mining areas.

- Chapter 5 sets out a novel method to detect small scale mining from satellite imagery, achieved by building and training a deep learning model based on multispectral imagery.

- Chapter 6 integrates the knowledge gained from the three case studies into a framework for applying these technologies to other mining sector tasks.

1.5 Contribution

This thesis aims to demonstrate how neural network based machine learning combined with remote sensing can be used to solve real world mining sector challenges. The research projects carried out in this PhD make contributions to knowledge both individually and collectively.

Individually, each project delivers new knowledge about how to design, implement and test effective neural network based methodologies for different types of remotely sensed mining datasets. Additionally, the datasets of detected objects generated by each case studies' research can be used for further investigation in their respective areas. Each of the case studies have been published as peer reviewed journal papers over the course of this research, demonstrating their academic contribution to knowledge in the field of remote sensing and artificial intelligence, in various deployment areas such as underground and surface mining operations.

Collectively, the knowledge gained from all the projects contributes to the creation of a framework for applied neural network and deep learning research in the mining sector. The implementation framework and the varied real world applications showcased in this research contribute practical knowledge to a wide range of stakeholders across the mining sector, from remote sensing professionals to mine managers. A greater awareness of the possibilities of these new technologies by mining stakeholders provides a solid foundation for further technology and policy research.

Literature review

2.1 Introduction

In the last decade, the mining sector has undergone unprecedented changes. Maennling and Toledano (2019) identified seven trends shaping the modern mining industry. These include access to new resources, developing new mining workforces, social contracts to operate, exploiting big data, new methods of financing mining, world geopolitics and transitions to low carbon economies. Technological drivers such as automation, remote sensing systems and big data analytics will be crucial for enabling agile adaptations by the mining sector. Over the full mining lifecycle, these technologies generate or process vast amounts of data, which will need new methods of big data processing to harness their benefits. Adapting knowledge from the technology, automotive and space sectors will aid progress in implementing these technologies in the mining sector.

Alongside technological changes, mining policies are also changing. Mining often has a negative public image, exacerbating the difficulties involved with gaining social license to operate (Prno and Scott Slocombe, 2012). Maintaining a social license to operate in a local community is becoming ever more critical. The push towards automation has led to decreasing amounts of local employment; this can mean the local community is left with more environmental concerns than financial gains (Prno and Scott Slocombe, 2012), leading to protests and sometimes the abandonment of projects. To address these concerns, stricter environmental regulations have been developed, particularly around emissions, water usage, tailings and mine closure (Tuokuu et al., 2019). Rehabilitation of end of life projects is a major issue; if closure bonds are insufficient, mines can be left in perpetual 'care and maintenance' status or abandoned without sufficient environmental protections (Ashby et al., 2016). Additionally, older mining regions were often not adequately rehabilitated, as the understanding of the impacts of mine abandonment was not fully appreciated for many centuries (CIRIA, 2019). In many developing countries, managing their growing Artisanal Scale Mining (ASM) sector is also a concern, where a balance must be found between livelihoods and environmental damages (Hilson and Gatsinzi, 2014). Figure 2-1 maps the connections in the

mining sector to these future drivers of change (Maennling and Toledano, 2019).
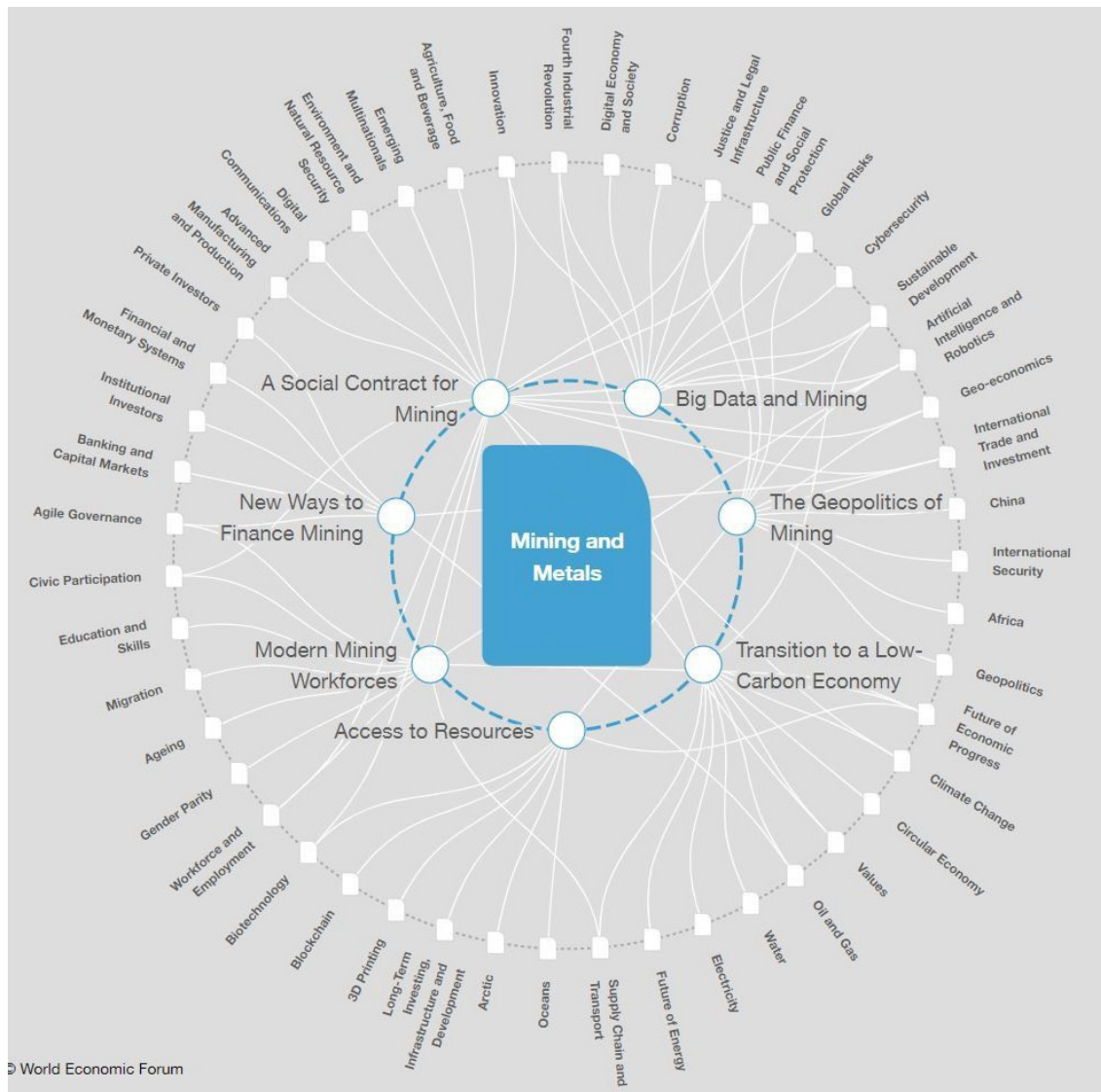
Figure 2-1: Mapping the connections in mining, World Economic Forum (Maennling and Toledano, 2019).

In response to these changes, the mining industry has seen legislative and technological reform in recent years. Many mining companies are now exploring how technology can be used to reduce costs, improve safety and environmental conditions. One such area that has experienced a significant advancement is remotely sensed data, applied to applications such as geotechnical analysis (Lato et al., 2009) and stockpile monitoring (Tong et al., 2015). Using remotely captured data reduces human exposure to hazardous environments; however, much of the data processing has been largely manual. Remote sensing datasets combined with deep learning algorithms could be leveraged to aid both the technological and social challenges facing the mining sector today.

This literature review first introduces remote sensing technologies and data types, before giving an overview of deep learning theory. The final section examines existing mining sector applications that leverage these technologies, alongside potential compatible applications from other sectors. The aim of this chapter is to give a broad overview of the technologies used in this research and an introduction to the breadth of applications which are currently utilising these technologies. To minimise duplication, references which are specific to the applications developed in the case studies are not discussed here, as they are included in the relevant chapter's introductory sections.

2.2 Remote sensing technology

Remote sensing is the broad scientific field concerned with acquiring information about an object without direct contact. By convention, the term remote sensing is commonly used to refer to geographic observations made from airborne and spaceborne platforms (Weng, 2012), while the terms high definition surveying (Frei et al., 2004) and close range photogrammetry (Luhmann et al., 2014) are more typically used to refer to observations made from a terrestrial platform. In this thesis, all types of spatial data collected remotely from terrestrial, aerial or orbital platforms will be collectively referred to as remote sensing data. In contrast to traditional land surveying methods, remote sensing systems of all types generate large volumes of data.

Remote sensing systems can be classified into two types, depending on whether they provide the energy source used for sensing the remote object themselves. Passive remote sensing systems record energy either emitted from the sun and reflected by the object or emitted by the object itself, while active remote sensing systems emit their own energy and measure the time it takes for this to travel to the object of interest and return to the sensor (Weng, 2012). Remote sensing systems all use electromagnetic (EM) radiation as their energy source. Electromagnetic radiation is formed of coupled electric and magnetic fields which travel in waves through a vacuum at the speed of light. EM waves are defined by their wavelength and frequency, which are inversely related because the speed of light is constant (Weng, 2012). Whilst EM radiation exists along a continuous spectrum, man-made categorisations are used to differentiate between the different regions (Weng, 2012). The regions used in

remote sensing are most conveniently described by a logarithmic plot such as that shown in Figure 2-2, as their wavelengths span multiple orders of magnitude (Lillesand et al., 2015).
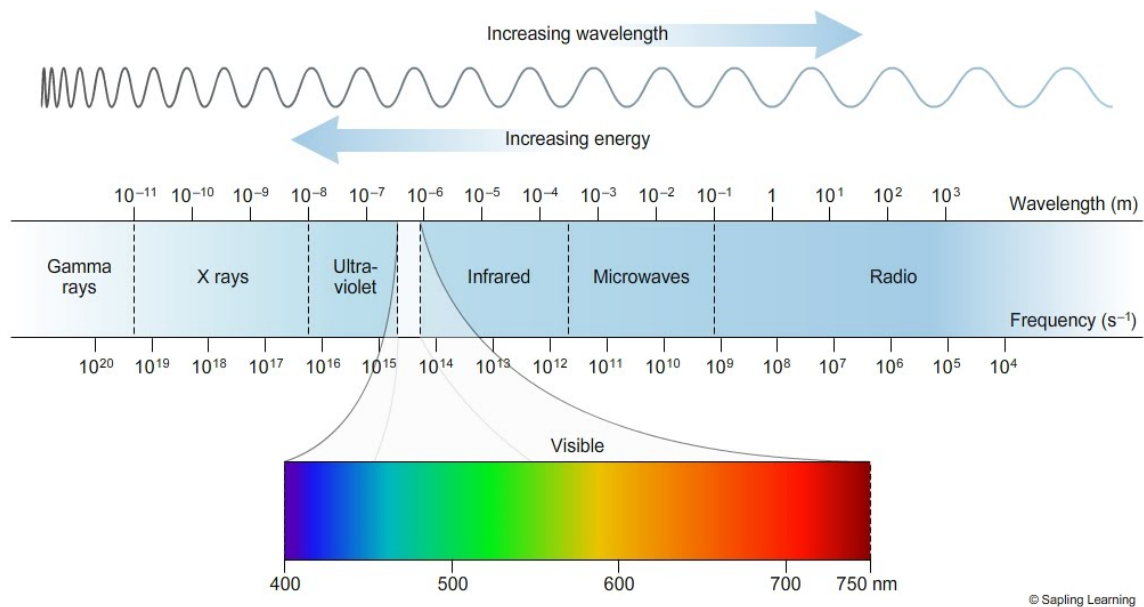
Figure 2-2: Electromagnetic spectrum, with inset showing visible light range (Sapling Learning, 2015).

The hardware used in remote sensing systems varies depending on whether it is a passive or active sensing system. Hardware systems also vary depending on the scale and accuracy required.

2.2.1 Passive remote sensing systems

Historically, passive remote sensing systems (cameras) used light sensitive photographic film to record a scene. Since 2010, modern systems using digital sensors have almost entirely replaced film (Lillesand et al., 2015). Digital systems use an array of either charge-coupled device (CCD) or complementary metal oxide semiconductor (CMOS) detectors which generate an electrical charge when exposed to EM energy (Jensen, 2007). The magnitude of the charge corresponds to the magnitude of the energy hitting the detector (McGlone et al., 2004). To split the observed EM radiation into discrete bands, filters or beamsplitters are used to ensure each detector cell receives radiation only from a specific band (Weng, 2012). Standard consumer cameras, small scientific cameras, drone mounted cameras and most aerial photogrammetric cameras primarily use a rectangular array of detectors with RGB (Red Green

Blue) filters, arranged in a Bayer filter mosaic (Bayer, 1975). These types of cameras are rarely designed to capture more than four EM bands.

Some aerial photogrammetric cameras, along with most spaceborne cameras utilise a different hardware design, made up of multiple lines of detectors sensitive to many different bands (Weng, 2012). These are either 'across-track' systems, where a rotating mirror deflects light from side to side, perpendicular to the platform's direction of travel, or 'along-track' systems, where an entire line of detectors is exposed at once (Jensen, 2007). Across-track systems are optically simpler and have been used on older remote sensing satellite systems such as the Landsat series (Williams et al., 2006), whilst more modern spaceborne systems such as Spot, QuickBird and the Sentinel-2 Multispectral Imager (MSI) use along track technology (Berger et al., 2012; Chevrel et al., 1981; Toutin and Cheng, 2002). Figure 2-3 shows the principles of a number of different passive remote sensing systems (Jensen, 2007). For a thorough overview of modern digital photographic hardware within the context of remote sensing, including advanced topics such as thermal infrared, passive microwave and hyperspectral systems not covered in this thesis, see Lillesand et al. (2015) Chapters 2 & 4 and Weng (2012) Chapter 6.

Analog Frame Camera and Film (silver halide crystals)

Negative: reversal of tone and geometry

Photograph

lens

a.

Digital Frame Camera Area Arrays

NIR
Red
Green
Blue

Detectors

Lens and filtration

b.

Scanner

Scanning mirror

NIR
Red
Green
Blue

Discrete detectors

c.

Linear Array "Pushbroom"

Detectors

Objective

d.

Linear Array "Whiskbroom"

Rotating mirror

NIR
Red
Green
Blue

n bands

e.

Hyperspectral Area Array

NIR

n bands

Blue

Dispersing element

Detectors

f.

Figure 2-3: Illustration of different passive remote sensing systems used to collect aerial photography, multispectral and hyperspectral imagery (Jensen, 2007).

2.2.2 Active remote sensing systems

Active remote sensing systems refer to any sensing technology which emits its own measuring wave energy. The energy pulse must travel in both directions and because the speed of light is known the distance to the object can be calculated using half of the roundtrip travel time[1] (Weng, 2012). Active systems do not use sunlight; therefore, they can measure at any time of day, allowing greater collection time flexibility (Dong and Chen, 2017). The most commonly

---

[1] Outside of a vacuum, atmospheric corrections must also be applied (Beraldin et al., 2010).

used active remote sensing systems are Radio Detection And Ranging (Radar) which measures using microwaves and Light Detection And Ranging (LiDAR) which uses shorter visible or near infrared wavelengths of light (Weng, 2012). The longer wavelengths used in Radar permit it to penetrate clouds, allowing reliable monitoring and guaranteed revisit times (Jensen, 2007). However, processing radar data into intelligible images is challenging, with experienced operators required for interpretation (Yumus and Ozkazanc, 2019). Radar data can be collected at different wavelengths and polarisations and is usually collected at a slanted angle in relation to the ground (Weng, 2012). Radar interferograms are generated by differentiating phase signals from two or more acquisitions; these interferograms allow small differences in topography to be observed (Jensen, 2007).

LiDAR systems use lasers, which are focused beams of coherent light (Weng, 2012). These can be used to measure the distance to an object, either by emitting a laser pulse and timing how long it takes for the pulse to reflect from the object of interest and back to the detector, or by modulating the phase of the emitted beam and measuring the phase offset on return (Beraldin et al., 2010). There are multiple types of laser ranging measurement technologies used in the surveying and mapping industries. Hardware systems include 3D terrestrial laser scanners, mobile mapping scanners and aerial LiDAR systems. Industrial processing and metrology also use laser measurement systems; these industries primarily use methods based on triangulation, structured light and interferometry (Beraldin et al., 2010). Due to the very close ranges and delicate calibrations involved with these types of measurement they are not considered further in this thesis.

Surveying and mapping grade laser measuring is carried out using either the time of flight of individual pulses or by continuous phase differencing (Beraldin et al., 2010). Time of flight systems calculate range ρ as follows in Equation (2-1) (Beraldin et al., 2010):

$$\rho = \frac{c}{n}\frac{\tau}{2} \tag{2-1}$$

Where: ρ = range, c = speed of light in a vacuum, n = refractive index of air and τ = round trip time.

As time of flight systems use discrete pulses, the entire field of view is scanned one point at a time, using a beam deflecting system to change the angle of the emitted pulse (Beraldin et al., 2010). Traditionally, time of flight instruments have been slower than their phase based counterparts, with the number of points per second limited by the fact that another pulse cannot be emitted until the previous echo arrives (Beraldin et al., 2010). However, the new Leica P-Series scanners can record up to 1 million points a second, by combining time of flight principles with full waveform digitising (Walsh, 2015).

Phase differencing systems measure by modulating the emitted incoherent laser light, then measuring the phase difference between the emitted and received waveforms as a time delay. The waveform can be modulated by sinusoidal modulation, amplitude modulation using phase difference, frequency modulation using beat frequencies, phase coded compression and chaotic LiDAR based on chaotic waveforms from a semiconductor laser (Beraldin et al., 2010). The phase difference can be related to the time delay t using Equation (2-2) (Beraldin et al., 2010):

$$t = \frac{\Delta\phi}{2\pi.f_{modulated}} \qquad (2\text{-}2)$$

Where $\Delta\varphi$ = phase difference and $f_{modulated}$ = modulation frequency

Time delay t can then be related to the distance equation where the range ρ is given by Equation (2-3):

$$\rho = \frac{(c.t)}{2} = \frac{c}{4.\pi}.\frac{\Delta\phi}{f_{modulated}} \qquad (2\text{-}3)$$

Time of flight systems have a longer range, as it is difficult to generate the continuous waves required for the phase differencing method at a high enough energy over long distances. The phase differencing method has historically provided higher accuracies at shorter ranges and faster collection speeds (Beraldin et al., 2010), although results from González-Jorge et al. (2018) indicate that the measurement method used in the newer generation of scanners is no longer a primary determiner of accuracy.

Laser measuring systems which are designed to be mounted on static platforms are known as Terrestrial Laser Scanners (TLS). These surveying scanners

generate a full 360° scan by deflecting the laser beam along both the vertical axis (using a rotating mirror) and the horizontal axis (by rotating the entire instrument) (Beraldin et al., 2010). A schematic of a Leica C10 TLS is shown in Figure 2-4.
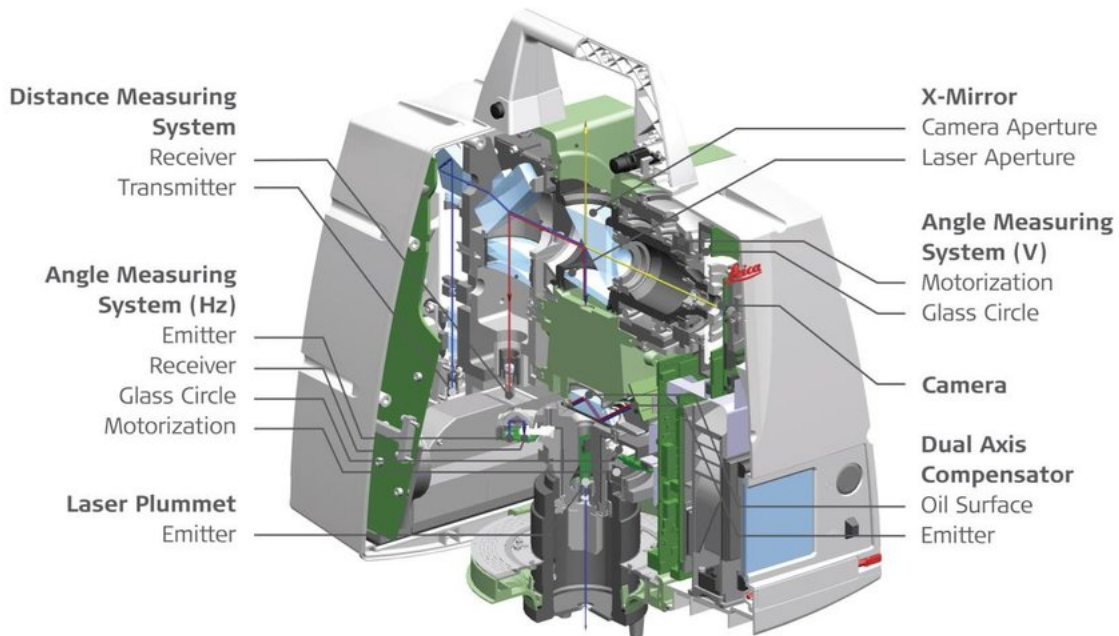
Figure 2-4: Schematic of a Leica C10 terrestrial laser scanner (Walsh, 2016).

Laser measuring systems can also be mounted on moving platforms such as vehicles. These systems are used for both mobile mapping applications and as an environment sensor for autonomous driving applications. The laser scanner typically only rotates on one axis (a 2D scanner), with the third dimension provided by the vehicle's forward motion (Puente et al., 2013). Systems designed for mobile mapping generally have two laser scanners mounted at the back of the vehicle for a more detailed field of view, while systems designed for autonomous vehicles generally have a primary laser scanner mounted at the highest point on the vehicle, as shown in Figure 2-5.

To measure ranges from a moving platform, the laser scanner must be precisely coupled to other sensors to determine the platform's precise location at all points in time (Puente et al., 2013). These sensors can include Global Navigation Satellite System (GNSS) sensors for absolute positioning, Inertial Measuring Unit (IMU) sensors for orientation and motion detection and distance measuring instruments (DMI) as a check on wheeled systems (Puente et al., 2013). These sensors are combined using a Kalman filter (Kalman, 1960) to

provide an estimation of the sensing platform's location. Once the location of the laser scanner's centre is known, the ranges can be calculated using the relevant equations (2-1) or (2-2 & 2-3) (Beraldin et al., 2010). For a thorough review of modern mobile mapping systems see Puente et al. (2013).
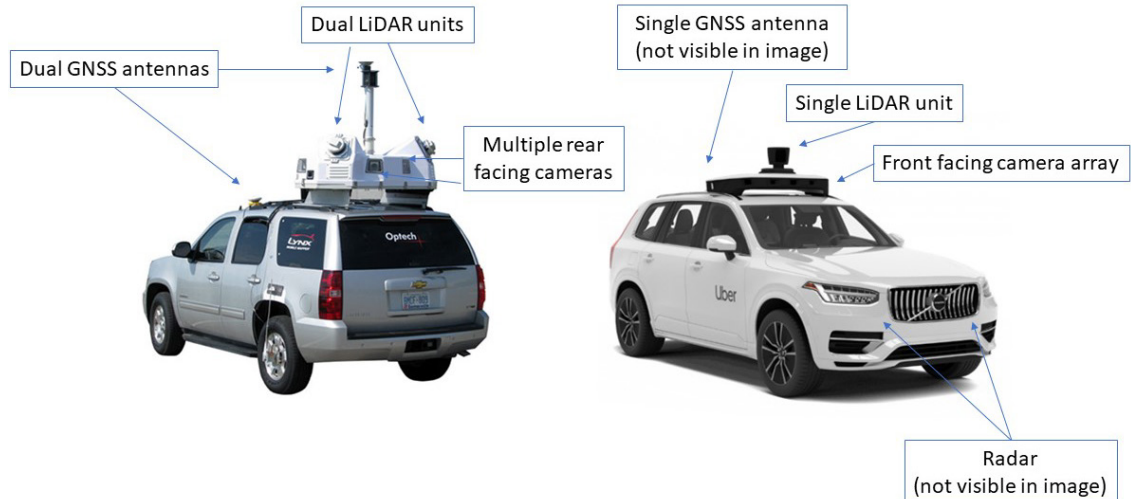
Figure 2-5: Mobile mapping (left) and autonomous driving (right) mobile laser scanning systems. Base images (Teledyne Optech, 2020 *and* Korosec, 2019)*.*

Laser measuring systems can also be mounted on aircraft to survey terrain from above. These systems are known as aerial LiDAR (Dong and Chen, 2017). The measurement principals are similar to the vehicle mounted mobile mapping systems, comprising of a 2D laser scanner moving forwards on a platform whose position is determined from GNSS and IMU sensors (Beraldin et al., 2010). Aerial LiDAR systems almost always use the time of flight measurement techniques due to the longer distances involved. Unlike most TLS systems, aerial LiDAR systems generally record multiple returns or even the full waveforms of the returned laser pulse (Beraldin et al., 2010). This allows measurement of more complex land cover such as forests as returns can be recorded for both the canopy and the ground. The differences between waveform recording systems are shown in Figure 2-6.
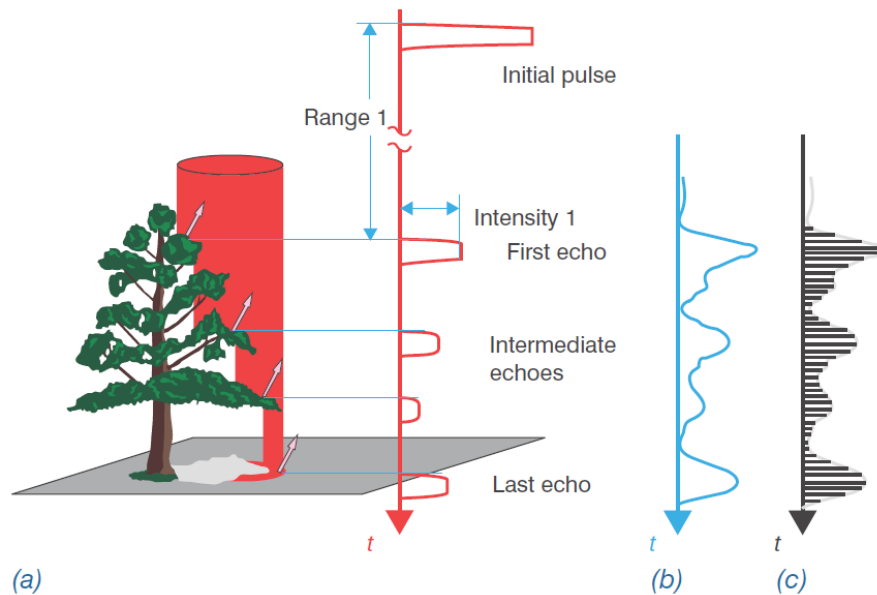
Figure 2-6: Illustration of the differences between multiple discrete pulses (a), full waveform (b) and digitised waveform (c) (Beraldin et al., 2010).

Along with the range from the scanner to the object, laser scanning systems also record the intensity of the returned pulse. Intensity values provide a valuable additional visual dimension to laser scan data. However, intensity values can pose difficulties for interpretation as they can be influenced by many factors, including distance from the scanner, incidence angle, wetness of surface and roughness of surface (Xu et al., 2018).

In order to develop applications across the wide range of sensing systems and scales, this thesis examines both active and passive remote sensing systems based on terrestrial, aerial and orbital platforms.

2.3 Data structures

The passive and active remote sensing technologies described in Section 2.2 generate large volumes of spatial data in differing structures depending on the sensing technology. The data structure is the greatest determiner of the type of deep learning that can be performed; therefore, it is essential to understand these data structures and to know which sensors are capable of generating which structures. Broadly, the data structures generated from remote sensing technologies can be divided into 2D images, structured 2.5D height or depth grids, structured 3D data and full unstructured 3D point clouds.

2.3.1 2D image data structures

The primary output data type from passive remote sensing systems are digital images, which are grids of stored numbers, where each number/location pair is called a pixel (Weng, 2012). This data is considered structured data, as it exists in a regular matrix structure with no irregular spaces. A monochrome image contains only one brightness value per pixel and takes the form of a *H x W x 1* matrix where *H* is the image height in pixels and *W* is the image width. Images with only one value per pixel location are single channel images. A true or natural colour composite image has three channels, corresponding to the red, green and blue (RGB) bands of the visible light spectrum, as shown in Figure 2-7. It takes the form of a *H x W x 3* matrix (Jensen, 2007). This same three channel structure can be used to store false colour composite images, where the near infrared band is mapped to the red channel, red light is mapped to the green channel and green light is mapped to the blue channel (Weng, 2012).
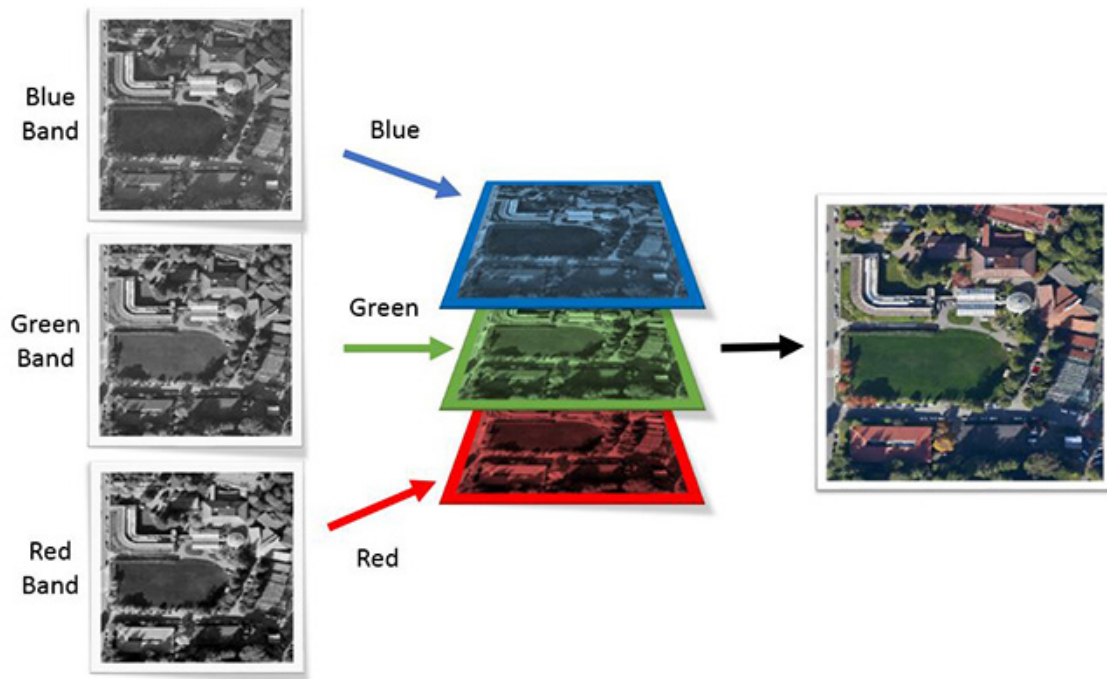
Figure 2-7: Example of a three band true colour image (Humboldt State University, 2019).

Images with more than three bands are not easily visualised by human eyes or computer monitors; however, arbitrary numbers of bands can be stored digitally in a *H x W x N* matrix where *N* is the number of channels (Jacobson and Gupta, 2005). These images are generally known as multispectral images if they contain 3 - 20 discrete channels, and hyperspectral images if they have many more channels, usually numbering in the hundreds and the sensor measures continuous spectral ranges without gaps (Giannoni et al., 2018). Hyperspectral image matrices are known as hyperspectral cubes (Lillesand et al., 2015). Whilst multispectral and hyperspectral data structures have many spectral dimensions, they can be classed as 2D data structures as they measure the scene in only the x and y spatial dimensions. Figure 2-8 illustrates the differences between multispectral and hyperspectral data types.
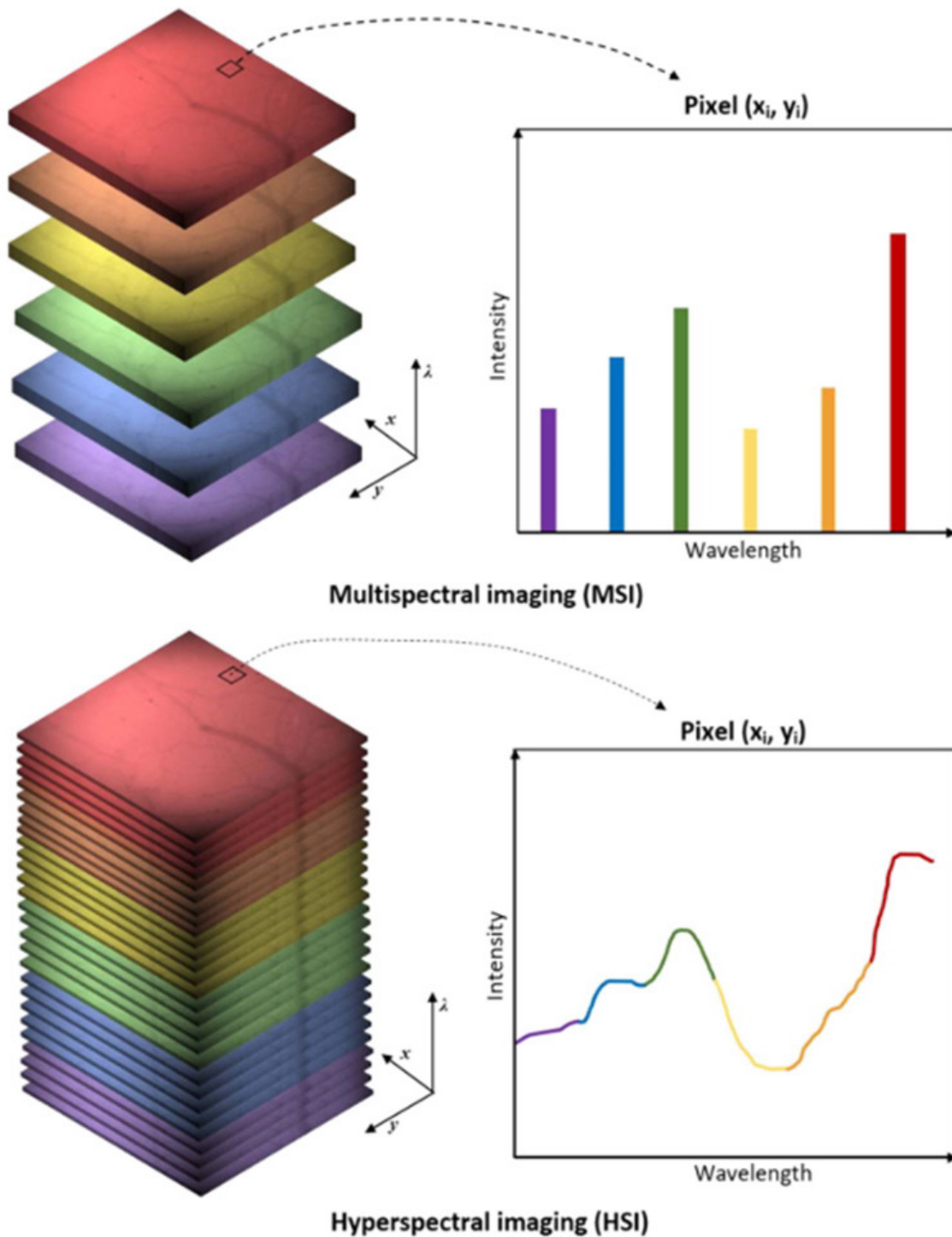
Figure 2-8: Difference between multispectral and hyperspectral data (Giannoni et al., 2018).

2.3.2 2.5D and 3D data structures

Unlike images, 3D datasets can have many different structures, primarily related to the sensing technology used for acquisition and also related to the desired final dataset function and whether or not it is true 3D data or 2.5D data.

2.3.2.1 Indirect 3D

The image data structures described in 2.3.1 measure the scene in two spatial dimensions, with no height or depth information.[2] This third spatial dimension can be generated either using the principles of photogrammetry, described in multiple textbooks (Lillesand et al., 2015; Linder, 2014) or by active remote sensing technologies such as LiDAR (Vosselman and Maas, 2010). With the exception of determining limited height information from the displacement of tall objects in aerial scenes, adding a third dimension from images requires obtaining multiple overlapping images of the scene (Weng, 2012).

The simplest method to generate 2.5D or 3D information is stereo photogrammetry. Stereo camera systems construct a model of the environment using a known baseline between two camera centres to reconstruct geometry (Luhmann et al., 2014). This is achieved by measuring the offset between left and right images (Luhmann et al., 2014). Stereo imaging systems are computationally simple and can be built using very small cameras, which makes them ideal for mobile robotics applications such as NASA's Curiosity rover (Grotzinger et al., 2012). Data generated from stereo vision systems is chiefly structured as 2.5D data, as there is only one third dimensional value recorded per pixel (Luhmann et al., 2014). 2.5D data maintains a structured grid format, with the height information saved either as an additional channel or separated out into a supplementary single channel image.

Indirect 3D information can also be obtained from combining multiple images. Downwards facing aerial photogrammetry uses a technique known as bundle block adjustment described in Aber et al. (2010), while multi-view terrestrial and drone based photogrammetry uses a technique known as Structure from Motion (SfM) described in a geoscience context by Carrivick et al. (2016). Conceptually these techniques are similar, where matches are made between points common to multiple images. These matched points are first used to locate the camera positions in space, before using the changes in camera position and the resulting shifts in image coordinates to reconstruct the geometry of the scene (Lillesand et al., 2015). SfM approaches need more computing power and

---

[2] By convention, height refers to adding another scene dimension to downwards facing airborne or spaceborne imagery, while depth refers to adding another scene dimension to forward facing terrestrial imagery.

higher resolution sensors than stereo vision systems and are popular for drone based photogrammetric surveys (Carrivick et al., 2016).

Of the indirect 3D techniques, only SfM is commonly used to produce true 3D data, where multiple third dimensional values are possible at every 2D location point. True 3D data such as this is most often represented as point clouds rather than gridded structures, these are described in more detail in the following section. Figure 2-9 helps to illustrate how SfM algorithms work, by showing the multiple 2D images that are oriented and projected to create a 3D model.
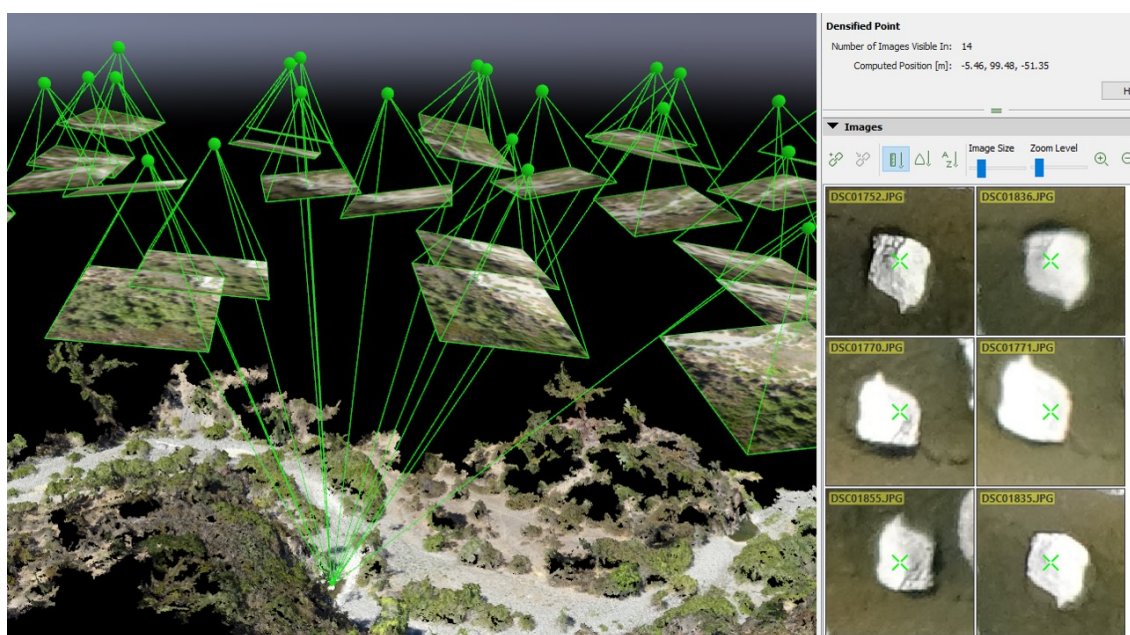
Figure 2-9: Screenshot from SfM software package showing how a model is built up from multiple 2D views of the same object (Humboldt State University, 2019).

2.3.2.2 Direct 3D

The primary data output from laser based active remote sensing systems is 3D point clouds. These are unordered sets of 3D coordinates, containing X, Y and Z values (Vosselman and Klein, 2010). This data format is considered true 3D as a single X, Y location can have multiple Z values, allowing tunnels and bridges to be recorded. An important property of 3D point clouds is that they are permutation invariant (Cherabier et al., 2016); therefore, the order of the points in the file does not change the point cloud itself. Alongside the coordinate information, other data can also be stored for each point, such as the intensity of the laser return and RGB information if the laser scanner has an integrated camera (Vosselman and Klein, 2010). Intensity data can add valuable

40

information about the scene, particularly underground where dust and uneven lighting can cause difficulties with image interpretation.

3D point clouds are unstructured data, which can prove challenging for computers to store and process (Vo et al., 2015; Xie et al., 2020). Point clouds can be converted to a structured format by voxelization, where 3D pixels known as voxels (Foley et al., 1990) are used to store the data representation. Voxelization simplifies the data by adding uniform structure but can also increase the dataset size (Gebhardt et al., 2009). Point clouds are by nature a sparse data representation. If using a gridded voxelization method the majority of the voxels will be empty; however, empty voxels take up the same amount of storage space as voxels which encode details of the scene (Gebhardt et al., 2009). Variations such as sparse voxelization and octree representations can mitigate this issue (Gebhardt et al., 2009; Xie et al., 2020).

The point clouds acquired from airborne LiDAR systems do not generally contain multiple Z values other than those related to multiple return systems, as discussed in Section 2.2.2. This is because the sensing platform has a fixed downwards facing field of view. This characteristic allows these point clouds to be converted to a simple and memory efficient 2.5D gridded structure known as a Digital Elevation Model (DEM) (Briese, 2010). DEMs can be generated using only the returns most likely to be ground points (by taking the last recorded laser return alongside other post-processing rules) to create a bare earth Digital Terrain Model (DTM) or they can be generated using the first return to creating a Digital Surface Model (DSM) (Briese, 2010). DSMs contain all ground surface objects regardless of category.

2.3.3 Data structures summary

The structure of the data is the primary factor for choosing an appropriate machine learning algorithm, with different data structures being more or less suited as input to different types of machine learning models. Because the data structure has such a powerful effect on the algorithm choice, the case studies in this thesis are organised around each of the three core data types: true 3D (Chapter 3), structured 2.5D (Chapter 4) and multi-band 2D (Chapter 5).

2.4 Deep learning

Machine learning is the discipline of computer science concerned with teaching computers to learn without explicitly programming them (Samuel, 1959). Instead of hard coding rules to enable computers to solve problems, machine learning allows the computer to acquire its own knowledge about the problem by extracting patterns from the input data (Goodfellow et al., 2016). The success of classic machine learning algorithms depends on the representation of the input data, known as features. Handcrafting these features allows for high accuracy results; however, the process is time consuming and relies on expert human knowledge combined with trial and error (Goodfellow et al., 2016). Deep learning is a subset of machine learning and examines how an algorithm can learn both the complex representations and the underlying data patterns. This is achieved by using many connected simple representations to model high level abstract features (Goodfellow et al., 2016). Figure 2-10 shows how rule-based systems, classic machine learning and deep learning relate to each other.
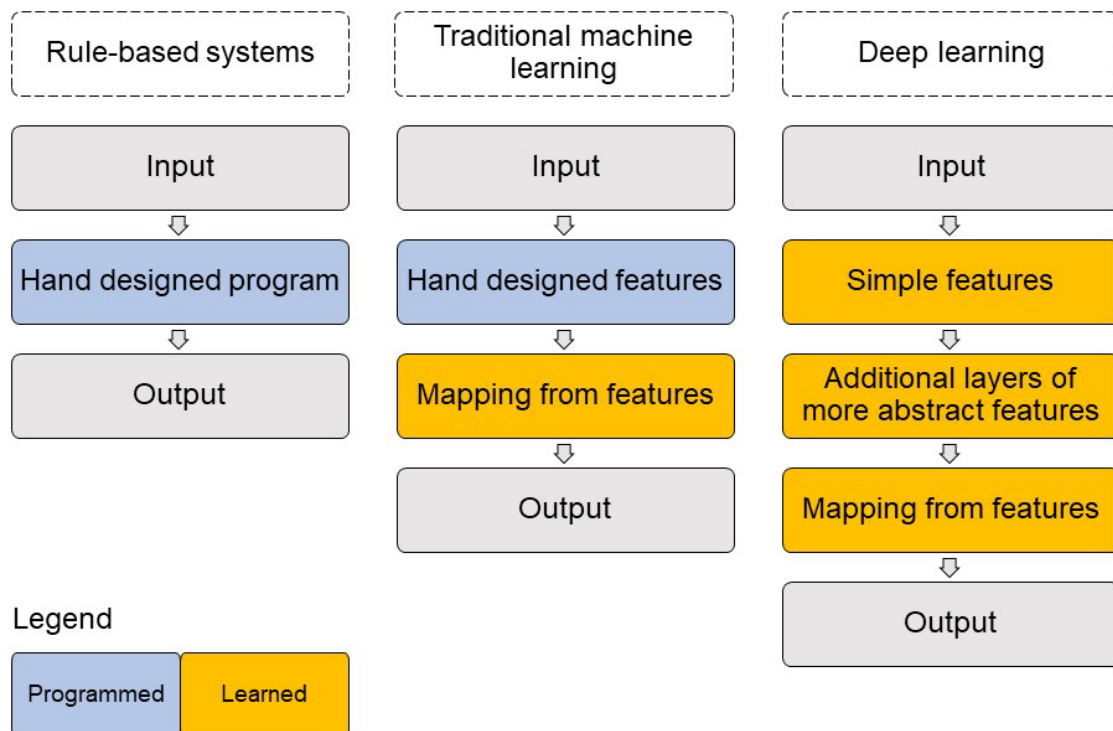
Figure 2-10: Differences between rule-based systems, classic machine learning and deep learning. Orange boxes indicate components that learn from data without human guidance. Image adapted from (Goodfellow et al., 2016)*.*

Machine learning systems can learn either from labelled examples or by determining patterns in unlabelled data; these techniques are known respectively as supervised and unsupervised learning. Supervised learning

methods can be divided into classification (if the result to be determined is categorical) or regression (if the result to be determined is continuous). Supervised learning is the most common method of machine learning (LeCun et al., 2015) and is the method explored by the different algorithms used in this thesis. Supervised learning can encompass both classic machine learning algorithms such as random forests (Breiman, 2001) and advanced modern deep learning algorithms such as convolutional neural networks. Unsupervised learning methods encompass clustering algorithms and dimensionality reduction algorithms such as principal component analysis. Other methods of learning include semi-supervised learning, where a small amount of labelled data is used in conjunction with large amounts of unlabelled data (Chapelle and Zien, 2005) and reinforcement learning, where autonomous machine agents learn the parameters of a task via trial and error (Kaelbling et al., 1996).

2.4.1 History of deep learning

The theories which underpin modern deep learning can be traced back to the 1940s where computational models to mimic biological learning were developed (McCulloch and Pitts, 1943), evolving into the perceptron (Rosenblatt, 1958). A perceptron is the building block of an artificial neural network, also known as a Multi-Layer Perceptron (MLP) and was the first model which could learn its own weights w for the function $f(x,w) = x_1w_1 + \ldots + x_nw_n$ (Goodfellow et al., 2016). However, a single perceptron could only learn linearly separable functions and computing technology in the 1960s was insufficient to iterate across multiple perceptrons. These shortcomings were publicised in Minsky & Papert (1969), leading to a slowdown in neural network research until the 1980s. At this time, the emerging field of parallel distributed processing (Rumelhart and McClelland, 1986) generated a resurgence in interest, alongside the revisiting by Rumelhart et al. (1986) of the technique of backpropagation for training neural networks. A model type known as a convolutional neural network, designed for image processing tasks also emerged in this period (Fukushima, 1988; LeCun et al., 1989a). Backpropagation provided an answer to the question of how to train a deep neural network; however, it was still computationally too expensive for the hardware available at the time, leading to the true capabilities of the algorithms developed in the 1980s to remain untapped for nearly twenty years (Goodfellow et al., 2016).

Modern deep learning can be considered to have emerged with a paper by Hinton et al. (2006) which demonstrated that a deep belief network could be trained effectively by a strategy known as greedy layer-wise pre-training (Goodfellow et al., 2016). Bengio & LeCun (2007) discussed how this strategy was successful on many types of deep neural networks and emphasised the importance of depth for generalisation on complex artificial intelligence tasks. The modern successes of deep learning can be attributed both to these improved architectures and to the increases in training dataset size and the sophistication of computer infrastructures (Goodfellow et al., 2016). Training dataset sizes have increased from 60,000 for the MNIST dataset (Lecun et al., 1998) to over 14 million for the ImageNet dataset (Deng et al., 2009). The invention of Graphical Processing Units (GPUs) in the early 2000s allowed far faster parallelized processing of simple matrix operations, leading to a significant increase in the speed, accuracy and usability of deep learning models (Cireşan et al., 2010; Raina et al., 2009).

2.4.2 Multi-layer perceptron

Before describing modern deep learning architectures such as convolutional neural networks it is beneficial to understand the foundational artificial neural network known as a multi-layer perceptron (MLP).[3] Neural networks can be thought of as function approximators; however, a single neuron can solve problems only where the solutions are linearly separable (Goodfellow et al., 2016). This limitation is removed when multiple neurons are connected together, allowing extremely complex functions to be modelled. Indeed, the universal approximation theorem (Hornik, 1991) states that a feed-forward neural network with only a single hidden layer containing a finite number of neurons can approximate arbitrary continuous functions. Any neural network containing more than two hidden layers can technically be considered 'deep'; however, it is more common to use the term 'deep learning' to refer to the

---

[3] Whilst these networks are named multi-layer perceptrons, the neurons used are normally sigmoid neurons rather than perceptron neurons. This is because they use a function with a calculable slope instead of the step function used in the original perceptron research (Nielsen, 2015). In this research the terms Fully Connected Neural Network and Multi-Layer Perceptron are used interchangeably to refer to modern implementations of the traditional fully connected neural network architecture regardless of activation function used.

modern architectures which are more complex than a simple fully connected neural network.

2.4.2.1 Structure

A schematic of a single neuron is shown in Figure 2-11. It contains multiple inputs connected via multiplicative weights to an activation function which outputs a result. The activation function takes the summation of the weighted inputs and transforms it to a fixed range, usually between 0 and 1 (Géron, 2017). Each neuron also contains a bias or threshold term. This is a real number which determines how sensitive the neuron is. The magnitude of the weights determines the relative importance each input plays in producing the final result. The very first artificial neuron (perceptron) used a step function for the activation function (Rosenblatt, 1958), which would only output binary values, creating difficulties for training. This has now been replaced by functions with differentiable gradients such as the sigmoid, tan-h and Rectified Linear Unit (ReLU) functions.
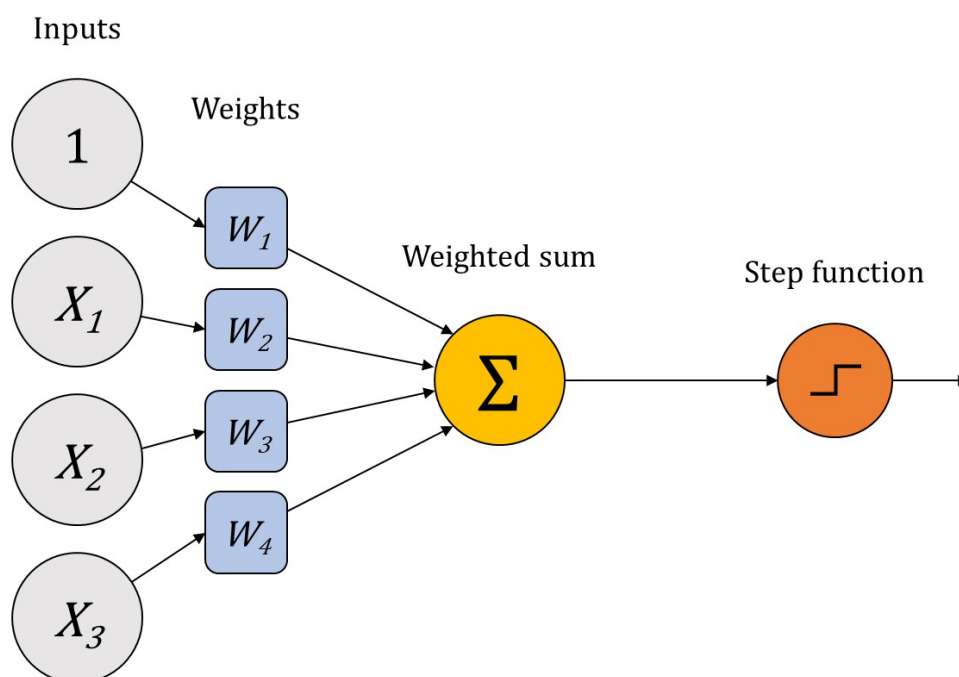
Figure 2-11: Structure of a single perceptron.

Moving on from a single neuron, an MLP is made up of multiple linked neurons arranged in layers. MLPs contain an input layer, an output layer and one or more hidden layers. A simple MLP is a fully connected architecture, where every neuron in each layer is connected to every neuron in the layers before and after it, as shown in Figure 2-12.

45
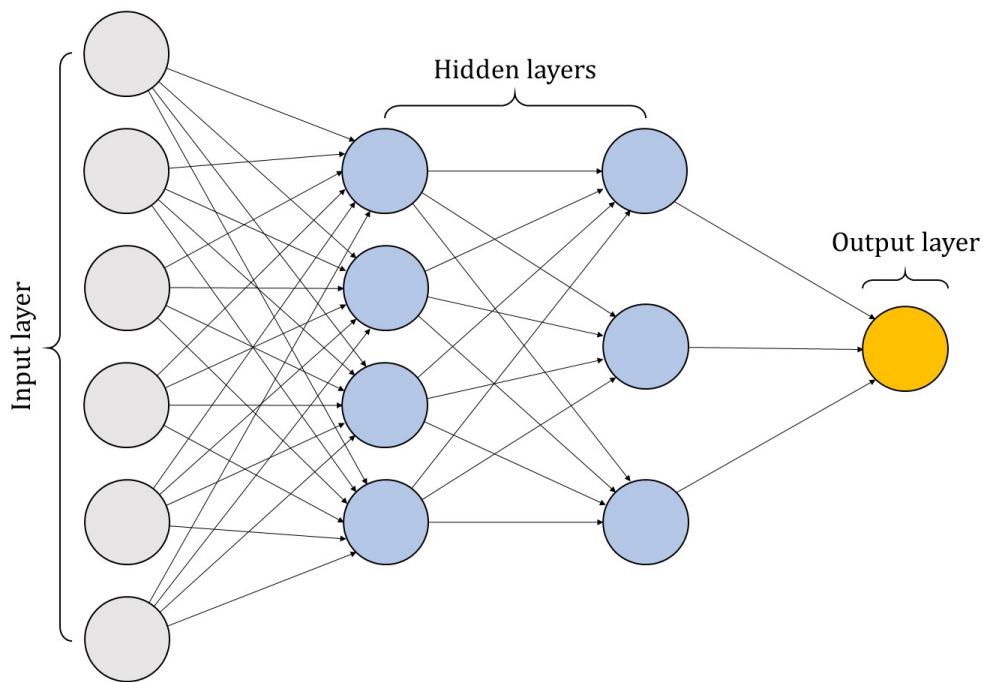
Figure 2-12: Structure of a fully connected artificial neural network

2.4.2.2 Training

A simple MLP is a supervised machine learning model which learns from labelled training examples. At the start of training, the weights for each neuron are randomly initialised and when the first training example is input to the model the resulting output will be incorrect, as the model has yet to learn to approximate the relationships between the inputs and the outputs. The model's performance can be measured using a cost function, a simple example of which is the quadratic cost function, as described by Nielsen (2015) in Equation (2-4):

$$C(w,b) = \frac{1}{2n}\sum_{x}\|y(x) - a\|^2$$

(2-4)

where *w* is the weights, *b* is the biases, *n* is the number of training inputs, *a* is the vector of outputs and *x* is the vector of inputs. It can be seen that the cost function becomes smallest when the computed function *y(x)* is closest to the labelled output *a* (Nielsen, 2015).

To minimise this cost, a mathematical technique called gradient descent by backpropagation is used to obtain the gradient of the cost function *C* and then move the values of *w* and *b* a small step in the downhill direction, incrementally decreasing *C*. The gradient of the cost function is computed by backpropagation (Rumelhart et al., 1986) which enables the partial derivatives of every weight to

46

be calculated using the chain rule, one layer at a time starting with the final layer and working backwards. Nielsen (2015) gives a clearly worked example of the mathematical derivation of gradient descent via backpropagation and Goodfellow et al. (2016) provide further derivation details. The magnitude of the downhill step taken at each iteration is known as the learning rate and is conventionally denoted as α.

Equation 2-4 shows the cost function as an average of the costs of each individual training example. One way to achieve this is to compute the gradient for every training example and then compute the average to determine the update to the weights; however, this can be very slow when the number of training examples is high and the network will take a very long time to learn. A more efficient method is to take randomly extracted mini-batches of the training data and compute the gradients over these, with the assumption that the mini-batch is reasonably representative of the entire training dataset. The model weights are then updated after each mini-batch has been calculated and the model is able to learn faster, without increasing the learning rate parameter α. This technique is known as stochastic gradient descent. Section 2.4.6 describes the practical aspects of training a deep learning model in more detail.

2.4.2.3 Towards deep learning

The type of fully connected neural network described above works well when the data inputs (features) are either naturally well suited to statistical machine learning or when adequate human generated feature extractors have been designed. Simple MLP models have been shown to outperform other classic machine learning algorithms such as random forests and support vector machines in applications such as generating landslide susceptibility maps (Tien Bui et al., 2016), detecting clouds in images (Taravat et al., 2015), classifying volcano-seismic events (Titos et al., 2018) and recognising document sentiments (Moraes et al., 2013). However, the differences in performance between the MLP and the other machine learning models are relatively small and in multiple studies the MLP model did not achieve the highest performance (Caruana and Niculescu-Mizil, 2006, Balabin et al., 2010, Weinmann et al., 2015). No simple MLP displays the paradigm shifting accuracies achieved by modern deep learning architectures such as Krizhevsky et al. (2012) and Hinton et al. (2012).

The success of models such as AlexNet (Krizhevsky et al., 2012) is due to their architecture, which allows them to leverage the spatial information present in structured data such as images. These architectures are known as Convolutional Neural Networks (CNNs). LeCun et al. (2015) define four key ideas that allow CNNs to leverage the properties of natural patterns: local connections, shared weights, pooling, and the use of many layers. By using spatial connectivity to model local pixel connections it is possible to share the model's weights across neighbouring patches, eliminating the need for every pixel and neuron to be fully connected. Consequently, model size is greatly reduced allowing very deep models to be trained. Other modern deep learning architectures include recurrent neural networks, recursive neural networks, deep generative models and structured probabilistic models. These types of deep learning are not examined in this research; see Goodfellow et al. (2016) for more details on these types of models.

2.4.3 Convolutional neural networks

The deep learning models used in this research are all variants of CNNs. CNNs are loosely inspired by the mammal visual cortex (Hubel and Wiesel, 1959, Hubel and Wiesel, 1968) where neurons have a small local receptive field, only reacting to patterns within limited regions of the visual field. Different collections of neurons react to different patterns and different sizes of receptive field. Larger receptive fields react to more complex patterns which are combinations of the lower level patterns (Géron, 2017). This biological architecture inspired Fukushima (1988) to create the Neocognitron, refined by Lecun et al. (1998) into the first recognisable CNN model, LeNet-5.

In a CNN, the local receptive field can be thought of as a small sliding window which moves over the entire input image one stride at a time. Each region of pixels covered by the local receptive field at each time is connected to a single neuron in the first hidden layer, as shown in Figure 2-13. Each neuron in the first hidden layer uses the same weights and bias from the local receptive field; therefore, the first hidden layer will show activations in the presence of the same pattern, just in different positions in the input image. This solves the problem of location invariance, i.e., a dog is still a dog whether it is the upper left or lower right corner of an image.

Figure 2-13: CNN stride and local receptive field. The yellow neurons represent the local receptive field and the stride in this example is 1. Image and caption adapted from Nielsen (2015).

This operation is also known as feature mapping, as it detects features in the input image and maps them to the hidden layers. CNNs are so named as the filtering operation used by the feature mapping is a discrete convolution[4] (LeCun et al., 2015). Each filter can detect only one type of pattern, for example a vertical edge; therefore, each convolutional layer is made up of multiple stacked feature maps designed to detect multiple types of patterns. A convolution layer increases the size of the output image along its 3rd dimension, as shown in Figure 2-14. After the convolutions, an activation function such as ReLU is applied to the feature maps to regularise the results.

---

[4] In the context of deep learning both true convolutions and cross-correlations (convolutions without the kernel flip operation) are by convention known as convolution operations (Goodfellow et al., 2016).

Input layer (16 x 16)   First hidden layer (13 x 13 x 3)
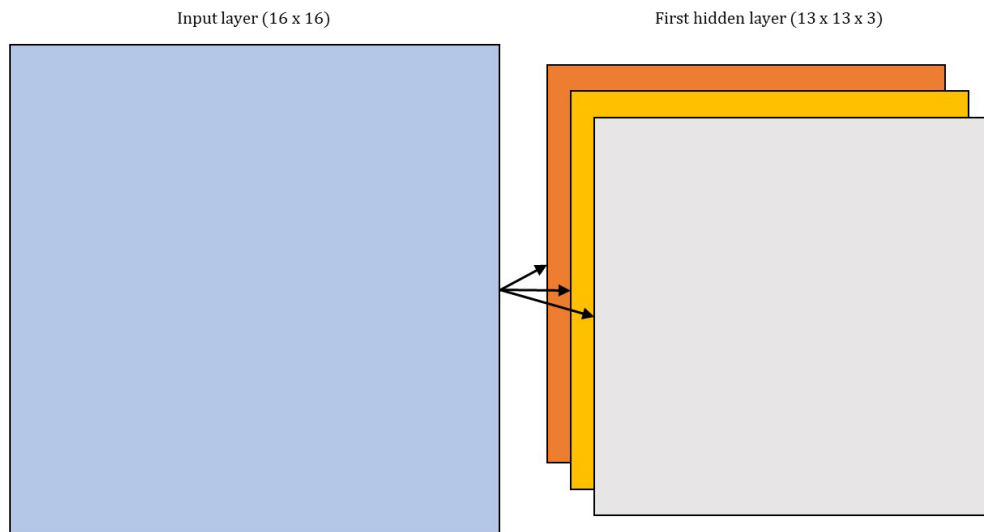
Figure 2-14: Multiple feature mapping. The process illustrated in Figure 2-13 generates a single dimension of the first hidden layer, additional filters generate additional dimensions. Image and caption adapted from Nielsen (2015).

As the convolution layer increases the input size, another operation known as pooling is used to simplify and reduce the resulting feature maps. Pooling is essentially a downsampling operation, reducing the size of the feature maps but maintaining the overall picture of what parts of the feature map have been activated. Typically pooling uses the max-pooling operation, where each group of pixels (for example in a 4 x 4 region) is replaced with the single highest value from that region, as shown in Figure 2-15.
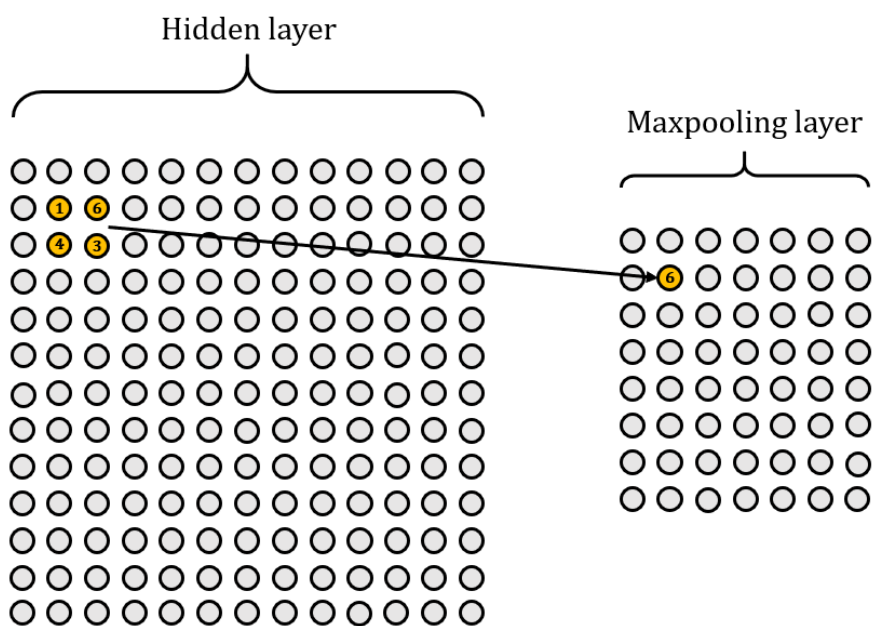
Figure 2-15: Illustration of maxpooling operation. Note: for clarity, this example shows only one of the hidden layer's multiple dimensions. Image adapted from Nielsen (2015).

Modern CNNs usually have multiple rounds of convolution and pooling layers, when visualising the stages, it is helpful to think of the images as 3D cubes, similar to the hyperspectral cubes described in Section 2.3.1. Each convolution layer increases the number of image channels and each pooling layer decreases the spatial resolution, as shown in Figure 2-16.
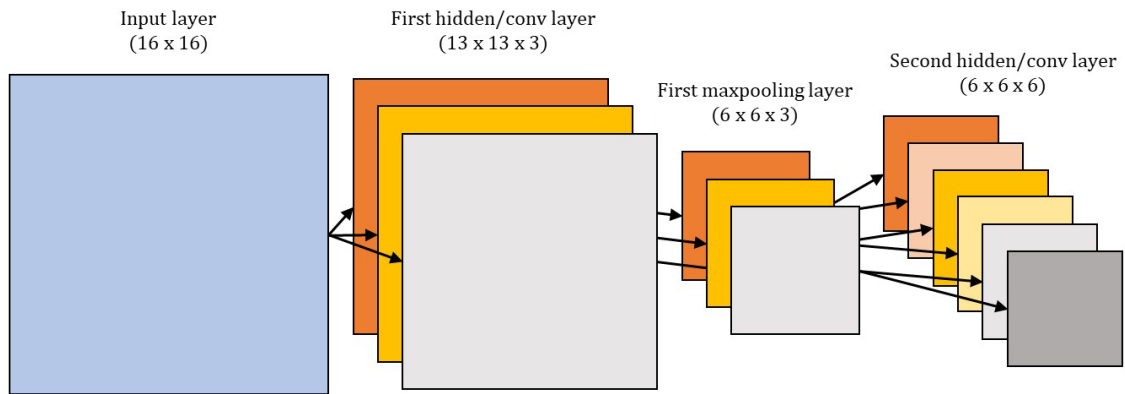
Figure 2-16: Example changes to image sizes throughout a CNN architecture. The images become spatially lower in resolution, but their number of dimensions increases as they move further into the model.

It can be seen that as we move deeper into the network the layers become spatially smaller but capable of ever more abstract representations of the input data. A fascinating paper by Zeiler & Fergus (2014) allows further understanding of how CNNs 'see'; some of the results from their deconvolution model are shown in Figure 2-17.
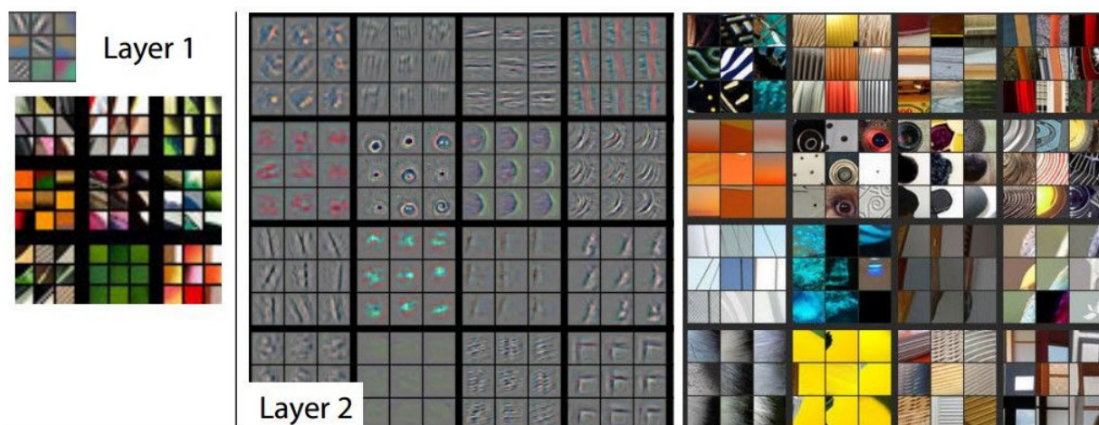
Figure 2-17: Visualisation of the features in the first two layers of a fully trained CNN, from Visualising and understanding CNNs (Zeiler and Fergus, 2014). For details on the visualisation method used to deconvolute and project the samples to pixel space see Zeiler and Fergus (2014).

2.4.4 Image processing tasks

There are five main meta-tasks which image processing CNNs are generally designed to solve. These are classification, semantic segmentation,

51

classification with localisation, object detection, and instance segmentation, illustrated in Figure 2-18. The last part of the CNN architecture depends on the type of task which it has been designed for.
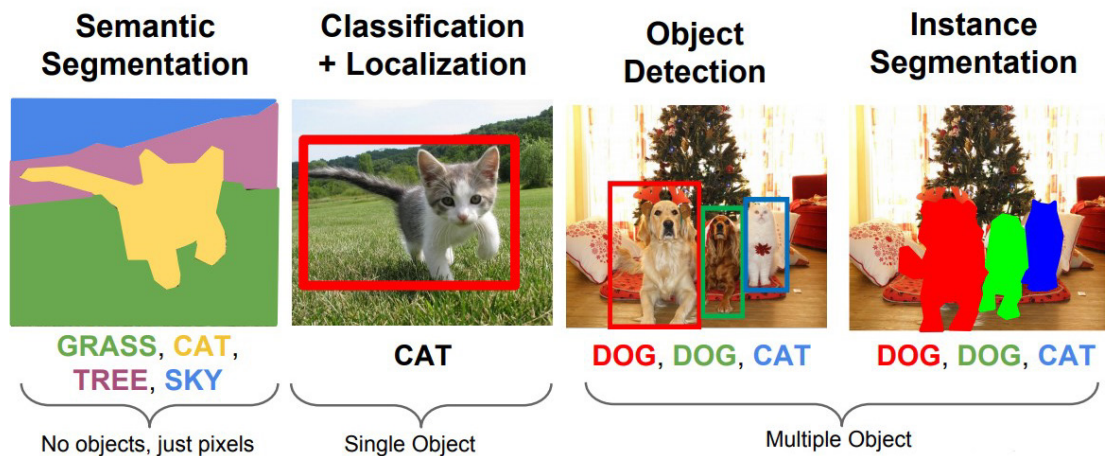
Figure 2-18: Examples of different image processing tasks (Li et al., 2020)

Classification is when the goal is to determine what type of object is depicted in the image. Classification is the simplest of the tasks, as the model does not need to know where the object is in the image, merely whether the image shows for example a cat or a dog. Generally, multiple labels are not supported in classification problems and benchmark datasets. The early ImageNet competitions were based on classification (Deng et al., 2009, Krizhevsky et al., 2012); however, they have since moved on to the more challenging tasks of object detection and localisation (Russakovsky et al., 2015).

To perform classification, the deepest convolutional and max pooling layer pair are followed by a small number of fully connected layers. These fully connected layers are identical to the simple MLP discussed above (Géron, 2017). In this way, it is possible to envision how a CNN grew from an MLP, where instead of inputting a feature representation of individual pixel brightnesses with no context (a poor feature representation), a strong feature representation is learned by the network itself. It is also possible to imagine the classic machine learning workflow in this context, where a human with expert knowledge of a task could hand design MLP input features which are more appropriate than the raw pixel values; however, this is both more time consuming and less robust than using a CNN to learn the representations itself.

Adding localisation to the classification task requires the model not only to classify the image but also to produce a single bounding box around the object.

For the classification and localisation task the assumption is made that each image only contains one object, if multiple objects are to be located it instead becomes an object detection task. This task is solved by using the same architecture as for classification, but along with the fully connected path leading to predicted category labels an additional fully connected path is used to calculate the bounding box location as a regression problem. In this architecture both paths are trained simultaneously using different loss functions, known as a multi-task loss problem.

If multiple objects are to be classified and located the task becomes known as object detection. In this task each image can have a differing number of objects of differing classes and the number of instances is unknown. There are two main approaches used to solve the object detection task, R-CNNs and single shot detectors. The original R-CNN paper by Girshick et al. (2014) proposed a method to detect objects by first detecting regions of interest using the image processing algorithm selective search (Uijlings et al., 2013) then applying classification CNN architectures to each region of interest. This method was later improved first by placing the convolutional feature mapping stage before the region proposal algorithm (Girshick, 2015) and later by replacing the selective search region proposals with a deep learning based region proposal network (Ren et al., 2017). The learning process for R-CNNs is complex, as the model must balance several multi-task loss objectives.
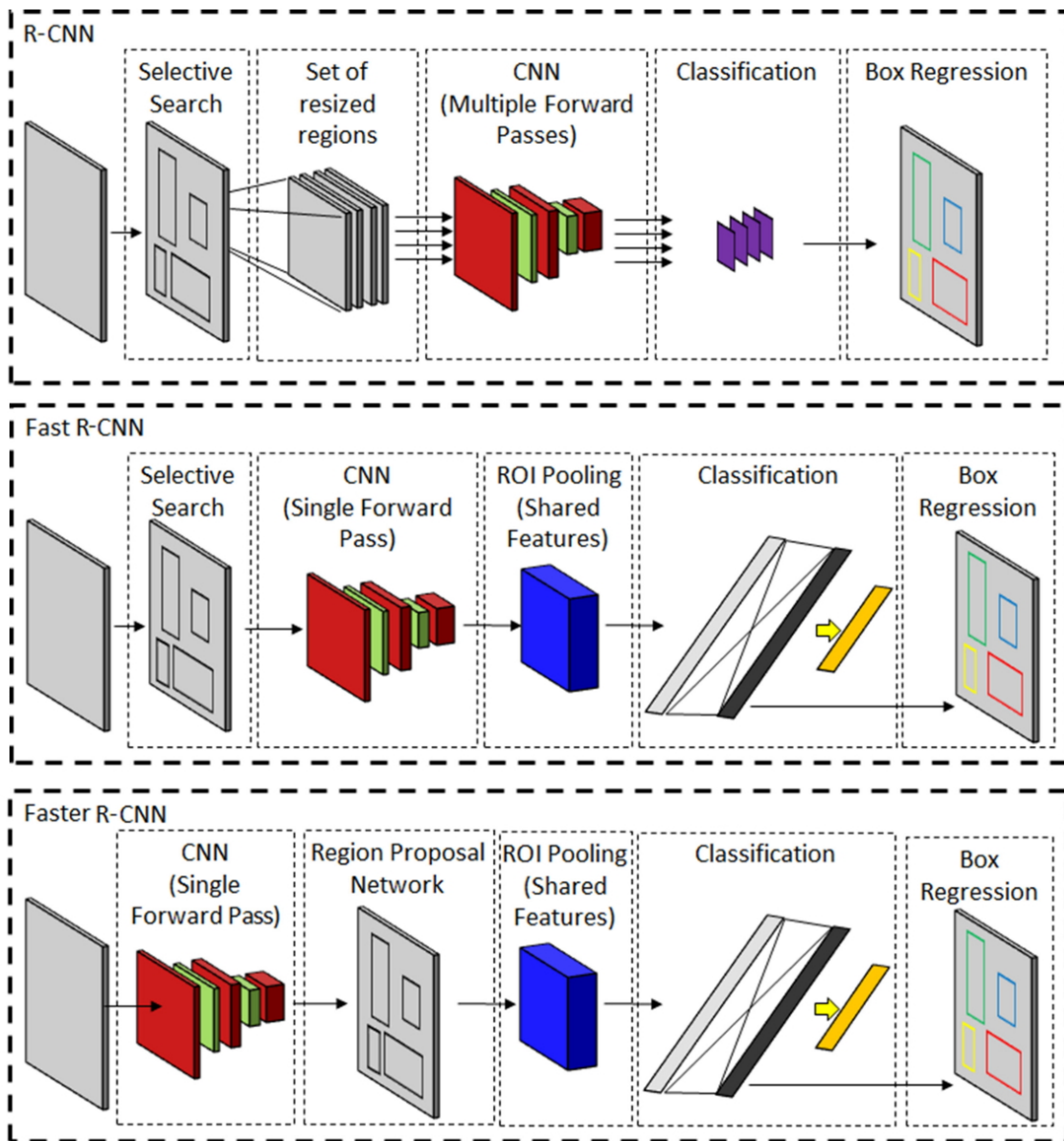
Figure 2-19: Faster R-CNN family of architectures, adapted from Ghosh et al. (2019).

Single shot detector methods do not use a region proposal network and solve for the object locations with one large feedforward network. To avoid searching near infinite potential sliding windows, these methods divide the input image into a discrete grid and solves for multiple bounding box shapes and classification confidence simultaneously (Redmon et al., 2016, Liu et al., 2016). In general, R-CNN methods are more accurate but single shot detectors are faster (Huang et al., 2017). Figure 2-19 shows the development of the R-CNN architecture and Figure 2-20 shows the principles of the single pass object detection method used by Redmon et al. (2016).
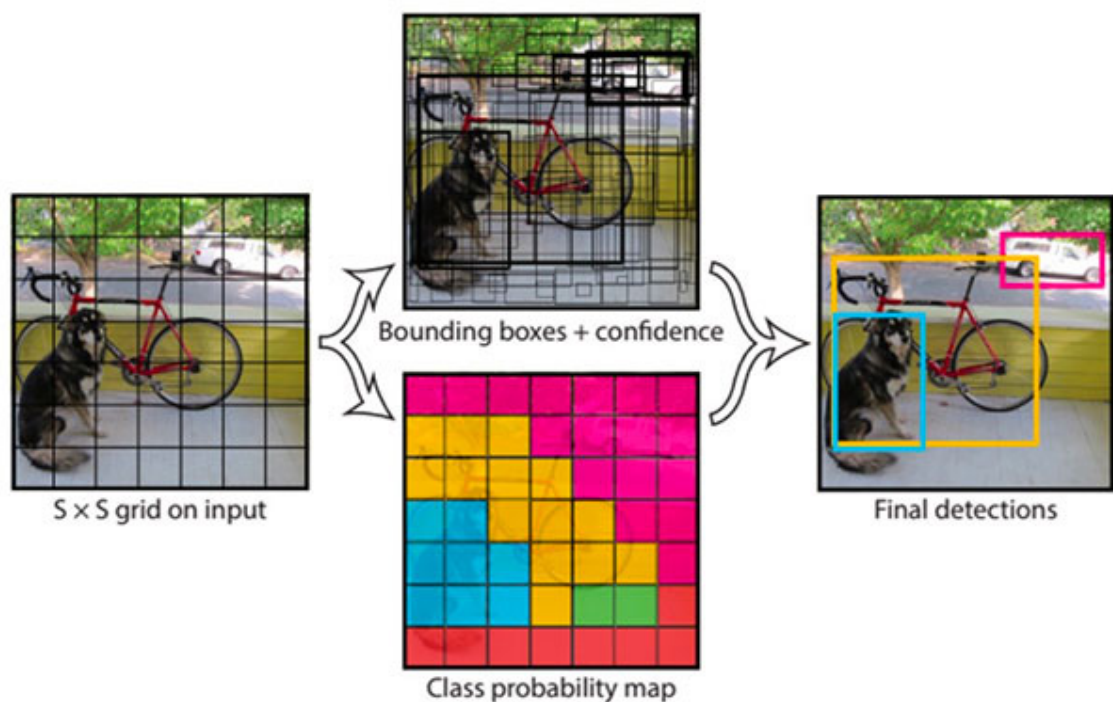
Figure 2-20: Principle of the YOLO single pass object detector (Redmon et al., 2016). *The model simultaneously learns the bounding box locations, confidence and class probabilities.*

A fundamentally different image processing task that can be solved by CNNs is semantic segmentation. This task involves determining a category label for every pixel in an image. Semantic segmentation does not differentiate objects or instances within an image, only pixel classes. Semantic segmentation could be solved using large fully convolutional networks with no pooling layers; however, computation would be costly if the images did not decrease spatially as they increase in depth. A solution is to use an architecture that progressively downsamples the input image (in the same way as a classification CNN) to capture the high level image understanding, before upsampling the deep feature maps in order to generate pixelwise labels at the same resolution as the original input image (Shelhamer et al., 2017). Cross-entropy loss is used to determine the loss across every pixel in the final layer, with no fully connected layers. The original paper by Long et al. (2015) used only one upconvolution layer, architectures such as SegNet (Badrinarayanan et al., 2017) and U-Net (Ronneberger et al., 2015) have a more symmetrical encoder-decoder structure where the number of upsampling layers is similar to the number of downsampling pooling layers. Greater precision in final resampled segmentation mask can be obtained by carrying information straight across from the corresponding downsampling layers, either by copying the pooling

55

indices (Badrinarayanan et al., 2017) or by using skip connections between corresponding resolution layers (Ronneberger et al., 2015). Figure 2-21 shows the architectures of the three discussed semantic segmentation networks, a review of over 30 semantic segmentation architectures is given in Garcia-Garcia et al. (2018).
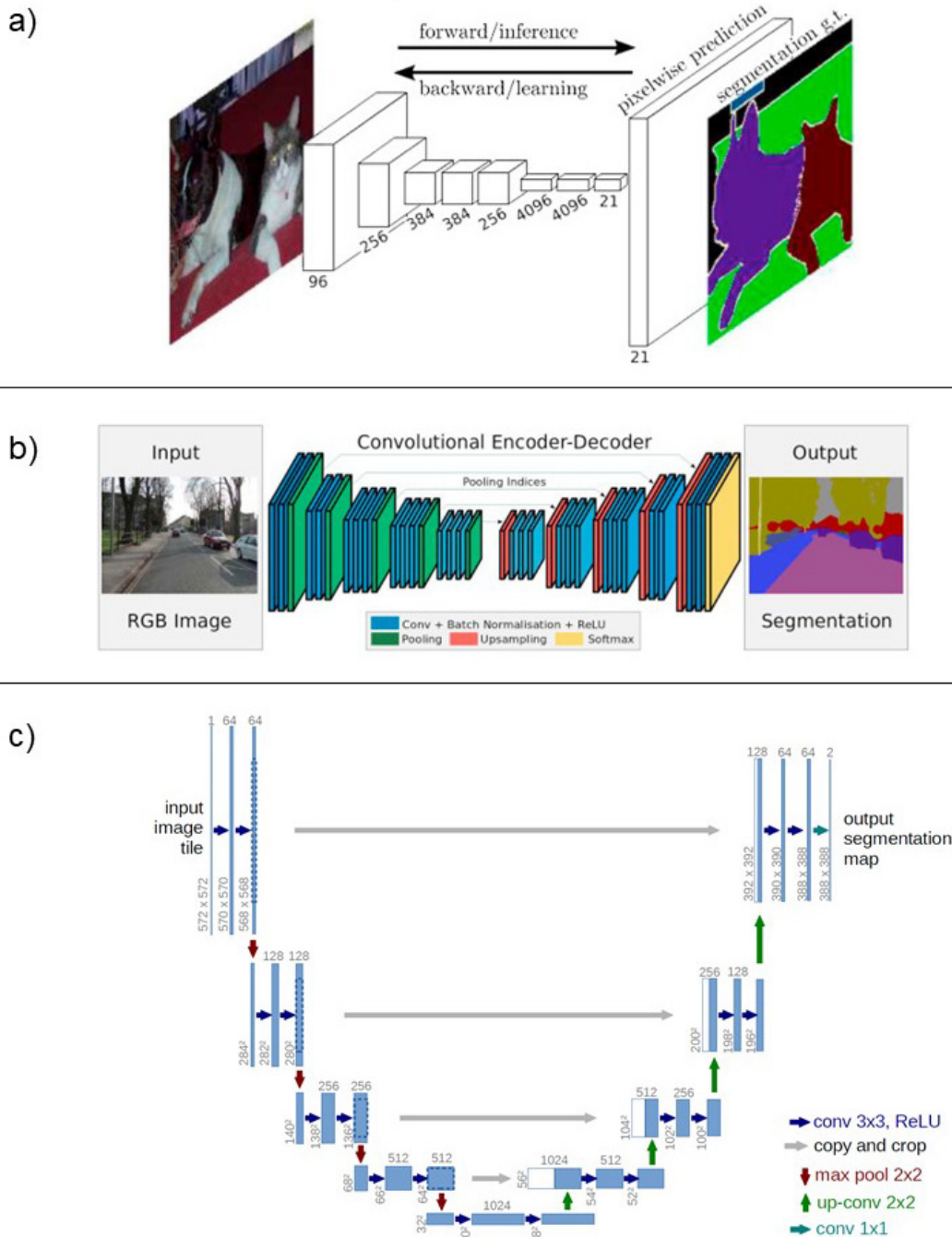
Figure 2-21: Architectures of different semantic segmentation CNNs: a) original fully convolutional network (Shelhamer et al., 2017), b) SegNet (Badrinarayanan et al., 2017) and c) U-net (Ronneberger et al., 2015).

The most complex image processing task is instance segmentation, which requires elements of all the other tasks. In instance segmentation, all individual

objects must be detected and every pixel comprising those objects must be labelled individually. He et al. (2017) proposed a modified version of Faster R-CNN (Ren et al., 2017) with an added branch which computes semantic segmentation within the regions of interest using a fully convolutional network. Solving these image processing tasks using deep learning is an extremely fast-growing field of research, with new architectures emerging rapidly across all tasks. To stay abreast of current developments it is recommended to monitor the leaderboards of the benchmark datasets such as Microsoft's COCO dataset (Lin et al., 2014) or the autonomous driving KITTI dataset (Geiger et al., 2013).

2.4.5 Deep learning for point clouds

Thus far, the deep learning methods discussed in this literature review are designed to work with images; however, the fundamental CNN building blocks can be used on 1D, 2D or 3D structured data types (LeCun et al., 2015). When dealing with 3D point clouds, the difficulty arises from their unstructured nature, as discussed in Section 2.3.2. Point clouds are not arranged in a gridded structure and are instead unordered lists of 3D cartesian coordinates; therefore, any model must be permutation invariant to address these characteristics. Additional difficulties arise from the size of the datasets, both in density of points and spatial extent. The primary tasks associated with point cloud processing are segmentation and semantic segmentation, Xie et al. (2020) give a review of the historical and state of the art approaches for solving these tasks. Segmentation involves dividing the point cloud into local objects based on geometric similarities; these methods do not usually use machine learning and limited semantic information is associated with the segmented sections (Xie et al., 2020). Segmentation algorithms can be a useful preprocessing step to improve the performance of machine learning models (Zhang et al., 2013, Vosselman et al., 2017).

Point cloud semantic segmentation can be carried out using either classic machine learning or modern deep learning. The classic approach can be divided into either point-based models or statistical context models. The point-based approach uses information about the point and its neighbours to generate a descriptive feature vector which is then used as input to some form of supervised classifier. Whilst not the first use of this method, Weinmann et al.

(2015) has become the definitive reference for this technique, due to its clear description of the methodology and thorough examination of relevant parameters. Figure 2-22 shows the main stages of this method. The advantages of this technique include a strong consideration for a point's neighbourhood and an understandable and repeatable supervised machine learning element. The disadvantages are the lack of overall context leading to noisy results, along with the time consuming nature of creating handcrafted features (Xie et al., 2020).
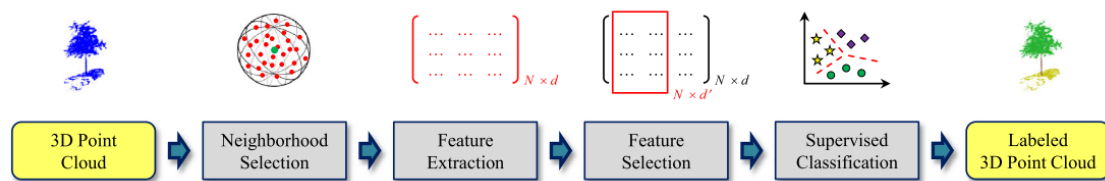
Figure 2-22: Classic point cloud semantic segmentation method (Weinmann et al., 2015a)*.*

Statistical context models address some of these issues, using Markov random fields (Geman and Geman, 1984) or conditional random fields (Lafferty et al., 2004) to model dependencies across different ranges. These methods can be employed either as a standalone method (Niemeyer et al., 2014) or applied to the results from the point-based methods (Landrieu et al., 2017a). Landrieu, Raguet, et al. (2017) showed that applying graphical models after pointwise classification improved on the accuracies obtainable by either method in isolation.

Since 2015, point cloud segmentation has attracted the interest of deep learning research (Xie et al., 2020). As described previously, point cloud data cannot be easily ingested by CNN models; therefore, the first strategies involved modifying the data rather than the fundamental CNN models. Su et al. (2015) proposed generating multiple 2D views of the 3D data to feed into a standard CNN, Boulch et al. (2018) extended this multi-view strategy to larger scenes by meshing the point cloud prior to generating the 2D views. Another method of modifying to point cloud is voxelization, described in Section 2.3.2. Once voxelized, point clouds can be processed using 3D convolutions in the same way as images are processed using 2D convolutions. VoxNet (Maturana and Scherer, 2015), VoxelNet (Zhou and Tuzel, 2018) and SEGCloud (Tchapmi et al., 2017) are examples of voxel based 3D CNNs, with the SEGCloud

architecture shown in Figure 2-23. SEGCloud achieved leading results in 2017 on the semantic segmentation benchmark dataset (Tchapmi et al., 2017); however, like all voxel-based models it suffers from high computation cost due to the need to store unoccupied voxels.
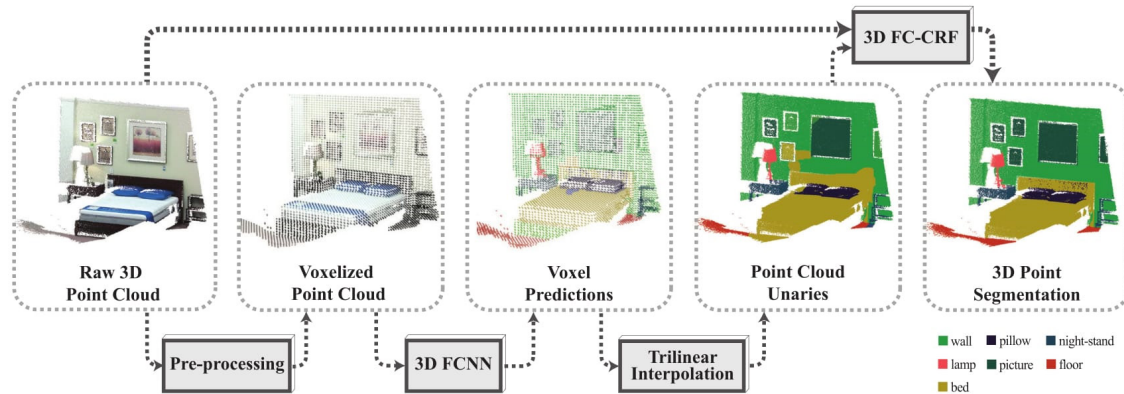
Figure 2-23: SEGCloud 3D CNN model architecture (Tchapmi et al., 2017).

PointNet (Cherabier et al., 2016) pioneered an entirely new approach to deep learning on pointcloud data by removing the convolutions, instead using max pooling as symmetric functions to solve the permutation invariance problem, combined with shared MLPs for spatial encoding and learned transformation matrices. The initial PointNet did not use local point context, which was added in the follow up model PointNet++ (Qi et al., 2017). Many other architectures either modify PointNet directly or reuse some of its intuitions (Bello et al., 2020).

A different approach is taken by Landrieu & Simonovsky (2018), where superpoint graphs are used to geometrically partition a pointcloud before applying a deep learning model which combines PointNets and graph convolutions. At the time of its publication, this method substantially improved on the state of the art on the benchmark Semantic3D dataset; however, it has since been overtaken by ever newer, faster and more accurate methods. As with deep learning for image processing, the field is evolving at an incredibly rapid pace; the results tables of the benchmark datasets such as Semantic3D (Hackel et al., 2017a), SemanticKITTI (Behley et al., 2019) and S3DIS (Armeni et al., 2017) are the best place to keep abreast of current developments. Figure 2-24 graphs the results from the Semantic3D benchmark's Reduced-8 dataset over time, with papers discussed in this literature review highlighted. It can be seen that each discussed paper represented a move forwards in achievable accuracy at each time.
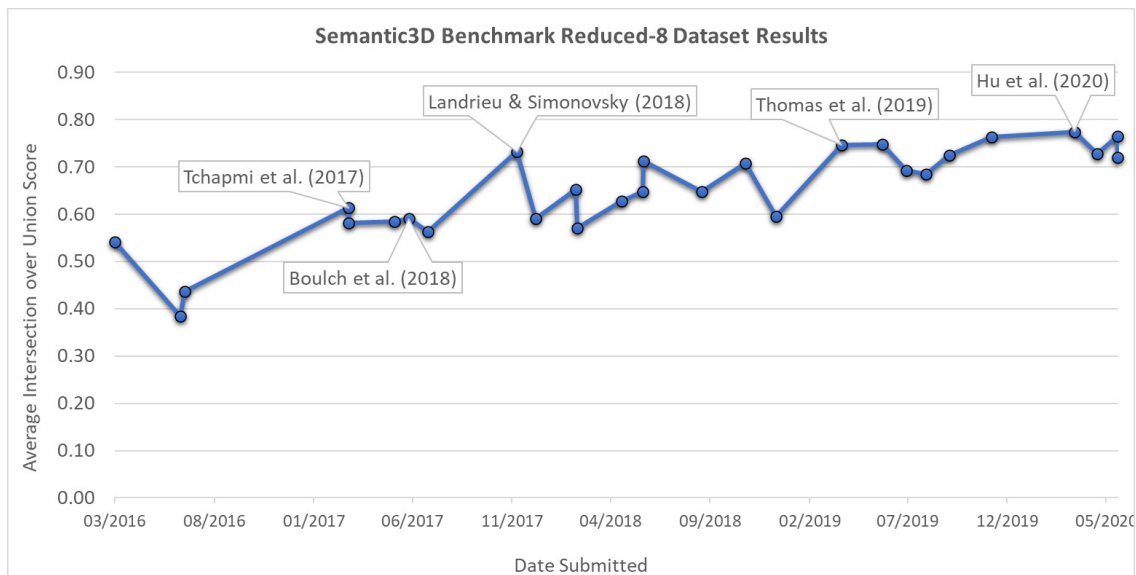
Figure 2-24: Graph of results on the Semantic3D's Reduced-8 dataset over time. Labels are added to highlight papers discussed in this literature review.

The current leaders on the two Semantic3D datasets[5] are ConvPoint (Boulch, 2019) and RandLA-Net (Hu et al., 2020). ConvPoint replaces the discrete convolutions used by grid based CNNs with a continuous convolutional kernel; this adaption allows unstructured point cloud data to be processed by a model with a very similar architecture to successful 2D CNNs (Boulch, 2019). RandLA-Net greatly improves processing speed by using random point sampling combined with local feature aggregators and does not use any pre or post processing operations (Hu et al., 2020). In a real world trial by the Alan Turing Institute it was found that whilst the KPConv method (Thomas et al., 2019) provided the most accurate results on the Birmingham test dataset, the training took over two weeks. RandLA-Net on the other hand achieved near comparable results with only a few hours of training (Data Study Group Team, 2020).

---

[5] The Semantic3D benchmark has two test sets named Semantic-8 and Reduced-8. The Reduced-8 dataset has been downsampled to 0.01m.

2.4.6 Practical deep learning

Sections 2.4.2 to 2.4.5 have introduced the basic theory of deep learning on images and point clouds. The design of cutting-edge architectures is primarily the preserve of mathematics and computer science; from an engineering perspective the priority must be given to practical implementation of these models. Historically deep networks were notoriously difficult to train; however, with larger training dataset sizes and more powerful hardware it has become more straightforward to apply these models to real world problems (Goodfellow et al., 2016). This section discusses several important practical elements of deep learning implementation in a general context, specific implementation details for the models developed in this thesis are given within their relevant chapters.

2.4.6.1 Generating training data

The type of learning examined in this thesis is supervised learning, therefore, the model must have access to a series of labelled examples to learn from, known as the training dataset. Goodfellow et al. (2016) offer a rule of thumb, that a supervised deep learning algorithm will achieve reasonable performance with > 5,000 training examples and will approach or surpass human level performance when provided with > 10 million examples. These training examples are primarily hand labelled by humans and are consequently very costly to create. There are several existing datasets with millions of labelled examples; these are used by researchers designing deep learning algorithms. For applied research, there is a choice between adapting models trained using these large datasets or creating new application specific training datasets. Adapting models trained on other datasets or trained for other problems is known as transfer learning and is discussed in more detail in Chapter 4.

2.4.6.2 Data augmentation

If a new training dataset must be generated from scratch, there are several techniques which can be used to increase the size of the dataset. Data augmentation is where multiple training examples can be generated from a single labelled instance, for example by mirroring or rotating both the source image and its label or by changing the pixel values with a filter, as shown in Figure 2-25. Augmentation generally improves performance on smaller

datasets, such as those found in medical imaging or remote sensing, though it cannot correct for missing instances in the original dataset (Shorten and Khoshgoftaar, 2019).
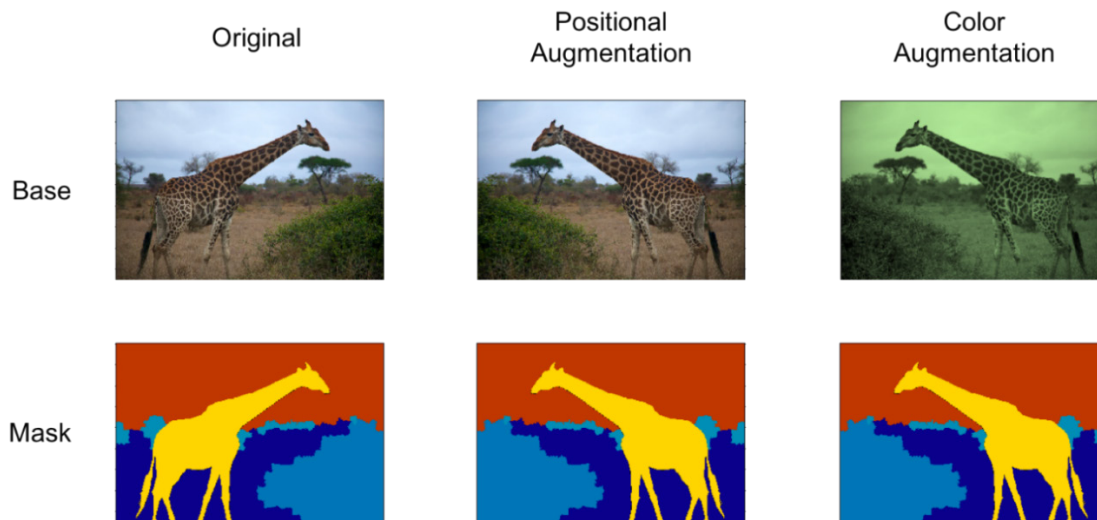
Figure 2-25: Image augmentation examples (MXNet, 2020)*.*

Data augmentation is itself a large and evolving field, with recent advances utilising deep learning methods to design augmentations based on neural style transfers. Shorten & Khoshgoftaar (2019) provide a review and a framework for categorising the many currently used types of data augmentation for deep learning.

2.4.6.3 Dataset structure

Once the labelled data has been generated it must be divided into samples which will be used for training and samples used for evaluating the model's performance. There are usually two evaluation sets, the cross-validation set and the final test set. The cross-validation set is the dataset used to assess the model's performance during development, it can be used during training to monitor overfitting and can be used to examine how changes to the model hyperparameters affect the results. Another strategy, common with smaller traditional machine learning classifies is n-fold cross validation. In this strategy instead of a separate cross-validation dataset the training set is divided into n segments and the model is trained n times, with a different segment acting as the cross-validation set each time. This is a robust way of evaluating the performance on a more varied validation dataset, but becomes impractical with larger deep learning models which take several days to train. Overfitting is a well-known issue in supervised machine learning, where the model becomes

too good at the training data at the expense of its generalisation ability (Géron, 2017), illustrated by Figure 2-26.

<div align="center">

Under-fitting
(not enough variance)

Appropriate fitting

Over-fitting
(too much variance)

</div>
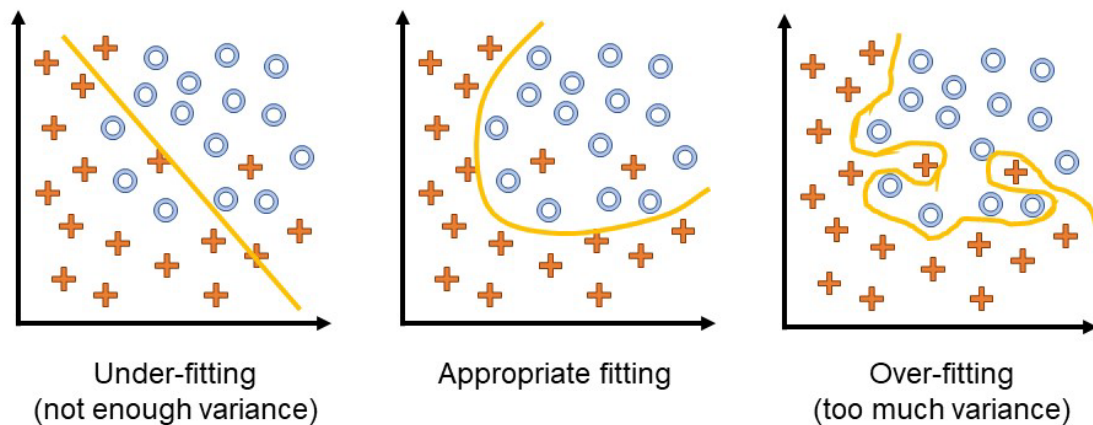
Figure 2-26: Illustration of how simple 2D data can be under or over fit.

A primary symptom of overfitting is a large gap between performance on the training data versus the validation data. Figure 2-27 shows typical accuracy curves for a model during training. Initially both training and validation accuracy is low as the model is yet to learn about the relationships between the data and the labels. As the model trains the performance improves on both sets, with training set performance usually above validation set performance as it is easier to make predictions on familiar data. If the model begins to overfit, the difference between the training and validation set accuracies increases. Overfitting can be minimised by a suite of regularisation techniques, discussed further in Section 2.4.6.5.
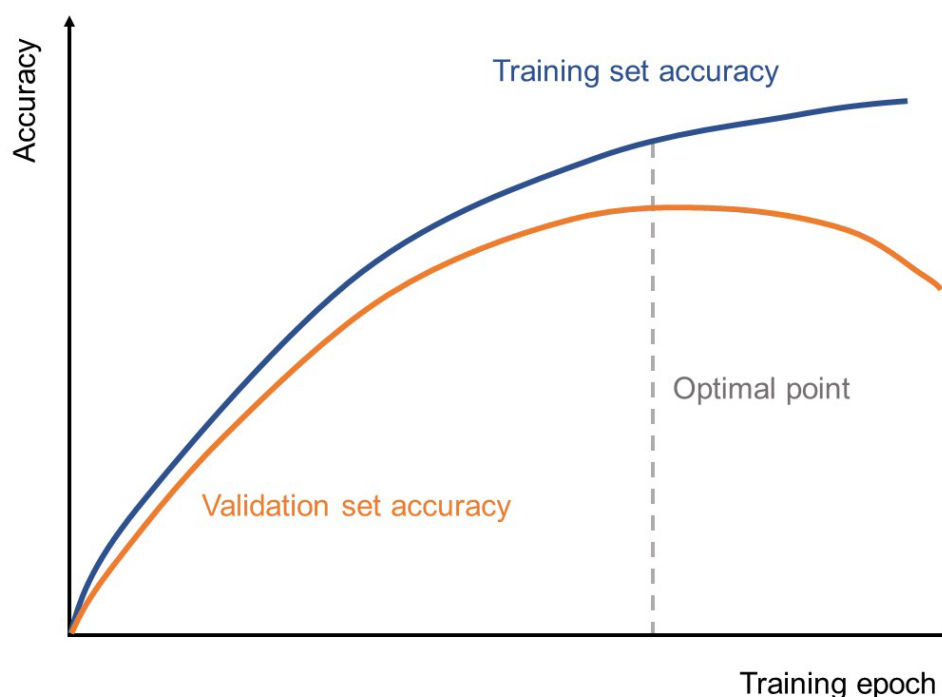
Figure 2-27: Typical accuracy curves for training and validation datasets, plotted against the number of epochs the model has been training for.

As the validation dataset is used to inform choices for model parameters, it is not truly unseen data and if extensive hyperparameter tuning is carried out, the model can begin to overfit on the validation set too. This is because design choices are made with the aim of increasing accuracy on this validation set. For this reason, a third set known as the test or hold out set is used to gain a measure of the model's performance on truly unseen data. This dataset is used only at the very end of a project and the results on this dataset are what is reported in publications and machine learning competitions. Traditionally, datasets were split into roughly 60% for training, 20% for validation and 20% for testing; however, with the advent of deeper models trained on larger datasets the modern splits can be closer to 98% for training and 1% each for validation and testing (Ng, 2017).

2.4.6.4 Hyperparameter tuning

Hyperparameters refer to the values in a neural network that are set by the operator rather than learned by the model. Choosing appropriate hyperparameters is a crucial step in creating successful deep learning models. Hyperparameters can refer to architecture settings such as the model depth and number of hidden units per layer and they can also refer to training options such as the learning rate and the mini-batch sizes. The hyperparameters to be tuned

64

depend on the machine learning model used; specific hyperparameters relevant to the models used in this research will be introduced in the model development sections of each chapter.

Hyperparameters can be selected either manually or automatically. Intuition and experience play a large part in selecting appropriate hyperparameters manually, whereas large amounts of raw computing power are required to select them automatically. A compromise can be to use domain expertise to narrow the search area before running multiple automatic tests. Some hyperparameters are continuous (learning rate), some are discrete (number of hidden layer units) and some are binary (normalisation on/off). Hyperparameter tuning can be considered an optimisation problem, where the objective is to find the optimum value for the hyperparameters which minimises model generalisation error (Goodfellow et al., 2016).

2.4.6.5 Regularisation

The idea of effective capacity is useful when considering model training. A deep learning algorithm's effective capacity is its ability to model complexity; good performance is achieved when its effective capacity is appropriate for the complexity of the task and the size of the available training data (Goodfellow et al., 2016). If it has more effective capacity than needed, it will tend to overfit. Overfitting can be minimised either by obtaining more training data, decreasing the complexity of the architecture or by regularisation. There are multiple regularisation strategies used in practice, the most common of which are weight decay, dropout and early stopping.

Weight decay, also known as L2 regularisation penalises model complexity by adding a regularisation term which multiplies the weight updates by a number smaller than 1, according to Equation 2-5 (Géron, 2017):

$$J(\theta) Regularised = J(\theta) + \alpha \frac{1}{2} \sum_{i=1}^{n} \theta_i^2 \qquad \text{(2-5)}$$

where $J(\theta)$ is the generic cost function and α is the hyperparameter controlling the amount of L2 regularisation required. If we consider w to be the vector of the feature weights θ_1 to θ_n then the second part of Equation 2-5 becomes $\frac{1}{2}(\|w\|_2)^2$

65

which is the L2 norm of the weight vector (Géron, 2017). L2 normalisation diminishes the sizes of some of the weights which makes the model behave similarly to a smaller model, reducing overfitting tendencies.

Dropout regularisation minimises overfitting by randomly eliminating nodes in the network. For every training example, different nodes are eliminated according to a pre-set probability. If the model makes multiple passes through the training data it is important that different nodes are dropped each time the same training example is passed through. Dropout is only used during training, not during validation or testing. Dropout minimises overfitting for two reasons. Firstly, it reduces the effective size of the model on each iteration and secondly it reduces reliance on individual features as they may be randomly eliminated, forcing the model to spread the weights more evenly (Ng, 2017).

Early stopping is a simple form of regularisation where the training is stopped at the highest validation accuracy score; before the training and validation scores diverge. A typical early stopping point is shown by the dotted line on Figure 2-27. A downside of early stopping is it couples the cost optimisation and reduces overfitting objectives making it difficult to examine either independently (Ng, 2017). Data augmentation, as discussed in Section 2.4.6.2 can also be considered a regularisation strategy as it contributes to minimising overfitting tendencies. In general, overfitting is only a problem when using small datasets, as it is essentially the inability of the model to differentiate between signal and noise in the training samples. When the training dataset gets sufficiently large, the model is able to determine the noise element itself.

2.4.6.6 Imbalanced classes

Many detection tasks contain highly imbalanced classes, which if unaddressed will lead to the model predicting only the majority class result. Imbalanced classes are generally addressed in two ways. Firstly, the input data can be resampled to increase the percentage of the minority class, either by downsampling the majority class or upsampling the minority class. Secondly, the loss function can be weighted to force the model to focus more on mistakes made on the smaller classes. the data

2.4.6.7 Software and hardware

There are many powerful open source and commercial libraries available for designing and running deep learning models. A selection of the most popular tools and libraries are detailed in Table 2-1.

Table 2-1: Common machine learning libraries. This list is not comprehensive, only a snapshot of the most commonly encountered libraries. See https://en.wikipedia.org/wiki/Comparison_of_deep-learning_software for a full list.

Name	Programming Language	Comments	Reference
Scikit-learn	Python	Traditional machine learning not deep learning, many algorithms implemented.	(Pedregosa et al., 2011)
MATLAB Deep Learning Toolbox	MATLAB	Not open source but allows fast prototyping with easy to use functions.	(The MathWorks, 2020)
TensorFlow	Python, C++, CUDA	Deep learning library developed by Google.	(Abadi et al., 2016)
PyTorch	Python	Based on the Torch library and primarily developed by Facebook AI Research.	(Paszke et al., 2019)
Keras	Python	High level library designed to be run on top of other libraries such as TensorFlow or PyTorch.	(Chollet, 2015)

Most software libraries run on Windows, Linux and macOS; however, newly released code is usually designed on Linux operations systems. As the code matures, edited Windows versions become available; this lag can be a challenge when developing applications based on the cutting-edge algorithms, if constrained to Windows machines.

Past progress in deep learning has been inextricably linked to available hardware (LeCun, 2019). Algorithm research and development is primarily carried out on High Performance computing (HPC) clusters as performance and flexibility are the primary concerns (LeCun, 2019). Applied research, using transfer learning or smaller specialised models can generally be carried out on the higher end of consumer workstations. Commercial inference[6] applications such as Facebook's image recognition CNNs are mostly carried out on data centre servers (LeCun, 2019). Demand for mobile, robotic and vehicle based

---

[6] Inference is where only the prediction stage is running, using a well understood pre-trained models

systems is accelerating development of extremely low power application specific integrated circuits (LeCun, 2019).

2.5 Remote sensing and machine learning in mining

Over the last decade, numerous mining specific remote sensing applications have been developed. Primarily these have involved human interpretation of the datasets, without the addition of machine learning algorithms. Recently, several applications utilising machine learning to analyse mining sector remote sensing data have been developed. This section of the literature review is structured as follows: firstly, an overview of the non-machine learning remote sensing based mining applications is given, followed by a review of the emerging integrated machine learning and remote sensing mining applications. Lastly, promising machine learning and remote sensing integrations found outside the mining sector are examined.

2.5.1 Non-machine learning remote sensing in the mining sector

Remote sensing technologies are used in many areas of the mining sector, from individual mine sites to government departments. In this section, remote sensing applications are grouped by the location of their sensing platform into the categories of terrestrial, aerial and orbital. This categorisation is effective as other attributes such as data format, scale of survey and stakeholders correlate well with each category. For consistency, this categorisation is also used across each of the three case studies in this thesis.

2.5.1.1 Terrestrial

Applications using terrestrial based remote sensing platforms are at an individual mine site scale and usually take the form of true 3D data. The innovations in this space are pushed by research institutions, equipment manufacturers and the mine sites themselves. Geotechnical applications make up a sizable part of terrestrial remote sensing research in mining and civil engineering. Of primary interest to mining is the study of rock mass characterisation from remote sensing data, both on surface (Lato et al., 2009, Sturzenegger and Stead, 2009, Coggan et al., 2007, Monsalve et al., 2019) and underground (Chen et al., 2018, Mcquillan, 2013, Fekete and Diederichs, 2013). Additionally, Fekete et al. (2010) give an overview of the range of underground operational and geotechnical applications that can benefit from 3D geodata, including calculation of shotcrete thickness, rock reinforcement bolt spacing,

leakage regions, rockmass characterisation and geometrical discontinuity characterisation. Terrestrial platforms have also been used for slope stability applications including rockfall analysis (Rosser et al., 2007, Abellán et al., 2010).

Due to lighting limitations, underground applications primarily favour laser scanning over close range photogrammetry, although some photogrammetric methods are successful for tunnel inspection, a review of these is given in (Attard et al., 2018). Tunnel inspection can also be carried out using laser scanning (Gikas, 2012, Xu et al., 2018, Tan et al., 2016), with a review of multiple tunnelling applications given in (W. Wang et al., 2014). Laser scanning also has been successfully used for measuring the precise underground environment of cross sections (Ganić et al., 2011), shafts (van der Merwe and Andersen, 2013) and areas where accidents have occurred (Eyre et al., 2015). Furthermore, investigations of mobile mapping for underground applications have been reported in Eyre et al. (2016) and Bissir et al. (2008).

Moving away from using remote sensing data for its more traditional purposes of supporting engineering through mapping, new applications have emerged using stereo cameras for rope shovel pose estimation (Lin et al., 2013), haul truck pose estimation (Borthwick, 2009) and other applications which support the move towards automated vehicles in mining (Frimpong et al., 2007, Ruff, 2004). Laser scanners have also been used to aid underground autonomous navigation (Bissir et al., 2008, Wu et al., 2018).

2.5.1.2 Aerial

The aerial category includes data captured from Unmanned Aerial Vehicles (UAVs) and manned aircraft. Due to payload size limitations, large format aerial photography and LiDAR campaigns are generally carried out from manned aircraft, while UAVs are used to generate SfM point clouds using smaller format cameras. UAVs are lower cost and require less personnel to pilot (Simic Milas et al., 2018), allowing faster revisit times in dynamic environments such as open pits. Lee & Choi (2016) provide a review of UAV applications in the mining industry, concluding that topographic surveying of pit benches and stockpiles is the largest primary use of UAV surveying (Tien Bui et al., 2018, Xiang et al., 2018, Q. Wang et al., 2014, Cryderman et al., 2014). Other UAV applications of

interest include measuring fracture orientations (McLeod et al., 2013), virtual structural analysis (Sayab et al., 2018), blast fragmentation (Bamford et al., 2017), hyperspectral monitoring of acid mine drainage (Jackisch et al., 2018), geological investigations (Kirsch et al., 2018), underground coal fires (Wang et al., 2015) and monitoring tailings subsidence (Rauhala et al., 2017).

LiDAR surveys are used for subsidence monitoring (Froese and Mei, 2008, Yu et al., 2011, Palamara et al., 2007), deformation monitoring (Hu and Wu, 2016) and rock mass characterisation (Gigli and Casagli, 2011). Other uses of airborne remote sensing data include studying legacy mine sites for reclamation (Maxwell et al., 2014), radiation mapping (Martin et al., 2015), monitoring restoration success (Moudrý et al., 2019, Padró et al., 2019) and documenting historic mining activities (Hanke et al., 2009, Fernández-Lozano et al., 2015).

2.5.1.3 Orbital

Orbital data is sensed from spaceborne platforms such as satellites. This type of data covers large areas and is primarily used by the mining sector to analyse land use changes related to mining. Spaceborne data is usually 2D and is often multispectral or hyperspectral. Common Ground Sampling Distances (GSD) for optical remote sensing satellite sensors range from sub 1m for the newest generation of Very High Resolution (VHR) sensors (Fu et al., 2020), to several hundred meters for the earlier multispectral sensors (Justice and Townshend, 2002). The mining sector applications include a large number of studies assessing the environmental impacts of mining, comprehensively reviewed in Werner et al. (2019). Other applications include hyperspectral geological analysis (van der Meer et al., 2012), mine subsidence (Wright & Stow, 1999, Demirel et al., 2011) and mine reclamation (Karan et al., 2016).

2.5.2 Machine learning and remote sensing in the mining sector

The integration of machine learning and particularly deep learning applied to mining sector remote sensing data is a newly emerging research area. Many of the contributions are driven by commercial entities and as such have minimal published references. The applications reviewed here are those currently in use in the mining sector but not directly examined in this thesis. Existing work directly related to the applications developed in this thesis are covered in the introductory sections to the relevant chapters.

2.5.2.1 Prospecting and exploration

The most mature area is geological prospecting and exploration. Many traditional machine learning algorithms such as random forests and support vector machines have been applied to tasks such as prospectivity modelling (Yeomans, 2018), lithological classification (Bressan et al., 2020; Cracknell and Reading, 2014; Yu et al., 2012), bedrock mapping (Hood et al., 2019) and mineral prospectivity mapping (Rodriguez-Galiano et al., 2015). Deep learning research to date has been primarily concerned with mineral prospectivity mapping (Sun et al., 2020; Xiong et al., 2018; Xiong and Zuo, 2020; Zuo et al., 2019); however, other applications such as surficial geology mapping (Latifovic et al., 2018), lineament interpretation (Naprstek, 2020) and close range rock identification (Liu et al., 2020) have also been examined. Whilst not strictly related to remote sensing data, other interesting geological applications based on deep learning algorithms have been used to analyse scanning electron microscope images (Chen et al., 2020) and thin sections (Pires de Lima et al., 2020). Alongside the academic work mentioned above, many commercial applications are being developed in the AI for prospecting sector, most notably by Goldspot Discoveries (Holmes, 2019), Earth AI (Barich, 2019) and the IBM joint venture Goldcorp (Moore, 2019).

2.5.2.2 Fragmentation analysis

Another area which has seen research interest is fragmentation analysis. Bespoke machine learning algorithms developed by Thurley and Ng (2008) demonstrated that it is possible to automatically classify broken material from laser scan data, either on conveyors (Onederra et al., 2015a), from rock piles in an open pit mine (Thurley, 2013, Onederra et al., 2015) or from underground draw points (Campbell and Thurley, 2017). MotionMetrics, a Canadian mining technology company, have extended this concept using deep learning with imagery to develop their PortaMetrics automated rock segmentation system (Azmin et al., 2016, Ramezani et al., 2017). Another commercial fragmentation analysis technique is proposed by Petra (Stewart, 2018), where laser scan data is classified using the eigenvalue based traditional machine learning algorithms described in Weinmann et al. (2015). Shao et al. (2020) uses a CNN algorithm to detect large coal blockages on a conveyor.

2.5.2.3 Automation

Autonomous technologies rely heavily on deep learning algorithms, for tasks such as object detection, hazard avoidance and vehicle localisation. Most mining sector research in this area is carried out by the large equipment manufacturers such as Caterpillar, Komatsu and Sandvik (Marshall et al., 2016) and the precise implementation details are not publicly available. However, the techniques used for autonomous surface haul trucks can be considered broadly similar to those employed by the automotive industry (Price et al., 2020), with the main differences relating to the scale of the vehicles and the harshness of the environment. For the automotive industry, Grigorescu et al. (2020) provide a thorough review of how different deep learning algorithms are applied in this sector.

In other areas of mine automation, Somua-Gyimah et al. (2019) propose a deep learning based vision system to classify terrain and recognise objects in an automated dragline excavation scenario. MotionMetrics have built on their fragmentation solution (described in Section 2.5.2.2) to develop an intelligent shovel bucket monitoring system, combining fragmentation analysis, missing tooth detection, tooth wear monitoring and foreign object detection (Shariati et al., 2019). A review of other emerging AI frontiers in mining is given by Ali and Frimpong (2020).

2.5.2.4 Environmental management

The increase in availability of high-resolution satellites has enabled governments and regional stakeholders to monitor land cover and land use changes related to mining at an unprecedented level of detail. Primarily these applications use traditional machine learning approaches to classify land use, with an extensive review given by (Chen et al., 2017). A broader review of how the impacts of mining can be studied using recent advances in remote sensing is given by Werner et al. (2019). Whilst not limited to machine learning and classification studies, this review gives a valuable overview of the different monitoring and impact assessment applications. Studies using deep learning algorithms are more scarce, though several interesting applications have been published in the last few years. Ferreira et al. (2020) developed a unique tailings dam detection benchmark dataset for Brazil before using it to test multiple CNN classifiers. Balaniuk et al. (2020) also apply CNN algorithms to

the problem of tailing dam detection. Maxwell et al. (2020) demonstrate a method using Mask R-CNN to identify geomorphological features associated with mountaintop removal coal mining and Chowdhary et al. (2019) use a U-Net based CNN to detect coal stockpiles.

2.5.3 Integrated applications outside the mining sector

Most of the deep learning based applications discussed in the previous section have been created by taking work carried out in the broader remote sensing and computer vision domains and adapting it for mining specific use cases. This section will briefly describe some of the other applications which could be of interest to the mining sector. Deep learning for general remotely sensed data, particularly VHR imagery has seen a large amount of research attention in recent years, with extensive reviews of the topic given in Ball et al. (2017) and Zhu et al. (2017). Figure 2-28 illustrates the research topics of the papers reviewed by Ball et al. (2017) as a word cloud, scaled according to the number of papers reviewed. Whilst not exhaustive, it is a useful way to visualise the breadth of applications and see how they could be relevant to the mining sector.

Geo-information extraction
3-D analysis Traffic flow analysis
Dimensionality reduction
Ship classification/detection
Vehicle detection/recognition
Environment and water analysis
Scene understanding
Object tracking ATR Segmentation
Object recognition and detection
Animal detection Image registration Pansharpening
Classification
Planetary studies Urban/building
Super-resolution
Human detection
Advanced driver-assistance systems
Data fusion Weather forecasting Land use
Image enhancement
Anomaly detection Disaster analysis/assessment
Change detection
Road extraction
Semantic segmentation
Underwater detection

Figure 2-28: Word cloud depicting the research topics which integrate remote sensing and deep learning reviewed by Ball et al. (2017). Word size is indicative of the number of studies reviewed.

It is easy to see the possible parallels with mining related challenges from the applications such as vehicle and ship tracking (Chen et al., 2014, Tang et al., 2015, Konoplich et al., 2016), off road driving (Procopio et al., 2009, Hadsell et al., 2009, Alamiyan-Harandi et al., 2020) and human detection (Ouyang and Wang, 2012, Tomè et al., 2016). However, other more general topics such as

change detection, semantic segmentation, object detection and land use classification could also be applied to mining sector data.

2.6 Conclusion

This literature review firstly has provided an introduction to passive and active remote sensing systems and the 2D and 3D data structures associated with these technologies. Next it has provided a theoretical background and review of major developments within the broad field of deep learning, with a focus on image processing tasks. A review of state of the art for using deep learning on point clouds is also given in this section. The deep learning section concludes with a review of the practical considerations required for successfully training algorithms of this type. Finally, this chapter examines how remote sensing and deep learning technologies, both separately and combined, are currently being used in the mining sector for a diverse range of applications and indicates where there could be further scope for adapting algorithms from other sectors to mining specific problems.

As this literature review shows, the breadth of applications across the wider scientific community for the integration of remote sensing and deep learning algorithms is large. Almost all of the applications shown in Figure 2-28 could be adapted for mining specific data; however, for this thesis, potential applications have been developed with the goal of demonstrating the broadest applicability of these new technologies. With this in mind, three applications have been developed based on each of the fundamental remotely sensed data types discussed in Section 2.3. These applications are aimed at different groups of stakeholders, as the benefits from integrating deep learning and remote sensing technologies vary depending on the end user, from equipment manufacturers using these new technologies to gain a market edge, to mine sites increasing productivity via increased automation, to governments gaining a better understanding of the environmental impacts of mining in their jurisdictions.

Using machine learning techniques to detect objects in 3D point clouds

Chapter overview

This chapter investigates how machine learning can be used with full 3D terrestrial laser scanner data for underground mining applications. Underground mining applications were chosen as the focus area for this chapter in order to fully exploit the nature of true 3D point clouds. The focus of the application developed in this chapter is the location of supporting rock bolts; however, similar techniques could be applied to a range of underground point cloud identification tasks such as hazard detection, pipe identification and as an input to route planning algorithms.

This chapter is primarily based on the paper **'A machine learning approach for the detection of supporting rock bolts from laser scan data in an underground mine'**, published in the journal Tunnelling and Underground Space Technology[7]. Additionally, this chapter contains the results from testing using a modern production mine dataset and it also includes the development of a bolt density assessment technique for pattern bolted mines.

3.1 Introduction

Rock reinforcement is a crucial element of underground construction. When operating with any underground excavation, an understanding of the rock mass characteristics as an engineering material is critical in ensuring that risks from tunnel collapse are mitigated through the use of ground control methods. Installation of rock bolts is the most widely used form of ground support (Li, 2017). The design of such a system is site dependent and based on the mechanical behaviour of the rock mass, the in-situ stress field and induced stress from the excavation (Hoek and Brown, 1982). In low stress conditions, compression of the ground is needed to ensure loose blocks do not fall. This can be achieved either by using spot bolting of discrete blocks or by a systematic bolting pattern. Spot bolting is carried out where needed without

---

[7] The candidate is the first author of this paper and the authorship contribution statement is as follows: **Jane Gallwey:** Methodology, writing (original draft & revisions), investigation, code development. **Matthew Eyre:** Conceptualization, writing (review & editing), supervision. **John Coggan:** Project administration, writing (review & editing), supervision.

following a set spacing, whereas systematic patterns are used to add a compression arch to the rock mass, reducing the potential for unravelling. Schach et al. (1979) show that an increase in bolt spacing leads to less interaction of neighbouring bolts, reducing the size of the compression zone to a point at which the bolts no longer provide a wide coverage leading to potential fall of ground. To ensure the required level of compressive cover is produced, it is important that correct installation of bolt patterns is carried out. Reconciliation of installed bolts is therefore an important part of the ground management process to ensure safe working underground.

There is no published literature on the current methods of documenting rock bolt installation; these are usually hand sketch based and not comprehensive (Öberg, 2013) due to the large volume of bolts that have to be recorded and the difficulty and time-consuming nature of manually surveying such data, along with the associated human error for this type of repetitive task. Another difficulty is that in many applications the entire surface is covered with shotcrete after installation, rendering the exact locations of the rock bolts unknown or challenging to discern after shotcrete installation (Öberg, 2013). Automatically detecting and recording the 3D coordinates of rock bolts either retrospectively or at installation would allow for greater quality assurance and quality control, providing a detailed record of exactly where rock bolts have been installed. These records also would be critical in a fall of ground situation, where the exact bolting configuration that was installed prior to the incident must be determined to verify the workings were adequately supported. Advancements in remote sensing techniques and machine learning algorithms could allow this bolting pattern information to be obtained. However, currently the mining sector is not fully utilising these new technologies despite being well placed to employ them due to a widespread adoption of laser scanners and other high resolution surveying technologies both onboard vehicles and as standalone survey technologies (Body, 2014).

To date, image based photogrammetric systems for automatically inspecting civil engineering tunnels have been the primary research focus in this area. A review of these techniques is given in Attard et al. (2018) and successful implementations for crack detection by Huang et al. (2018) and moisture mark detection by Zhao et al. (2020), demonstrating the power of remote sensing and

machine learning for underground infrastructure management. However, passive remote sensing methods such as those used in the above referenced studies can be problematic underground, particularly in mines, due to challenges from uneven illumination and dust (Gikas, 2012). Active systems such as laser scanning can circumvent these issues, by measuring using multiple high speed laser pulses emitted from the instrument itself therefore removing the need for external illumination (Eyre et al., 2016). The data obtained from a laser scanner is in the form of a 3D point cloud which records the X, Y, Z coordinates of the reflected point in 3D. Most scanners also record the intensity of the laser return and some also use cameras to store an RGB colour value for each point. The primary issue with laser scanners compared to cameras is the size of the data collected and the subsequent difficulty in efficiently processing it. The raw output from the laser scanner is a large unordered set of 3D coordinates with no semantic knowledge of the object they are surveying. This 3D point cloud data is currently used by mines directly for surveying tasks such as change detection, geometric analysis and as-built to design comparison (van der Merwe and Andersen, 2013). In order for this data to be utilised in a wider range of applications such as automated machines, mine information databases and infrastructure monitoring a level of semantic information needs to be added to the data, along with a reduction in the dataset size.

The most directly applicable prior work on this topic is by Martínez-Sánchez et al. (2016). In their work they built and trained an autoencoder based model to detect not only the rock bolts from laser scan data, but also their orientations and the shotcrete thickness. Their work achieved a bolt detection rate of 91% showing that geometric neighbourhood based machine learning algorithms have great potential to solve this engineering and monitoring problem. However, their dataset consisted of clean, generally planar shotcreted surfaces with minimal confusion objects.

Laser scanners also have been used in tunnel inspection (Tan et al., 2016; Xu et al., 2018) however, these studies have used the laser scan data to generate intensity images rather than detecting objects from the 3D point cloud data. These methods are successful in infrastructure tunnelling projects where the tunnels are relatively empty and the images can be generated using a fixed

perspective distance; however, these methods are unlikely to work in mines as there are many confusion objects present at varying distances from the scanner. Soilán et al. (2019) give a full review of the use of laser scanners for infrastructure monitoring.

Whilst there is minimal published work on detecting discrete objects in an underground environment from laser scanned data, automatically generating an understanding of a scene from point cloud data has been the topic of much research in recent years. Most application oriented work in this field focuses on either identifying roadside objects and road characteristics from surface mobile laser scan data (Balado et al., 2018; Lehtomaki et al., 2016; Soilán et al., 2017; Yang et al., 2013) or on ground cover classification from aerial LiDAR data (Blomley et al., 2016; Niemeyer et al., 2014; Rau et al., 2015). Properties of these types of surface scenes, such as proliferation of regular vertical objects in streetscapes and a mostly fixed view angle in aerial LiDAR can be leveraged to aid in detecting these types of objects, unlike in the underground environment. Underground terrestrial and mobile laser scan data is complex as it is true 3D data, with the possibility of multiple points sharing the same XY location but possessing different Z values. Approaches used for identification of discrete objects on roads, such as Weinmann et al. (2017) for trees and Lehtomäki et al. (2010) for poles can be considered the closest neighbours, and techniques from these studies can be adapted to the problem of identifying underground features or objects with regard to the particular properties of the underground environment.

These close research applications use variations on the classical point cloud machine learning method, described in section 2.4.5 of this thesis. Other less common methods include directly classifying using Markov networks (Agrawal et al., 2009; Anguelov et al., 2005; Triebel et al., 2006), spectral hashing (Behley et al., 2010) and most recently, approaches using deep learning. For this application, the classical approach similar to Weinmann et al. (2017) was selected, allowing a low computational burden which is more appropriate for time critical applications such as those deployed on underground vehicles and equipment. Whilst deep learning approaches have shown impressive results (Cherabier et al., 2016; Maturana and Scherer, 2015; Riegler et al., 2017), the additional model complexity, computational power, training time and the size of

the training data required for successful deployment make these methods less attractive for an efficient vehicle-based solution.

This chapter will describe an automated approach for rock bolt identification from laser scan data using machine learning, based on the classical point cloud semantic segmentation but implemented using a more extensive set of features from both the robotics and remote sensing communities, alongside adaptations for the geometry of underground environments. The machine learning element of the research compares a random forest, which was the highest scoring classifier in the literature, with the fully connected neural network model developed in this research. Following the classification, the bolt objects are extracted via clustering and centroid generation.

3.2 Datasets

A large amount of labelled data is required to train a machine learning classifier to detect objects. As there are no available datasets of labelled laser scanned rock bolts, two datasets were collected and annotated specifically for this study.

3.2.1 Cornwall dataset

The primary dataset used in this research was collected from a 250m section of underground workings from a currently disused small tin and copper mine. This is a good training area, as the slaty nature of the country rock manifests itself as a fair to poor quality rock mass, resulting in extensive spot bolting based on observations of potential block fallouts. The area of interest was surveyed using a terrestrial laser scanning workflow. The scanner was mounted on a static tripod to perform a scan, next the scanner was moved to a position approximately 12m further down the tunnel and another scan was taken. This process was repeated for 25 scans. The individual scans were registered together to make a unified dataset in the point cloud processing software Leica Cyclone. The hardware used was a Leica C10 laser scanner, as used in other underground studies such as Ganić et al. (2011), S. Chen et al. (2018) and Long et al. (2018). This instrument has a specified accuracy of ± 6mm per point (Long et al., 2018) and the scan resolution at the chosen setting provides a point spacing of 5mm at 5m from the scanner, allowing small variation in tunnel geometry to be captured. The scanner was set to record only laser intensities not optical imagery values. This is due to the poor illumination in the mine and

the training data required for successful deployment make these methods less attractive for an efficient vehicle-based solution.

This chapter will describe an automated approach for rock bolt identification from laser scan data using machine learning, based on the classical point cloud semantic segmentation but implemented using a more extensive set of features from both the robotics and remote sensing communities, alongside adaptations for the geometry of underground environments. The machine learning element of the research compares a random forest, which was the highest scoring classifier in the literature, with the fully connected neural network model developed in this research. Following the classification, the bolt objects are extracted via clustering and centroid generation.

3.2 Datasets

A large amount of labelled data is required to train a machine learning classifier to detect objects. As there are no available datasets of labelled laser scanned rock bolts, two datasets were collected and annotated specifically for this study.

3.2.1 Cornwall dataset

The primary dataset used in this research was collected from a 250m section of underground workings from a currently disused small tin and copper mine. This is a good training area, as the slaty nature of the country rock manifests itself as a fair to poor quality rock mass, resulting in extensive spot bolting based on observations of potential block fallouts. The area of interest was surveyed using a terrestrial laser scanning workflow. The scanner was mounted on a static tripod to perform a scan, next the scanner was moved to a position approximately 12m further down the tunnel and another scan was taken. This process was repeated for 25 scans. The individual scans were registered together to make a unified dataset in the point cloud processing software Leica Cyclone. The hardware used was a Leica C10 laser scanner, as used in other underground studies such as Ganić et al. (2011), S. Chen et al. (2018) and Long et al. (2018). This instrument has a specified accuracy of ± 6mm per point (Long et al., 2018) and the scan resolution at the chosen setting provides a point spacing of 5mm at 5m from the scanner, allowing small variation in tunnel geometry to be captured. The scanner was set to record only laser intensities not optical imagery values. This is due to the poor illumination in the mine and

the additional time required to take photographs with the inbuilt camera. The final dataset is representative of real world underground scan data, containing laser noise, occlusions and many objects that are neither tunnel nor bolt and it has not been manually cleaned and simplified for improved machine learning results. A sample of the data is shown in Figure 3-1.
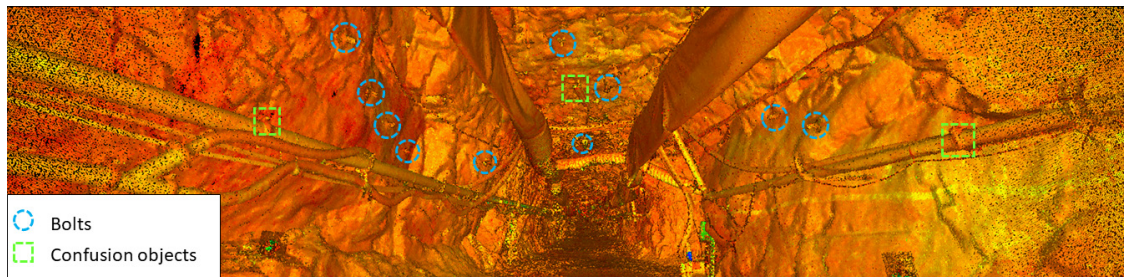
Figure 3-1: A view of the underground data. Many challenging objects are present including pipes, brackets, ventilation bagging and electrical boxes. The colour scheme is taken from the strength of the laser return.

3.2.2 Production mine dataset

To further assess the effectiveness of the method, an additional dataset was surveyed in a large modern zinc and lead mine. This mine is primarily shotcreted before or after bolting and the majority of the rock bolting in this mine is installed to conform to a pattern rather than the spot bolting seen in the first dataset. This dataset was surveyed with the Leica BLK360, a lower resolution and lower cost laser scanner. This was chosen to more closely resemble the datasets which might be obtained from mobile mapping systems onboard mining machines. The scans were surveyed at the medium resolution setting, providing an approximate point spacing of 12mm at 10m from the scanner. This second dataset was not measured as one contiguous block, instead multiple scans were taken throughout different areas of the mine and registered area by area, with the areas shown in Figure 3-3.

3.2.3 Pre-processing

To generate the training data, the rock bolt points were manually separated from all other points and given the class label 1 'bolt'. All other objects were labelled 0 'not-bolt', including confusion objects such as pipes, brackets and ventilation bagging, alongside the hanging wall, side wall and foot wall surfaces. The Cornish dataset was then split into sections for training, cross-validation and testing, as shown in Figure 3-2. The production mine dataset was not in one contiguous section, the different sub areas used for training and testing are

82

shown in Figure 3-3. Four fold cross-validation, as described in Section 2.4.6.3 was used instead of reserving a dedicated cross validation sub-area. In both figures the grey areas are unused and have been reserved for future algorithm testing.

Figure 3-2: Tunnel showing the areas for training (blue), cross-validation (green) and testing (red) from the Cornaish dataset.
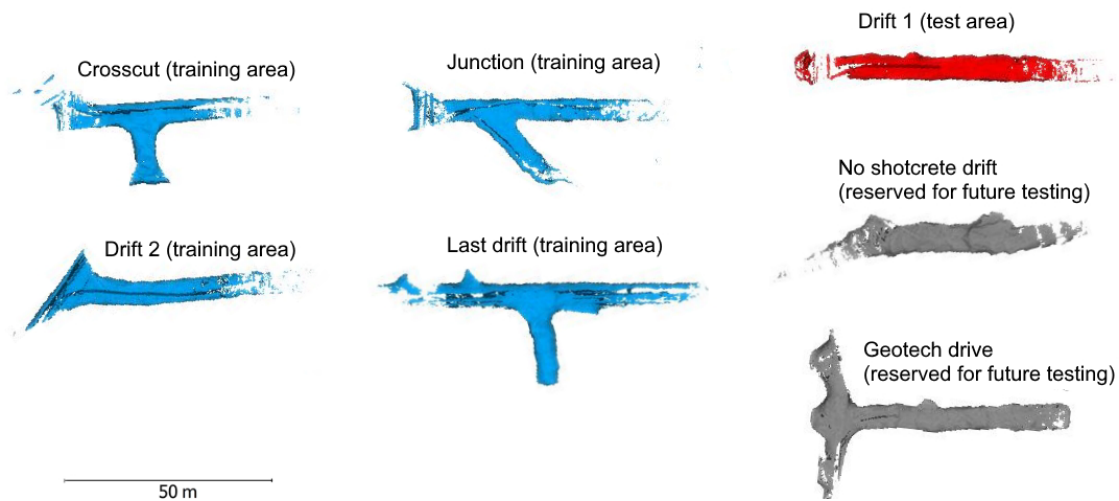
Figure 3-3: Production mine dataset areas. Four fold cross-validation was used instead of a dedicated cross validation dataset.

Before the point cloud dataset features can be generated a number of preprocessing steps are carried out, using the open source software CloudCompare (Girardeau-Montaut, 2016). Firstly, the point clouds are shifted from their real-world coordinates to a position near the origin to avoid potential precision loss from processing very large numbers. Next, denoising is carried out using CloudCompare's noise filter tool, this works similarly to a low pass image filter. This tool removes points which are further than a set factor of their neighbours reprojection error onto a plane, where the plane itself is fitted to all points within a specified radius (Girardeau-Montaut, 2016). The denoising settings used a radius of 10cm and a relative error factor of 1, the relative error factor of 1 is standard for clouds without excessive range noise and the 10cm value was chosen as it is 10 times the final required cloud spacing. The final step in the base dataset creation is density reduction. Point clouds acquired from laser scanners have a large variation in density due to many factors

83

including an object's distance from the scanner, the scan angle, overlap between neighbouring scans and occlusions. Whilst it is difficult to create new points in areas of low density, it is straightforward to remove points in areas of high density using resampling techniques. For this application, the point cloud was spatially resampled to a density of 1 point per cm maximum. Figure 3-4 shows the distribution of point densities on a section of the Cornish dataset before and after resampling.
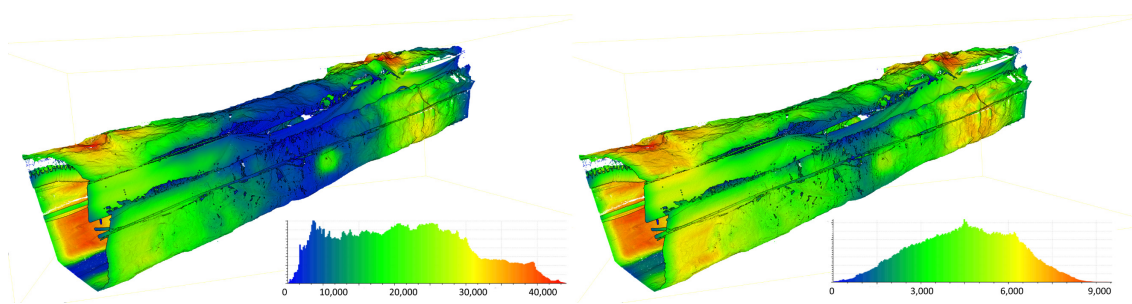
Figure 3-4: A section of the training data showing the density before (left) and after (right) spatial resampling. The density is measured as the number of points per square meter of tunnel surface. The graphs below each image show the range of data densities.

The resampling algorithm also reduced the total number of points by ~40%. As shown in Figure 3-4, the density range is now closer to a normal distribution, but still not constant across the point cloud. This is because a constant density is undesirable for real world data, as there will always be areas of low sampling due to occlusions, however, if the majority of the cloud is downsampled to match the lowest density much of the useful detail can be lost.

3.3 Methods

The workflow for detecting bolts from the laser scanned point cloud dataset has three primary components: feature descriptor creation, machine learning classification and object creation. An overview of the processing workflow is shown in Figure 3-5.
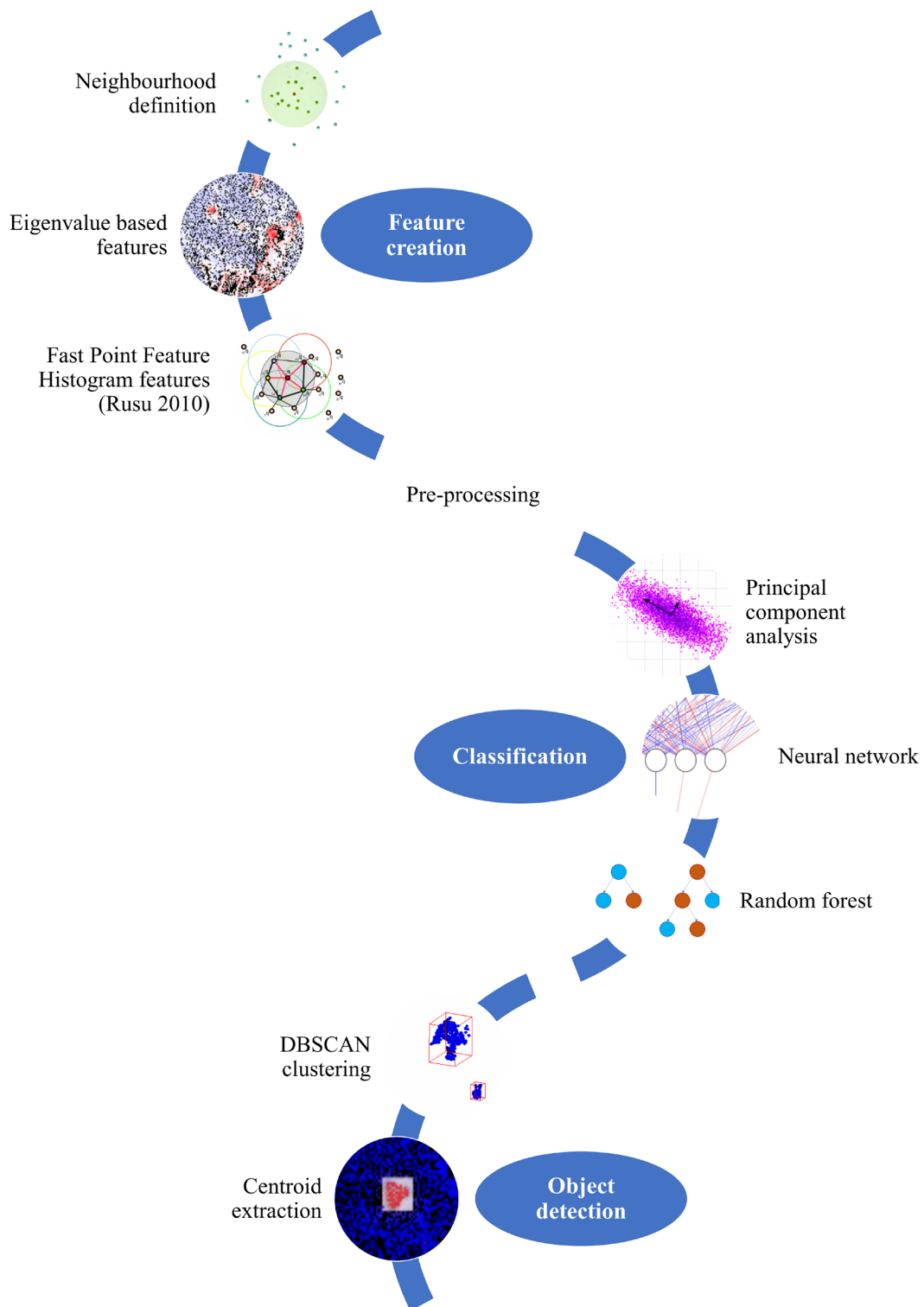
Figure 3-5: Methodology diagram outlining the pipeline used for the task of identification of rock bolts from the laser scan data.

3.3.1 Feature creation

Single laser scanned points are not adequate descriptors of the data they represent, as they contain only 3D cartesian coordinates and an intensity value. In isolation, this information is insufficient to describe what type of object this point belongs to; therefore, the point cloud data must be encoded in a way that

allows a machine learning algorithm to differentiate between object types. This can be achieved by describing each point in relation to the geometry of its neighbouring points, these descriptors are known as features. The most popular features in the remote sensing community are based on the eigenvalues of the point neighbourhood. Early work by Pauly et al. (2003) and Vandapel et al. (2004) introduced the concept, which was extended by Jutzi and Gross (2009) and Weinmann et al. (2015b). The other common features are proposed by Rusu (2010) and implemented in Point Cloud Library (PCL)(Rusu and Cousins, 2011). This approach computes a Fast Point Feature Histogram (FPFH) based on the angular variations between the normals of the points using a Darboux frame (Rusu et al., 2009).

For choosing a point neighbourhood, the dimensions of the object to be detected and the spacing of points in the point cloud determine the optimum value. A typical mechanically anchored rock bolt measures 16cm across the faceplate. Computing the number of neighbours per point over the resampled point cloud using an 8cm radius found the mean number of neighbours to be close to 100, therefore, this is a suitable neighbourhood size to adequately capture the geometry of a rock bolt. Once the neighbourhoods have been defined, descriptive features can be constructed for each point using its neighbours.

Two types of feature sets are calculated for each point in the cloud. The first are the 'Geometric' features, described fully in Weinmann (2016). These include simple 2D and 3D properties of the neighbourhood (density, vertical difference, minimum bounding box), eigenvalue based features which describe the local shape properties of the neighbourhood and 2D accumulation map based features, an overview of each individual feature is given in Table 1. These features were calculated using python code adapted from the MATLAB script published by Weinmann et al. (2015a). The 2D accumulation map features have the highest processing overhead and also are potentially less descriptive for an underground scenario where the hanging wall and footwall share the same XY coordinates, to investigate, the feature sets were generated both with and without these features. The geometric feature set is powerful as it is understandable and can be easily visualised, Figure 3-6 shows a small section of hanging wall with the points coloured by the magnitude of different features. It

can be seen that certain features are intuitively better at differentiating between 'bolts' and 'not-bolts' for a human interpreter; however, some of the less obvious features may still be strong descriptors as they can help to separate between false positives and true positives. As can be seen in Figure 3-6 (a) the areas of high curvature change clearly correspond to rock bolt locations, whereas in 3-6 (b) the 3D density appears to be more related to the distance from the scanner than the bolt location, indicating that is probably not a particularly effective feature for locating rock bolts. The omnivariance feature shown in Figure 3-6 (c) is high for the bolts but also high for other areas of discontinuities, especially visible in the vertical lines near the centre of the image, whereas in 3-6 (d) it can be seen that the verticality feature, despite does not spotting rock bolts has high values in the same areas of non-bolt discontinuities that were highlighted in 3-6 (c).
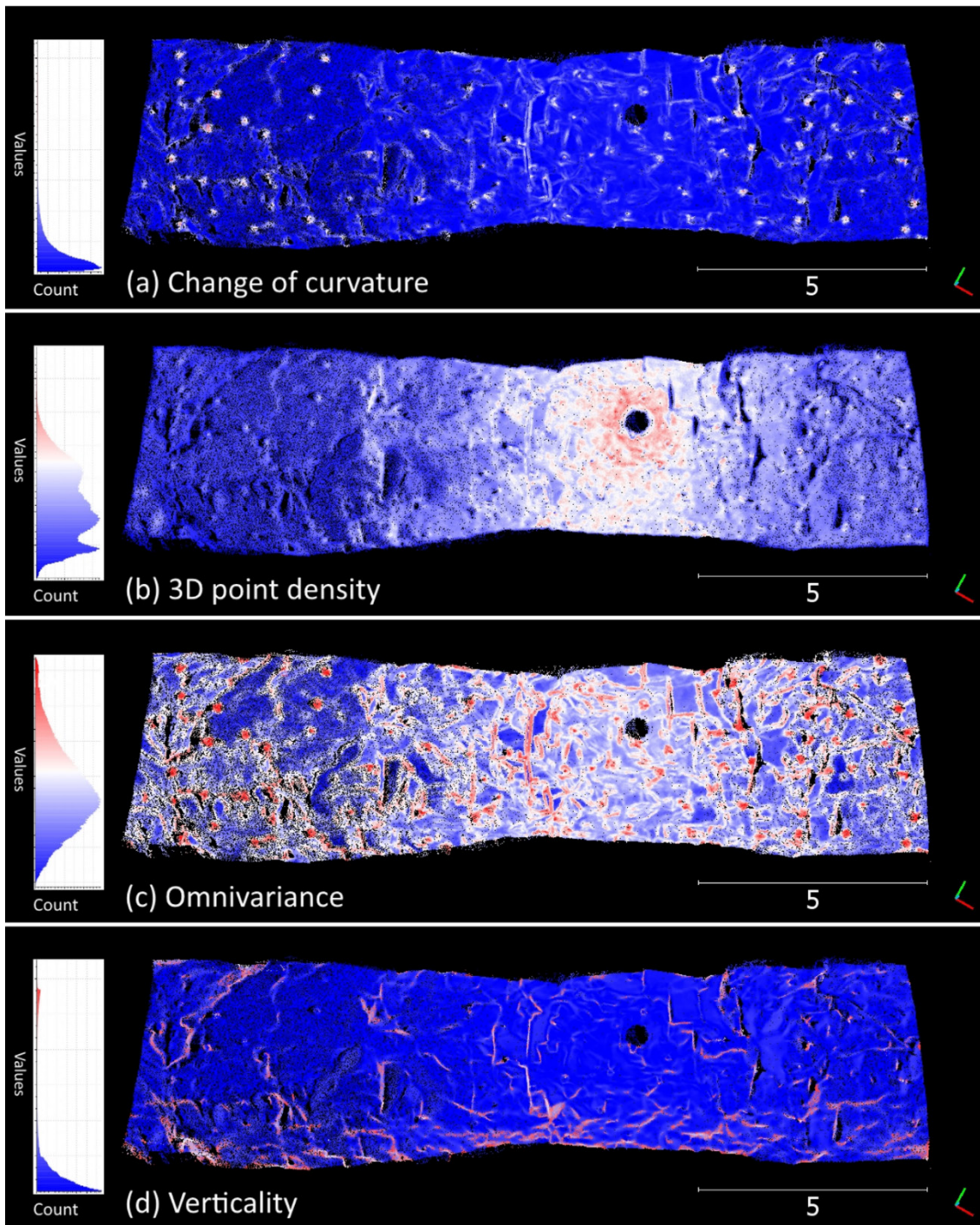
Figure 3-6: A section of hanging wall showing each point coloured by its feature value. Features shown are (a) change of curvature, (b) point density, (c) omnivariance and (d) verticality. All scales are relative, and the colour scheme banding runs from blue (lowest) to red (highest) with white as the median value.

The second type of features used are the fast point feature histogram features (FPFH) proposed by Rusu (2010). This type of feature representation uses the relationships between the points in the neighbourhood and their normal vectors to describe the local geometry around the point. This is calculated for each pair of points by defining a fixed Darboux coordinate frame at one point and using it to compute the three angles which define the difference between the normal

vectors. The complexity is then reduced by not computing the same neighbourhood pairs for multiple points and instead using a weighting scheme. Finally, the values are binned into a 33 bin histogram. Full derivation of the FPFH is found in Rusu (2009). This step was implemented in C++ with the Point Cloud Library (Rusu and Cousins, 2011).

As the intensity data adds further valuable information about the object, especially underground (Xu et al., 2018), two additional features; the intensity of the point itself and the average intensity of the neighbourhood are computed and added to the feature set. As all sets of features are computed individually for each point using the same set K number of neighbours the geometric, FPFH and intensity features can be concatenated, along with the X, Y, Z data for the point and the true class label. The result is a 65-dimensional vector describing the local geometry in a way that can be statistically interpreted by the machine learning classifiers in the next stage, shown in Table 3-1 overleaf.

Table 3-1: Composition of the generated 65 dimensional vector including equations, where k = number of neighbours and λn = eigenvalue n. Eigenvalue equations shown are adapted from Weinmann (2016). For brevity, each FPFH value is not shown as they are numbered elements from the same histogram.

No	Name	Description	Equation
1	X	X coordinate of point	n/a
2	Y	Y coordinate of point	n/a
3	Z	Z coordinate of point	n/a
4	Label	Point label	n/a
5	Intensity	Reflectance intensity of point	n/a
6	Linearity	How much variance can be explained by only the largest eigenvalue	$(\lambda_1 - \lambda_2)/\lambda_1$
7	Planarity	How much variance can be explained by the two largest eigenvalues	$(\lambda_2 - \lambda_3)/\lambda_1$
8	Scattering	How much neighbourhood variance can be explained by the smallest eigenvalue	λ_3/λ_1
9	Omnivariance	Volumetric point distribution	$\sqrt[3]{(\lambda_1.\lambda_2.\lambda_3)}$
10	Anisotropy	Directional dependence	$(\lambda_1 - \lambda_3)/\lambda_1$
11	Eigenentropy	Order/disorder	$-\lambda_1 \ln(\lambda_1) - \lambda_2 \ln(\lambda_2) - \lambda_3 \ln(\lambda_3)$
12	Sum EVs 3D	Sum of eigenvalues	$\lambda_1 + \lambda_2 + \lambda_3$
13	Curvature change	Local change in curvature	$\lambda_3/(\lambda_1 + \lambda_2 + \lambda_3)$
14	Z values	Absolute height of point	Z
15	KNN radius 3D	Size of the neighbourhood sphere	r_{knn-3D}
16	Density 3D	Points per m^3	$k + 1/(4/3.\pi.r_{knn-3D}^3)$
17	Verticality	The difference from vertical of the Z component of the normal vector	$1 - n_z$
18	Change in Z	Maximum height difference	$Z_{max} - Z_{min}$
19	STD of Z	Standard deviation of heights	$\sigma_{Z,knn-3D}$
20	KNN radius 2D	Size of the neighbourhood circle	r_{knn-2D}
21	Density 2D	Points per m^2	$k + 1/\pi.r_{knn-2D}^2$
22	Sum EVs 2D	Sum of eigenvalues from 2D structural tensor	$\lambda_{1-2D} + \lambda_{2-2D}$
23	EV ratio 2D	Ratio of the 2D eigenvalues	$\lambda_{2-2D}/\lambda_{1-2D}$
24	2D map	Frequency accumulation map	n/a
25	D_Z	Change in Z in accumulation map	n/a
26	Std_Z	Standard deviation of Z in accumulation map	n/a
27	EV$_{3d-1}$	First 3D eigenvalue	λ_1
28	EV$_{3D-2}$	Second 3D eigenvalue	λ_2
29	EV$_{3D-3}$	Third 3D eigenvalue	λ_3
30	EV$_{2D-1}$	First 2D eigenvalue	λ_{1-2D}
31	EV$_{2D-2}$	Second 2D eigenvalue	λ_{2-2D}
32	Mean_I	Mean intensity	$\sum \dfrac{(i_1 + i_2 \ldots + i_k)}{k}$
33	FPFH$_1$	FPFH value from bin number 1	n/a
-	- - -	- - -	- - -
65	FPFH$_{33}$	FPFH value from bin number 33	n/a

3.3.2 Classification

Once the data has been transformed into meaningful features it can be classified into categories using a variety of machine learning techniques. However, prior to classifier training several pre-processing steps must be carried out to improve the machine interpretability of the data. For the problem of finding rock bolts, the classifier is trained on large hand-labelled datasets where less than 1% of the observed points are rock bolts. If this data were directly used for training, even if the classifier always predicted 'not-bolt' it would achieve 99% accuracy. Of the several possible methods of class rebalancing discussed in section 2.4.6, the one chosen for this study is down-sampling the majority class. Empirical testing on the cross-validation data found a full downsampling (99% reduction) to match the minority class is not as effective as a less severe 80% reduction of the majority class. After downsampling, each feature is standardised by removing the mean and scaling to unit variance. The final classifier inputs are now a collection of m vectors of dimension n where m corresponds to the number of laser scanned points and n is the number of features in the feature set.

For learning the point representations, Weinmann et al. (2015a) tested many of the most popular types of classifiers including instance based, rule based, probabilistic, max-margin, ensemble and a simple neural network. They found that the ensemble method random forest performed best, which was the method also chosen by Chehata et al. (2009), Niemeyer et al. (2014), Landrieu et al. (2017b) and Hackel et al. (2017b). For our study, a preliminary test was carried out using multiple machine learning classifiers including Random Forests (RF), Multi-Layer Perceptron (MLP), Support Vector Machines (SVM), Quadratic Discriminant Analysis (QDA), Linear Discriminant Analysis (LDA) and Naive Bayes (NB), results shown in Table 3-2. The Linear and Quadratic Discriminant Analyses, along with the Naive Bayes proved unable to effectively classify the bolt points and were not considered further. When comparing the remaining three classifiers, the Random Forest produced higher accuracies on the minority bolt class than the Support Vector Machines; these results agree with those found by Bassier et al. (2019), Kogut and Weistock (2019) and Weinmann et al. (2015a). However, the MLP outperformed both the SVM and the RF, this is in contrast to the results observed by Bassier et al. (2019) and

Weinmann et al. (2015a). It is hypothesised that this difference may be due to the larger number of hyperparameters required to produce a stable result from the MLP classifier, as discussed by Nygren and Jasinski (2016). Based on this initial testing, the classifiers chosen for this work were the Random Forest and the MLP. The Random Forest was chosen as it is one of the highest performing classifiers in the literature and has been proven to be capable of achieving robust high accuracy classifications for problems of this type. The MLP was chosen as it showed the best performance in the initial tests and indicated strong generalisation potential when paired with appropriate hyperparameters.

Table 3-2: Results from the preliminary classifier testing

Classifier	Precision	Recall	F1
NB	0.08	0.97	0.14
QDA	0.25	0.94	0.40
LDA	0.48	0.79	0.60
SVM	0.81	0.60	0.69
RF	0.87	0.79	0.83
MLP	0.82	0.83	0.83

A random forest is a powerful machine learning algorithm based on a randomised forest of decision trees (Breiman, 2001). It has a low number of hyperparameters to tune and is resilient to noise in the data, making it an appropriate choice for remote sensing applications (Pal, 2005). An additional benefit of the random forest classifier is the ability to output a feature importance ranking, allowing for the relative contribution of individual features to the final prediction result to be observed (Strobl et al., 2008). The second classifier, an MLP or fully connected neural network, is a node-based architecture which can approximate complex functions by learning weights for every node by a process known as backpropagation (Hecht-Nielsen, 1992). Recent advances in processing power and vast dataset sizes have led to deep learning networks many hundreds of layers deep performing increasingly complex tasks (LeCun et al., 2015).

The structure chosen for the neural network used in this research is informed by the concept of effective capacity. A deep learning algorithm's effective capacity is its ability to model complexity; good performance is achieved when its effective capacity is appropriate for the complexity of the task and the size of the available training data (Goodfellow et al., 2016). If it has more effective

capacity than needed, it will tend to overfit. The task of finding bolt points from multi-dimensional feature vectors requires a relatively small effective capacity, as there are limited generalisation requirements. Combined with the small bespoke training set, an appropriate starting point for the structure was defined as containing no more than three hidden layers with no more than 40 nodes in each layer. Empirical testing was then carried out using a variety of values within this parameter space; stable, effective performance was obtained when the network contained two hidden layers with between 20-30 nodes in the first layer and 5-10 in the second layer. The final chosen structure contained 25 nodes in the first hidden layer and 5 nodes in the second hidden layer.

To decrease processing time, a Principal Component Analysis (PCA) dimensionality reduction (Wold et al., 1987) is performed on the data prior to input, reducing the features from 65 to 40 whilst maintaining 99.4% of the variance. These 40 features are then used as the input to the neural network and are joined to every neuron in the first hidden layer by a weight, with the value of the neuron being the weighted sum of all the features, transformed by the non-linear ReLU function. The second hidden layer has the same structure, with every neuron in each layer connected by weights, and the final output is a binary ('bolt' or 'not-bolt') decision. The network learns by backpropagation using the L-BFGS solver. Both classifications were carried out using the Scikit-learn libraries in Python (Pedregosa et al., 2011).

During model training, suitable values for hyperparameters of the classifiers were determined using a dual strategy. Firstly, a randomised search of the probable value space was carried out, using the Scikit-learn model selection tool 'RandomisedSearchCV' (Pedregosa et al., 2011). Taking the results of this search, empirical testing was then carried out above and below the best random search values to determine the exact hyperparameters choice. This hyperparameter tuning was carried out on the cross-validation section of the dataset via two-fold cross-validation. For the random forest, it was found that only the 'number of estimators' hyperparameter affected the results to any appreciable degree. Therefore, to ease repeatability, the random forest hyperparameters were all kept at the Scikit-learn default values except for the 'number of estimators' hyperparameter which was changed to 200.

The neural network hyperparameters examined included the solver, the activation function and the L2 regularisation term. There was no appreciable difference in accuracy observed from using different solvers, however, the LBFGS converged faster and required fewer additional hyperparameters. Figure 3-7 shows the results from the empirical testing of the L2 regularisation term and activation function, showing that the best accuracies are obtained with an L2-regularisation term of 1e-4 and the ReLU activation function.
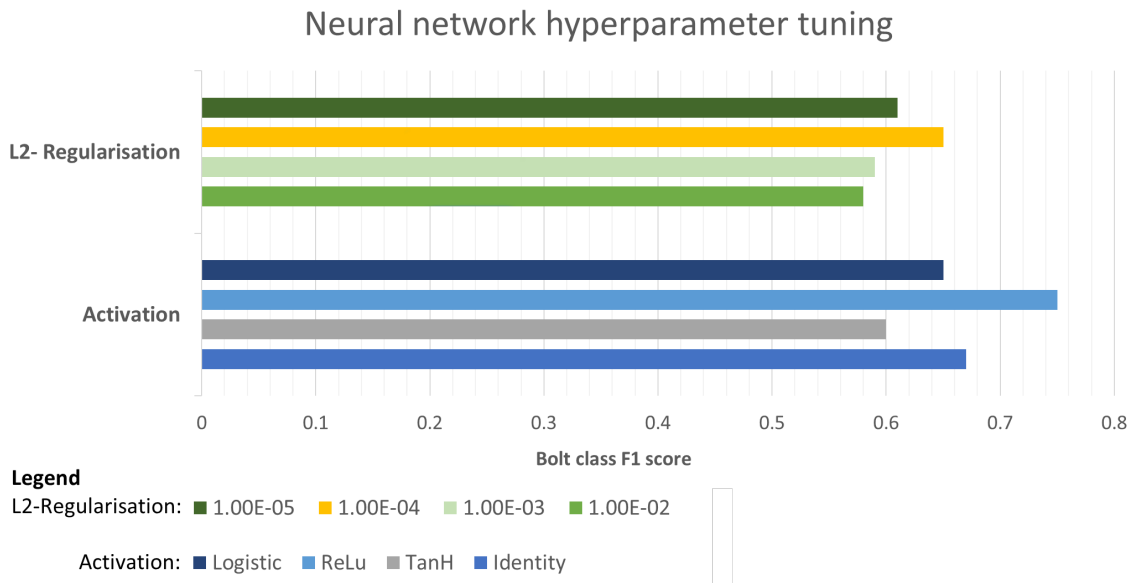
Figure 3-7: Results from the neural network manual hyperparameter tuning.

3.3.3 Object creation

The type of machine learning used in this research acts on the features derived for each individual point in the cloud. Because there is no spatial connectivity, they suffer from noise due to isolated misclassified points. In the processing pipeline, after the point wise classification, the resulting point cloud is split using the predicted values and the points that have been labelled as 'not-bolt' are now discarded, greatly reducing the dataset size. The remaining cloud now contains all the correctly predicted bolt points and the falsely predicted non-bolt points. From visual examination of this remaining cloud, it can be seen that the point cluster separation is good, with adequate empty space visible between the clusters of predicted points.

Cloud segmentation was carried out using DBSCAN (Density-Based Spatial Clustering of Applications with Noise). This algorithm finds core samples and generates clusters from high density areas adjacent to them, allowing for

clusters of any shape (Ester et al., 1996; Schubert et al., 2017). The maximum distance between neighbourhood samples parameter (ε) was set to 5cm and the minimum cluster size was set to 10 points. The ε value was chosen based on the heuristic proposed by Ester et al. (1996) of a suitable value being approximately the distance to the 4th nearest neighbour, in this case 5cm for the 1cm resampled point cloud. The minimum cluster size was set to 10 points; as the ground truth bolt clusters contained between 20-400 points a number set at 50% of the sparsest bolt cluster was a suitable choice of parameter. The Euclidean distance metric was used as the inputs were coordinates in 3D space and the K-D tree algorithm was used to compute the neighbours as the data dimensionality is low.

The next processing step was to calculate the centroid of each cluster to use as the predicted bolt location. The final step was to export these cluster centroids as a X, Y, Z file of only a few kilobytes that can be easily shared with machines and surveyors. This clustering greatly reduces the algorithm's sensitivity to misclassifications in the individual points. Provided at least 10 points from a bolt have been classified correctly the bolt will be detected, reducing missed detections.

3.3.4 Generating additional value

To further demonstrate the applications of automated recognition algorithms an additional post processing workflow was developed to generate bolting quality assurance reports. This was carried out by performing a nearest neighbours search on the extracted centroids to determine how many other bolts are within a user defined distance of each bolt; bolts with very few neighbours can then be flagged as requiring further investigation from the geotechnical team to determine whether sufficient bolting has been carried out in that area. Figure 3-8 shows the results of the nearest neighbour search on the production mine test dataset. Text reports detailing the bolt location and the distances to the nearest k neighbouring bolts can also be produced. If these reports could be linked to the rock bolting machine's own records this would provide comprehensive reporting on the as-built bolting pattern installed in the mine. Other potential applications could include recording spot bolting locations for

geotechnical reference and for linking the onboard hole information recorded by bolting machines to real world coordinates from georeferenced laser scans.

Figure 3-8: A visualisation of the bolt density reporting. The red bolts have less than 1 other bolt within 1.5 meters and have been flagged for further inspection.

3.4 Results

The performance of the proposed methodology was assessed on both the raw point prediction accuracy and also on the number of bolts correctly detected. The metrics used for evaluation are precision and recall and F1 score. These metrics are derived from the relationships between True Positives (correctly detected pixels), False Positives (incorrectly detected pixels) and False Negatives (undetected pixels). These metrics were chosen as others such as the overall accuracy are inadequate in cases such as this, where large class imbalances are present in the data. The precision is defined as the measure of what proportion of the positive predictions are correct; it is the number of True Positives divided by the number of True Positives and False Positives. The recall is a measure of what proportion of actual positives were correctly identified; it is defined as the number of True Positives divided by the number of True Positives plus False Negatives. The F1 score is the harmonic mean of the precision and recall.

96

3.4.1 Cornwall mine results

The first experiment, carried out using the Cornish dataset, tested which set of point feature descriptors provided the most accurate results. It compared the full geometric feature set proposed by Weinmann (2016) consisting of 26 features, a reduced version of this feature set with the accumulation map features removed (23 features), the FPFH features (33 bin histogram), the combined feature sets (59 features) and finally the combined features plus the intensity features (61 features). Table 3-3 shows the results of the feature set comparison on both classifiers, with the F1 score used as the performance metric. For this test, the PCA reduction was not carried out on the neural network dataset to more clearly isolate the effect that feature sets have on the results. The random forest classifier also outputted the feature importance rankings, shown in Figure 3-11 and discussed in Section 5. As the combined features with intensity achieved the highest accuracy, this was the feature set used for the final model which was applied to the unseen test data.

Table 3-3: F1 scores for differing feature sets evaluated on the Cornish mine dataset. The reduced geometric features refer to the set with the 3 highest computation time features (accumulation maps) removed.

Feature set	Geometric features full	Geometric features reduced	FPFH features	Combined features	Combined features and intensity
No. features	*26*	*23*	*33*	*59*	*61*
Neural network	0.42	0.41	0.51	0.63	0.64
Random forest	0.49	0.43	0.37	0.56	0.58

Once the feature set choice was finalised, the per point prediction results were examined against the human generated ones for the Cornish test data, totalling almost 1.5 million point predictions. These results are given in Table 3-4. Figure 3-9 gives a graphical view of the neural network's point prediction results. In this figure the footwall has been removed and the viewing angle is directly vertical towards the hanging wall. The predicted bolt points are shown in red and the overlaid white squares show the true bolt locations. Where that the red points do not have a corresponding white square overlay indicates incorrect objects classified as bolts, and any white squares without corresponding red points indicate missed bolts.

Table 3-4: Results from the point-wise classifiers on the Cornish test dataset.

Neural network	Predict not-bolt	Predict bolt		Precision	0.59
Not-bolt	1471791	6586		**Recall**	0.70
Bolt	4071	9370		**F1 score**	0.64
Random forest	*Predict not-bolt*	*Predict bolt*		**Precision**	0.72
Not-bolt	1475540	2837		**Recall**	0.38
Bolt	6809	6632		**F1 score**	0.58
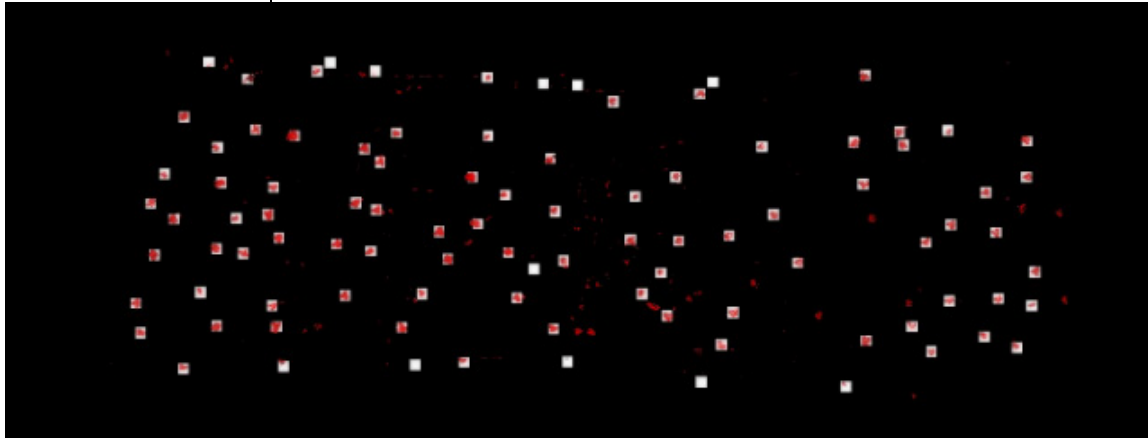
Figure 3-9: Graphical view of the neural network point cloud classification. The red points are those that the classifier predicts are bolts and the white boxes indicate the actual bolt locations.

As can be seen in Table 3-4, the results, whilst overall positive still contain many misclassified points. To investigate whether the DBSCAN clustering can extract individual bolt object locations to a greater degree of accuracy, the extracted centroids were overlaid with the 101 true bolt centroids and the number of true positives, false positives and false negatives were counted. For this test, the bolt was classed as detected if the human generated and machine generated centroids were within the nominal bolt faceplate radius size of 8cm of each other. These results are given in Table 3-5.

Table 3-5: Results of bolt detection algorithm on the Cornish test dataset.

Neural network	Predict Not Bolt	Predict Bolt			
Not Bolt	n/a	6		**Precision**	0.94
Bolt	13	88		**Recall**	0.87
				F1	0.90
Random forest	*Predict Not Bolt*	*Predict Bolt*			
Not Bolt	n/a	3		**Precision**	0.95
Bolt	46	55		**Recall**	0.54
				F1	0.69

3.4.2 Production mine results

To test the approach developed in this chapter in a new environment, the data from the modern production mine was run through the same pipeline as the Cornish dataset. As the extended eigen features are not optimised to be

deployed in the mining cycle without impacting production constraints, this test used only the reduced geometric feature set of 23 features, combined with the 33 FPFH features. Four scenes (of 2-4 scans each) were used for training and a 5th scene was used to assess the performance. To address the many confusion objects present in the scans which might not be captured by the majority resampling, during the 4 fold cross validation the confusion 'not-bolt' objects were extracted from each of the training datasets. These were then combined into an additional 'confusion' point cloud which was added to the training data without resampling. The pointwise results are given in Table 3-5 and the object based results in Table 3-7. Figure 3-10 displays an example of the predicted results.

Table 3-6: Results from the point-wise classifiers in the production test dataset.

Neural network	*Predict not-bolt*	*Predict bolt*	**Precision**	0.61
Not-bolt	2253747	10624	**Recall**	0.78
Bolt	4748	16402	**F1 score**	0.68
Random forest	*Predict not-bolt*	*Predict bolt*	**Precision**	0.76
Not-bolt	2260288	4083	**Recall**	0.60
Bolt	8536	12614	**F1 score**	0.67

Table 3-7: Results from the bolt detection algorithm on the production test dataset.

Neural network	*Predict not-bolt*	*Predict bolt*	**Precision**	0.83
Not-bolt	n/a	35	**Recall**	0.96
Bolt	7	170	**F1 score**	0.89
Random forest	*Predict not-bolt*	*Predict bolt*	**Precision**	0.89
Not-bolt	n/a	17	**Recall**	0.80
Bolt	35	142	**F1 score**	0.84
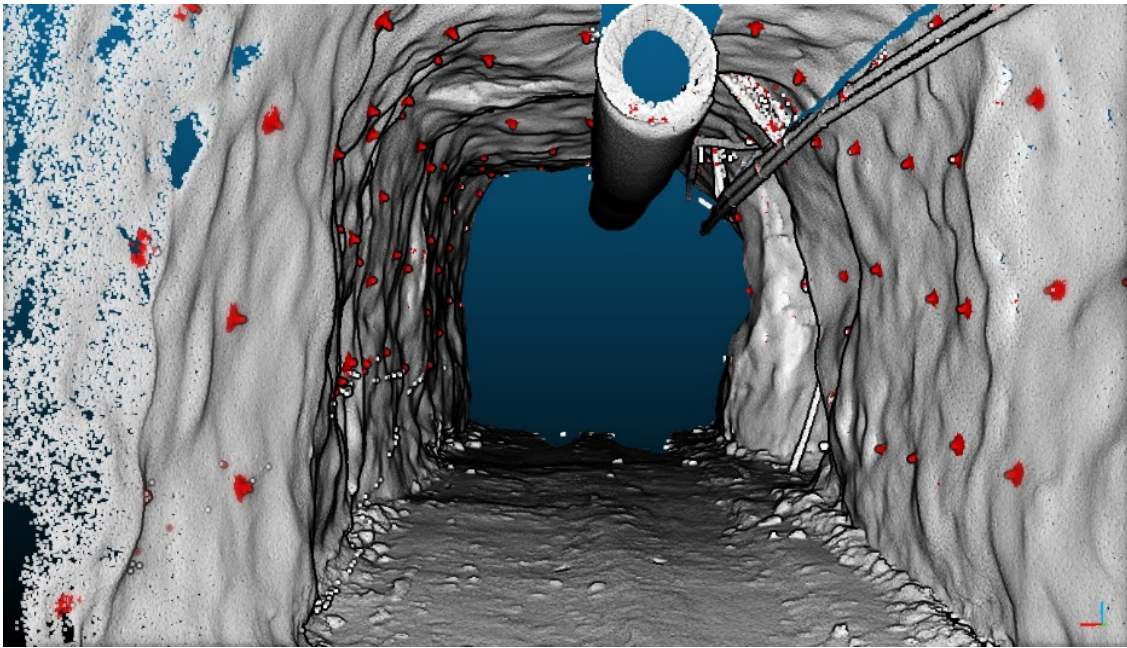
Figure 3-10: Results from the production mine test dataset. Clear bolt detections are visible, with only small amounts of noise present at the rim of the vent bagging and along the pipe hanging wall strop.

3.5 Discussion

The feature set test shows that the combined feature sets are more effective than either the geometric or FPFH based feature sets applied separately. Using only the geometric feature set, the random forest outperforms the neural network; this agrees with the results obtained by Weinmann et al. (2015a) using the same feature types. Using FPFHs the random forest scores relatively poorly, though the combination does still improve on the score recorded from just the geometric feature set. These results infer that the addition of the FPFH features does contribute to the overall accuracy of the random forest, but that they are less important than the geometric features. To examine the feature contributions further, the feature importances were calculated using the Gini importance method. This technique measures how much the Gini impurity is reduced when using a particular feature, averaged across all trees in the forest (Géron, 2017). The feature importances are then normalised so that the sum of all importances equals one. Figure 3-11 graphs the feature importances across the classification vector, this shows that the more important features are primarily from the geometric set, though several from the FPFH set also score highly. The highest ranked features (above 0.05) are scattering, absolute height, mean intensity and anisotropy.

100
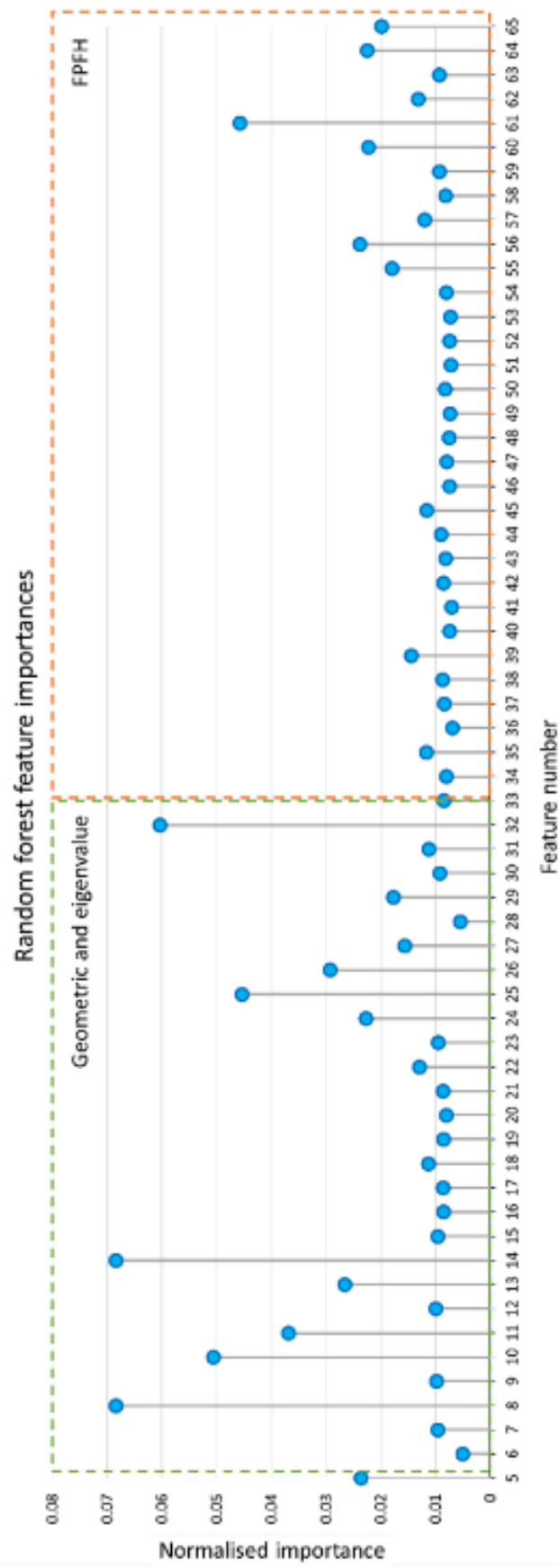
Figure 3-11: Graph showing the individual feature importances for the random forest classifier. The green box indicates the geometric and eigenvalue based features and the amber box indicates the FPFH features. For details on feature numbers see Table 1.

101

The neural network classifier cannot output a feature importance ranking; however, from examining the results it appears that the neural network is utilising more of the FPFH set features, as this was the highest non-combined score for all classifier and feature set combinations. The intensity features provided an improvement of 0.02 to both classifiers' scores; these intensity features are some of the simplest to compute and are therefore a strong addition to the feature sets.

The point-wise results are positive despite outlier misclassifications. This is due to the challenging dataset and the many confusion objects presenting geometry similar to a bolt object. Most importantly, they contain enough positively identified points to enable the DBSCAN algorithm to detect the actual bolt objects, as demonstrated by the precision scores of 0.87-0.96. Primarily, the incorrectly identified bolt points (false positives) occurred as isolated points, allowing them to be easily removed by the clustering operation. Only rarely (< 10% of Cornish dataset bolt detections), as in the instance of pipe mounting steelwork which closely resembles a bolt, did the algorithm misclassify enough points in close proximity to create a false positive cluster, as seen in Figure 3-12, where the cluster inside the red box is large enough to make it through the DBSCAN stage. The isolated incorrect points visible on the hanging wall in Figure 3-12 will all be removed by the DBSCAN process.
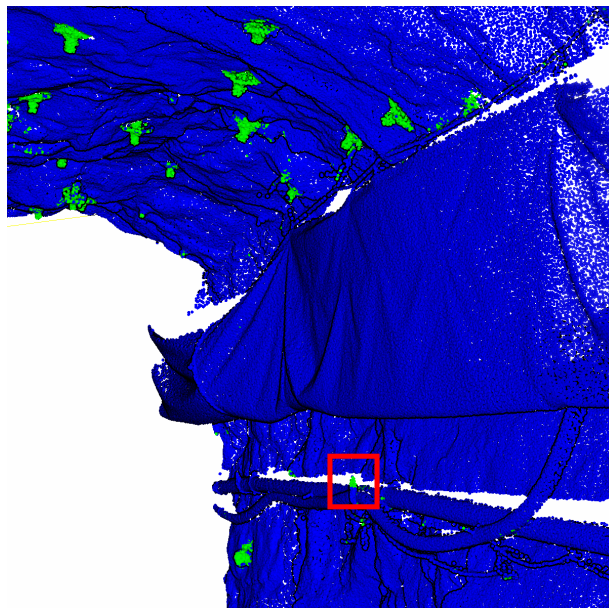
Figure 3-12: Instance of misclassified cluster of points by the random forest classifier. Blue points are predicted not-bolt, green points are predicted bolt and the red box indicates a piece of pipe mounting bracket incorrectly classified as a bolt.

At the object extraction stage, as both classifiers had only a few false positives, these were manually checked in the original highest resolution scan data to determine if there was in fact a bolt present at that location which had been missed at the labelling stage. From this examination, it appeared that the neural network correctly identified 5 bolts from the Cornish dataset and 7 bolts from the production dataset which were badly scanned and highly obscured, leading them to be not picked up by the human operator at the dataset creation stage. This demonstrates the value of machine learning technologies for automated quality assurance and quality control as in these difficult cases the neural network surpassed the human inspector. The Cornish testing dataset was then used to estimate the level of label noise present in the training datasets. The test dataset label noise was ~5% at the cluster/object level (5 missed out of 101 total) and ~3% at the individual point level (471 missed out of 13,912 total). The figures are expected to be far lower for the training dataset as the mislabelled points are all in the 'not bolt' class, which has been randomly resampled to contain only 20% of its original points. Neural networks and random forests have been shown to be highly robust to label noise below 10% (Folleco et al. 2009, Pelletier et al. 2017), therefore, the small number of mislabelled points in the training dataset is not expected to have has a meaningful impact on the classifier training. Comparing the human result to the neural network, the human is still superior with a precision of 1 and a recall of 0.95; however, in a real world inspection case, the human takes much longer to identify the bolts, suffers from fatigue and still cannot detect every bolt. Figure 3-13 (a) and (b) show an example of a bolt missed by the human operator but found by the neural network and Figure 3-13 (c) shows an actual incorrect detection by the neural network.

The false negatives from the neural network also were examined, and it was found that in the Cornish dataset 10 out of the 11 missed detections were low bolts on the sidewall. From this, we can infer that the Z values and the relatively few sidewall bolts compared to roof bolts in the training data are influencing the model's decision making. The production dataset contained many more sidewall bolts for training and subsequently there were very few missed bolts, the recall value of 0.96 demonstrates that when furnished with representative training data the model can detect almost all bolts.
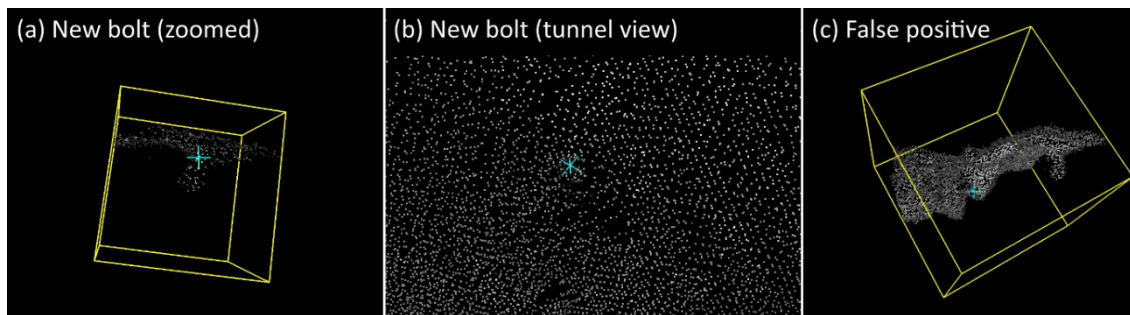
Figure 3-13: Examination of false negatives and false positives. (a) shows the false negative bolt zoomed and extracted to a specific angle and (b) shows how the false negative appears to a human in the full tunnel dataset. (c) shows a sharp discontinuity (false positive) that has been mistaken for a bolt by the classifier.

3.6 Summary

This chapter described a methodology to automatically detect supporting rock bolts from laser scan data. After the scans have been extracted from the instrument, the workflow is implemented entirely with open source software. The methodology is customised to the underground environment and improves upon previously published surface applications by utilising a larger feature set and robust clustering algorithms to address the challenges from noise, confusion objects and multiple Z values present in a typical underground mine environment.

The neural network classifier produced the strongest point-wise classification results, allowing the DBSCAN clustering algorithm to successfully locate the candidate bolt objects. Further work extended this approach to other mining datasets gathered with lower cost 3D laser scanners more closely resembling machine mounted scanners, verifying that the method presented here is suitable for a wide range of sensors and mining types.

An additional processing algorithm developed in this chapter used the bolt location output files to verify that bolting patterns that have been installed to specification and a technique to locate a machine based on bolter patterns has been demonstrated. Applications such as this offer mining companies valuable opportunities to embrace new technologies for improved productivity and safety in a digitally connected world.

Analysis of legacy mining landscapes from LiDAR data using deep transfer learning

Chapter overview

This chapter investigates how deep transfer learning can be applied to LiDAR data to automatically detect landscape signatures such as legacy mine shafts. These signatures can then be used to provide landscape insights, as a target for further investigation or as a proxy to infer other landscape features. This chapter develops techniques that can be used with 2.5D gridded LiDAR data at a range of resolutions. The research underpinning this chapter comes from three projects. The first project (4.3) arose from a research collaboration with Cornwall Consultants, financed by the Cornwall Aerospace Innovation fund. This project examines the different mining landscape signatures that can be detected in LiDAR data and explores how advanced visualisation techniques can be applied to enhance these signatures.

The second project (4.4) investigates how a deep convolutional neural network originally designed to detect craters on the moon can be repurposed through transfer learning to detect historic mine workings on Dartmoor. This research formed the core of the paper **'Bringing Lunar LiDAR Back Down to Earth: Mapping Our Industrial Heritage Through Deep Transfer Learning'** published in the journal Remote Sensing[8]. The final project in this chapter (4.5) is adapted from further work with the Dartmoor dataset, using automated fitting algorithms to infer mineralisation trends. This work has been published as a conference paper **'Using Deep Learning and Hough Transformations to Infer Mineralised Veins from Lidar Data over Historic Mining Areas'** in the International Society of Photogrammetry and Remote Sensing (ISPRS) Archives[9].

---

[8] The candidate is the first author of this paper and the authorship contribution statement is as follows: **Jane Gallwey:** Conceptualisation, investigation, methodology, writing (original draft & revisions), software, validation. **Matthew Tonkins**: Software. **Matthew Eyre:** Conceptualization, writing (review & editing), supervision. **John Coggan:** Project administration, writing (review & editing), supervision.

[9] This article was first authored by the candidate who wrote 80% of the paper, the 20% of the paper concerning geological information was written by the second author. The paper was revised based on co-author, supervisor and reviewer comments.

4.1 Introduction

Across the UK, there is a long history of mining and it is estimated that over 15% of land can be considered within mining areas (CIRIA, 2019). Mining creates significant changes to a landscape, the scale of which are determined by the mine size, mining type and commodity type. After a mine ceases production, over time, these changes become less visible as the landscape revegetates and features such as small spoil heaps erode. Until 1872, there was no requirement to record mine plans or secure workings, leading to large areas of land in the UK harbouring unrecorded historic mine workings (CIRIA, 2019). These workings create potential risks from subsidence, slope instability, fault reactivation and egress of mine water and gas. They also present a hazard to members of the public, particularly in areas such as West Cornwall and Dartmoor that today have extensive recreational usage. Remote sensing data can be used to rediscover these sites, helping to manage the risk of the mining legacy, alongside providing potential mineralogical information to future developers and adding to an area's historic record.

Some of the past mining landscape features are visible in aerial imagery; however, others are more subtle and are better observed as topographic changes in LiDAR datasets. There has been little published work on using LiDAR for identifying mining related landscape signatures. Work by Mcdonald (2011 & 2013) uses LiDAR for visualising abandoned coalfields in Ohio; however, the LiDAR data is only used as a hillshaded background layer, without any further post processing and manipulation. However, there is a large body of literature for using LiDAR to detect archaeological landscape signatures and many parallels can be drawn between these two tasks. Primarily, both are searching for small relative differences in 3D terrain which would indicate unnatural features; indeed many archaeological studies include mines as one of the features which they are attempting to locate using LiDAR (Fonte et al., 2014; Historic England, 2018).

Technological advances in LiDAR processing can be used to aid historic mining detection in two ways. Firstly, advanced methods of data visualisation and representation pioneered by the archaeological community can be used to enhance the data for human or machine interpretation, and secondly, new deep

learning and transfer learning algorithms can be used to automatically or semi-automatically search landscapes to detect signatures of interest. The three projects presented in this chapter examine many aspects of the current state of the art in LiDAR data analysis in relation to historic mining landscapes. These include multiple data representations, automated detection models and inferred analysis from automated detections.

4.2 Datasets

Across England there are several publicly available LiDAR datasets which could be used for mining signature detection. The Environment Agency oversees the dissemination of these datasets through the Department for Environment, Food and Rural Affairs (DEFRA) data services portal[10], however, the datasets themselves have been collected by multiple parties for diverse projects. Figure 4-1 shows the availability of different scales of data across England. The highest resolution data is 0.25m with limited coverage, followed by 0.5m across some larger areas and 1m datasets with almost complete coverage. The Environment Agency has pledged full coverage of England at 1m resolution by 2020[11] (Winter, 2017).
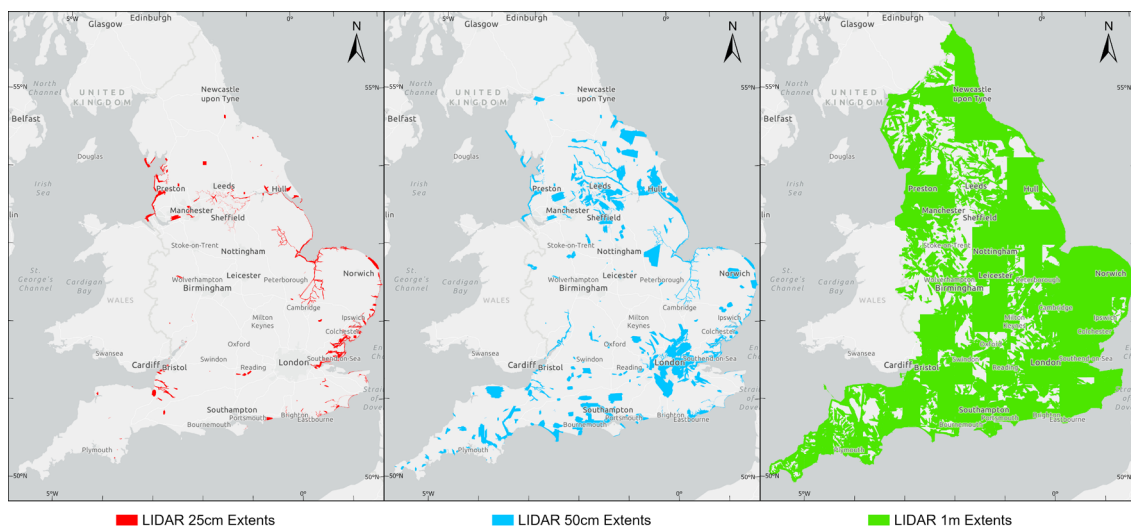
Figure 4-1: Environmental Agency LiDAR coverage across England at 25cm (left), 50cm (middle) and 1m (right).

The most extensive dataset covering the mining regions of the South West of England was collected as part of the Tellus South West project in 2014. This

---

[10] Available at https://environment.data.gov.uk/DefraDataDownload/?Mode=survey
[11] Now delayed to 2022

data has a resolution of approximately 1 point per meter (PPM) and a vertical accuracy of 25cm. It is supplied as either a DSM or DTM in raster format with a grid size of 1m. The data collection was funded by the Natural Environment Research Council and the full dataset is managed by the British Geological Survey (BGS) and the Centre for Ecology and Hydrology as an open access dataset (Yeomans, 2017), alongside being made available through the DEFRA portal.

Many of the mining remains prevalent across the landscape of the South West are small and obscured by thick vegetation, making detection from a coarse 1m dataset such as Tellus challenging. To investigate whether higher resolution data would improve performance, sample areas from the 0.5m and 0.25m scale datasets were also examined. The 0.5m dataset covers the mining areas of Dartmoor and was flown in 2009, the 0.25m dataset covers the Grassington mines of Yorkshire and was flown in 2012. Across all scales, the DSM was chosen in preference to the filtered DTM due to concerns that the filtering algorithms used to produce the DTM can excessively smooth small features (Haslam and Howard, 2017).

4.2.1 St Just area

The Cornish test area used for the visual landscape signature analysis and initially trialled for the deep learning algorithms is the area around the Balleswidden Mine of St Just, West Penwith, Cornwall. This mine was one of the largest in the district in the 1800s with more than 10 major shafts providing access to extensive underground workings. Rowe and Foster (1887) recorded that it produced 11,828 tons of black tin between 1837 and 1873. Underground activity ceased in 1877, however, large sections of the waste tips were removed and reprocessed between 1913-1916, further modifying the landscape in this area (Dines, 1988)
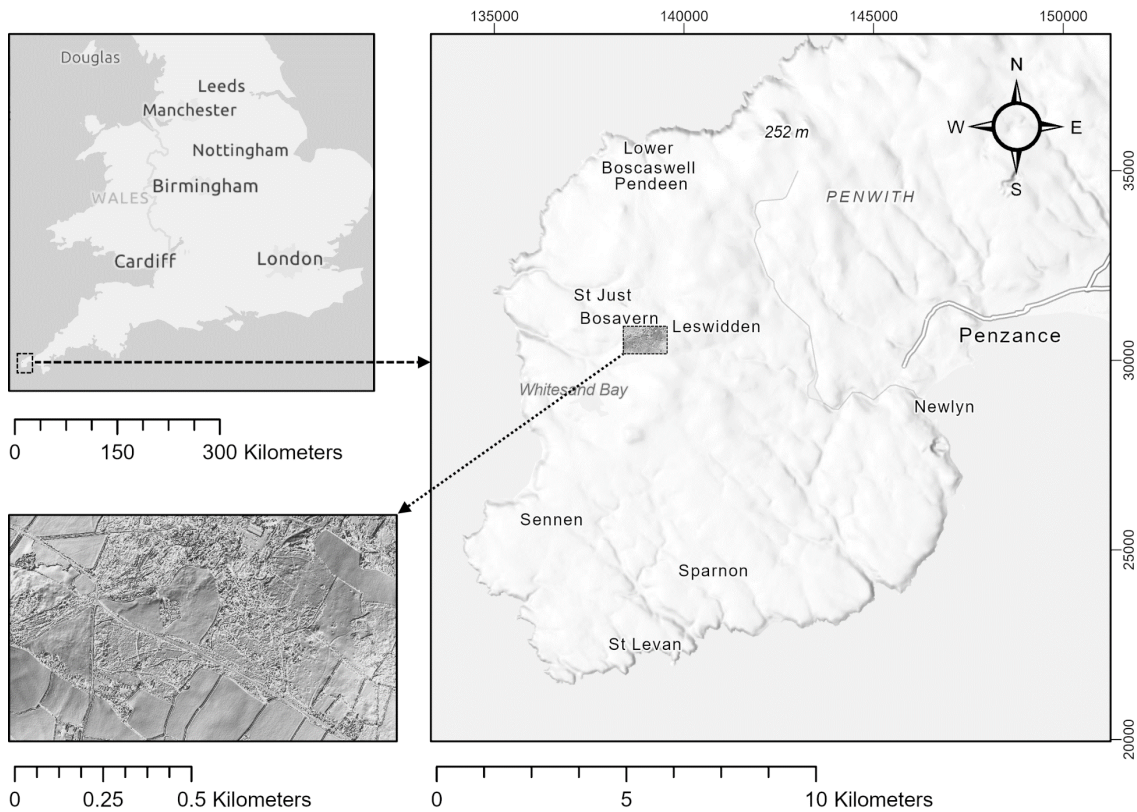
Figure 4-2: Overview map of the St Just mining test area.

4.2.2 Dartmoor area

The primary study area for the deep learning section of this research is Dartmoor National Park, an upland area of moorland studded with exposed granite hilltops known as tors. The ground cover is primarily low vegetation, including heather, bracken, gorse, fern and marsh grasses. Tin and copper mining on Dartmoor has taken place almost continuously from the 12th to the 20th centuries and the remains are pervasive and visually striking throughout the landscape (Newman, 2010). Three areas of concentrated historic mining activity were used to develop this deep learning model; these are shown in Figure 4-3. The different colours in Figure 4-3 refer to the distribution of training, validation and testing data.
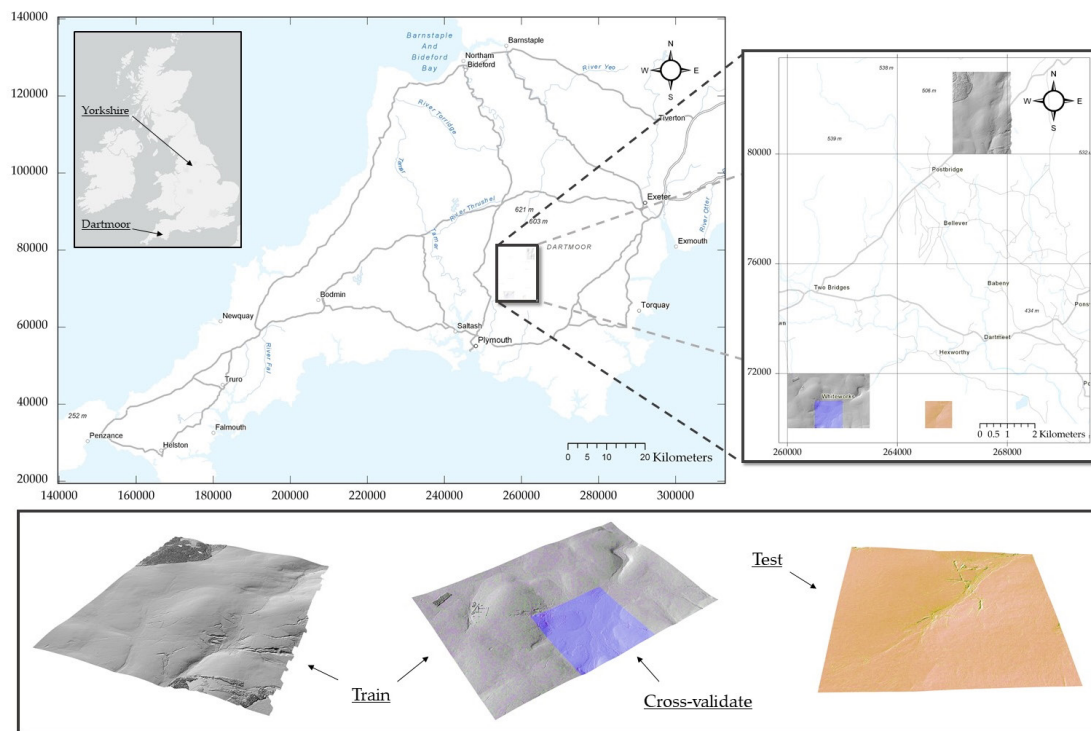
Figure 4-3: Overview of the Dartmoor dataset. Grey areas represent training data (14 tiles), the purple tile shows the cross validation area and the orange tile shows the test area. Coordinate system British National Grid, image data © Environment Agency 2015 & Getmapping Plc. Basemap © ESRI 2019

The training and validation areas include in the north the old Birch Tor Mine (1726-1928) (Dines, 1988) and in the south the former Whiteworks Mine. It is believed that the Whiteworks area was being mined as early as 1180 although the mine was expanded substantially around 1790 towards the beginning of the industrial revolution when the demand for tin increased (Dines, 1988). The mine was owned by the wealthy Tavistock mining entrepreneur Moses Bawden and operated for just under 100 years until 1880, briefly reopening in early 1900 before finally closing for good by 1914 (Hamilton Jenkin, 1974). The test area for Dartmoor is the site of Hexworthy Mine (1891-1912). This is an interesting site as it displays remains from multiple eras of mining; from the early unrecorded openworkings, through traditional 19th century mining to semi-modern 20th century workings (Richardson, 1992). The mine operated productively until the call up for men in 1914, during the war it was placed in care and maintenance before a large storm in 1920 destroyed the waterwheel flume, causing the underground workings to flood (Hamilton Jenkin, 1974).

4.2.3 Yorkshire area

A further testing area for the deep learning model was selected in the Yorkshire Dales National Park more than 500km from Dartmoor to examine the model's ability to generalise to new locations, mine types and data resolutions. This test area is part of the site of the former Grassington Moor lead mine and is shown in Figure 4-4. The first known exploitation of lead at Grassington was by the 4th Earl of Cumberland in the early 17th century, although it is thought that some primitive extraction and smelting had taken place earlier. The early exploitation involved the digging of shallow shafts along the vein. The first mill to process the Grassington lead ore was the Low Mill built in 1605. The test area covers the western part of the Yarnbury mine, including Tomkins, Barretts and Good Hope shafts (Northern Mine Research Society, 1980).
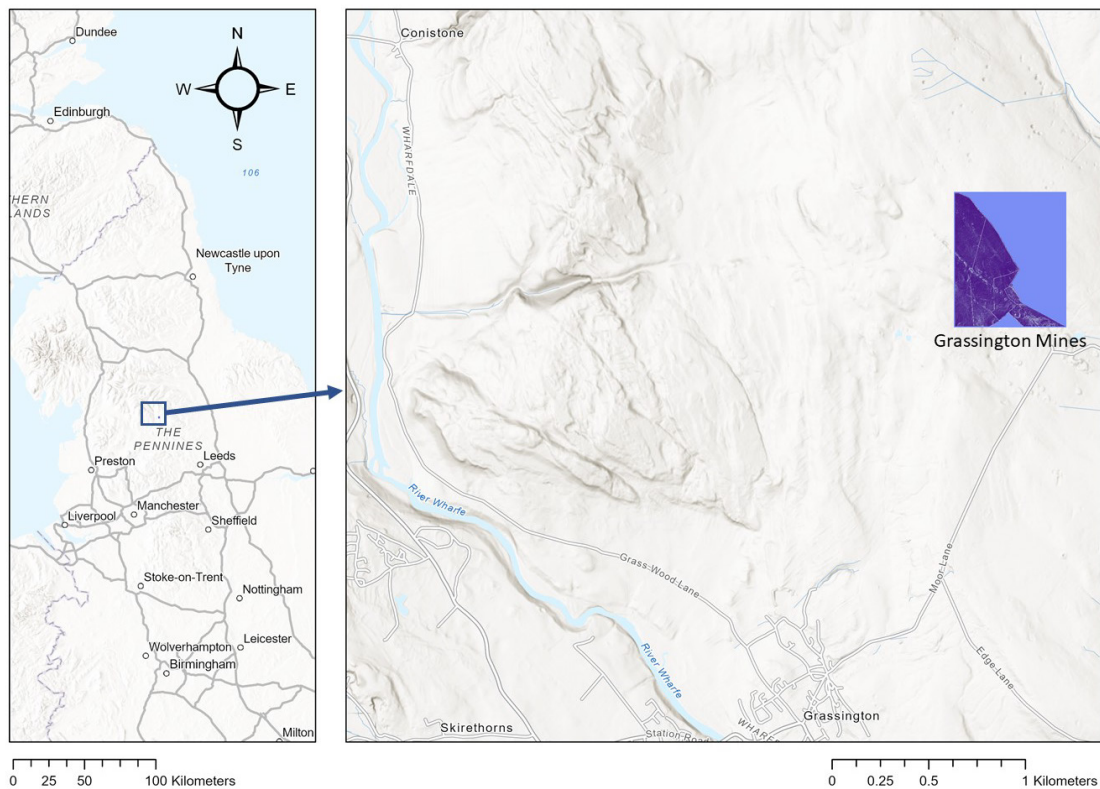
Figure 4-4: Yorkshire test area. Basemap © ESRI 2019.

4.3 Advanced visualisations – project one

In its simplest form, raster LiDAR data is a pixel coordinate and an elevation value for each cell. If these elevations are normalised to the range 0-255 an 8-bit greyscale image can be produced; however, these images often lack

contrast and can be difficult to discern landscape signatures from. There are many advanced visualisation processing techniques used by archaeologists that can improve the interpretability of LiDAR data; an excellent reference for these is the document 'Airborne Laser Scanning Raster Data Visualisation: A Guide to Good Practice" (Kokalj and Hesse, 2017). This section gives an overview of the main types of advanced visualisations relevant to historic mining hazard identification, along with examples of each visualisation generated for the mining test area. Figure 4-5 shows a satellite image of this test area, overlaid with many types of historic mining features identified by Cornwall Consultants from their existing database. This area was chosen as it contains known lodes, shafts and pits which are visible to some degree on the LiDAR data.
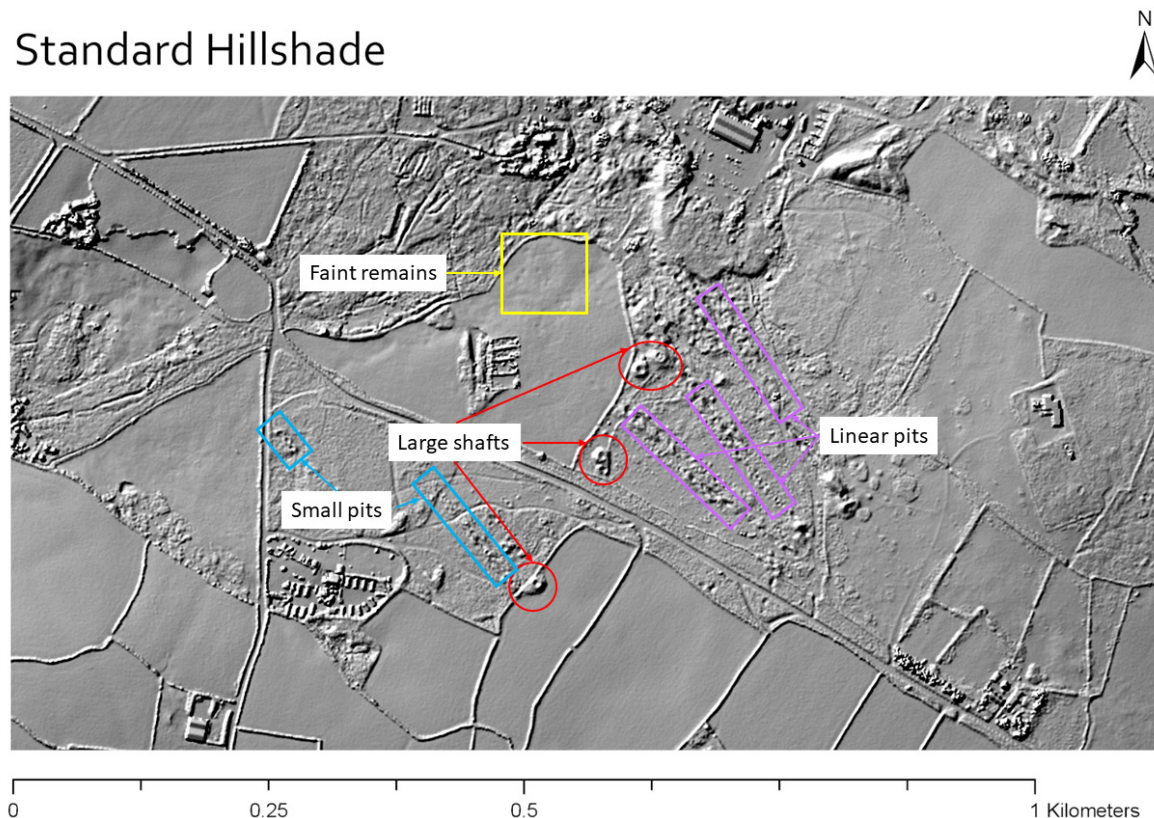
Figure 4-5: Overview of the Balleswidden Mine demonstration area showing some of the visible mining landscape features.

The visualisations shown here are performed on the freely available Tellus 1m dataset described in Section 4.2, following the workflows described in Kokalj and Hesse (2017). This dataset was chosen as the UK has broad coverage at a 1m resolution level, allowing methods developed in this project to be applied nationwide. Throughout this section, it can be seen that many of the features

that show up in the various LiDAR visualisations can be attributed to old field boundaries, unmarked paths and other non-mining related causes. For interpretation tasks carried out by a human, the LiDAR data is examined in conjunction with other datasets in order to identify any LiDAR abnormalities that cannot be attributed to other causes.

4.3.1 Techniques and discussions

The LiDAR representation which most closely resembles how a human eye perceives relief is known as hillshading, shown in Figure 4-6a. It is created by virtually illuminating the scene from a defined azimuth angle and sun elevation. Any features facing towards the light will be illuminated, while features facing away will be shadowed. To accentuate faint features in flat terrain a very low sun angle can be used as seen in Figure 4-6b. From examining the hillshades of the mining area manually it can be seen that whilst they provide an easily interpretable view of the landscape the mining features are not particularly prominent.

A potential issue with all single illumination source hillshades is that features parallel to the light source will not be discernible and that features facing away from the light may be too deeply shadowed to interpret. A method to negate this is to create composite multi-directional hillshades by combining multiple sun angle hillshades into a single image using 3 different hillshades mapped to the RGB channels to generate a colour composite image, shown in Figure 4-7a. Another method of visualising the results of multiple hillshades is Principal Component Analysis; because each different view is of the same scene, the results are highly correlated and the first 3 principal components can be mapped to the RGB channels to produce a false colour image showing the three primary variances between the multiple hillshades (Kokalj and Hesse, 2017), as displayed in Figure 4-7b.

a) Standard Hillshade

N

b) Low Sun Angle Hillshade

N

0   0.25   0.5   1 Kilometers
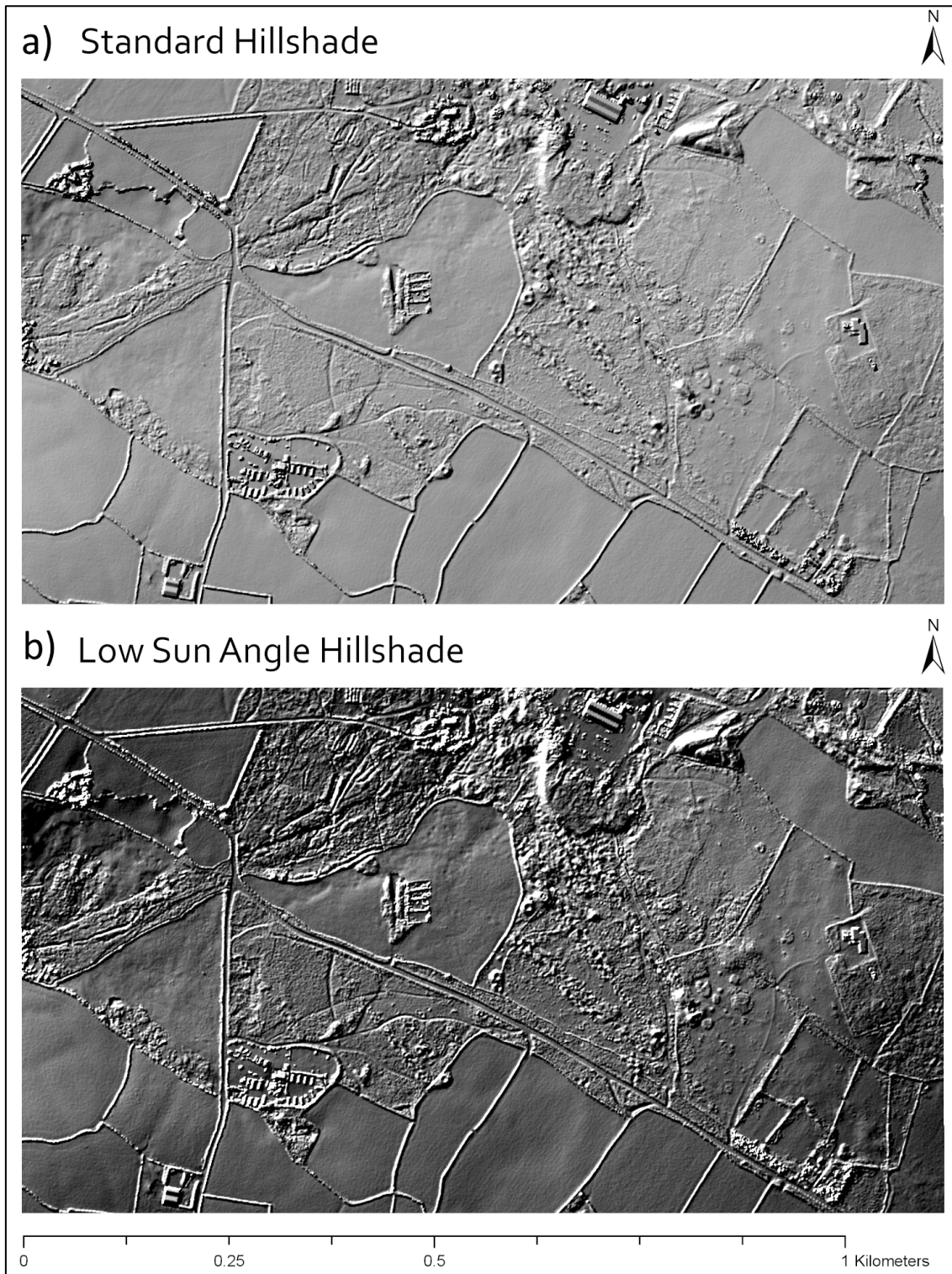
Figure 4-6: Hillshading from differing sun angles. Both images are shaded from a 315 degree azimuth, a) from a 35 degree sun angle and b) from a 5 degree sun angle.
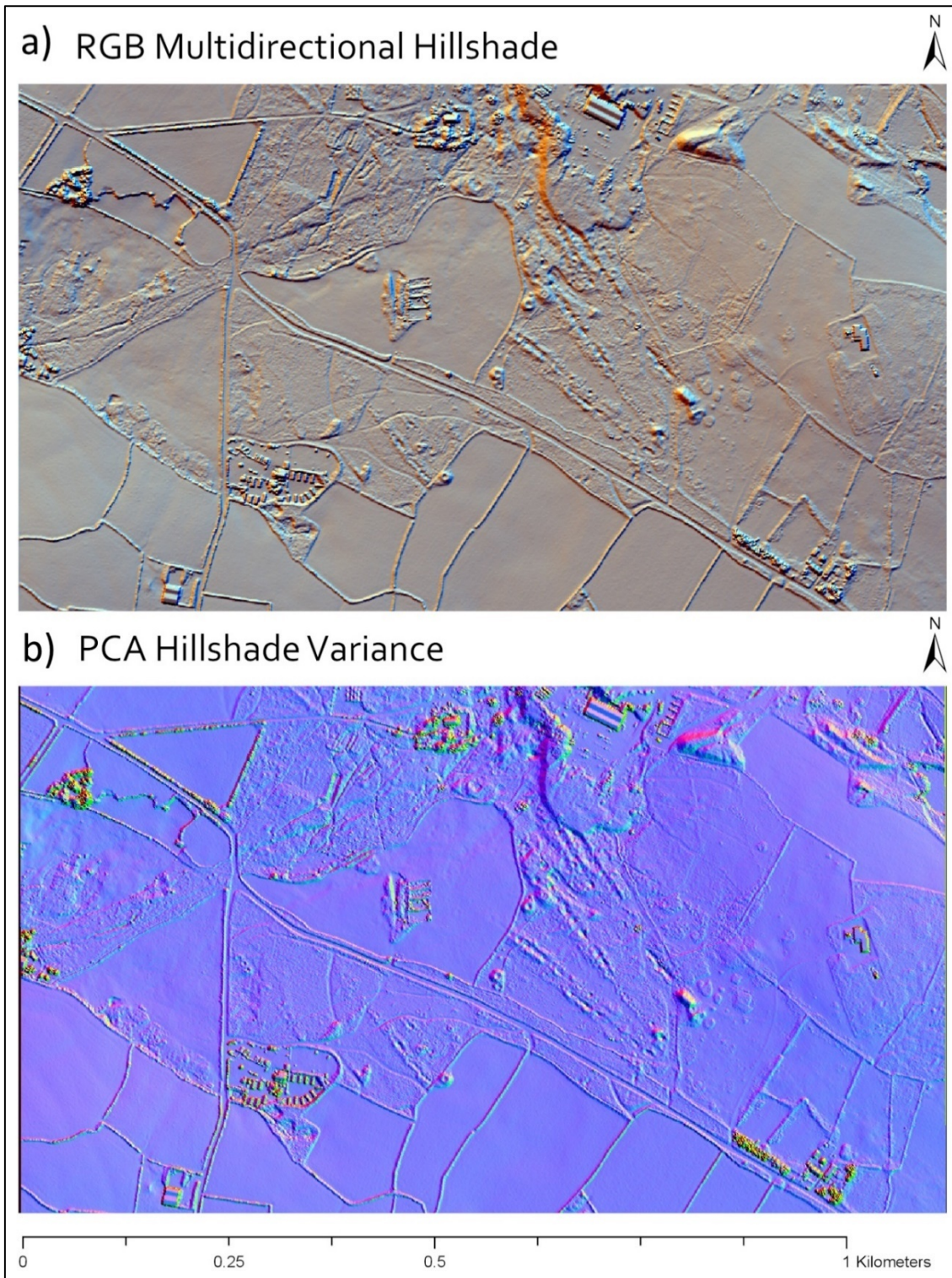
Figure 4-7: a) shows a multidirectional RGB false colour hillshade, where the 315 degree azimuth is mapped to red, the 0 degree azimuth is mapped to green and the 45 degree azimuth is mapped to blue. b) shows the alternative principal component hillshading taken from 16 azimuth directions and mapped to RGB for the three largest principal components.

Moving on from hillshading and its variations, another classic method to conceptualise a surface is from its slope. This is the rate of change between each grid cell and is not linked to which direction it faces. It is a powerful and intuitive way to visualise topographical variations. Another visualisation

technique related to slope is surface accessibility, which is computed by calculating the largest size sphere that can be placed at every pixel without being intersected by neighbouring topography. This technique is particularly effective at highlighting pits and variations along slopes such as terraces and depressions but is weak at detecting small changes along mainly flat terrain (Kokalj and Hesse, 2017). Figure 4-8 displays the slope, where it can be seen that the larger mineshafts are particularly prominent however, smaller shafts are less well defined. Surface accessibility is shown in Figure 4-9, as this landscape is not on a significant slope this representation does not appear to aid interpretation of mining features in this case, however, changes in surface texture related to vegetation cover appear highlighted. Similar in some ways to surface accessibility, Sky View Factor (SVF) is a representation of the portion of the sky visible from each point in the image and is displayed in Figure 4-10. A ridge would have a SVF close to 1 (for a full view) whereas a ditch would have a SVF close to 0 (for a mainly obstructed view).
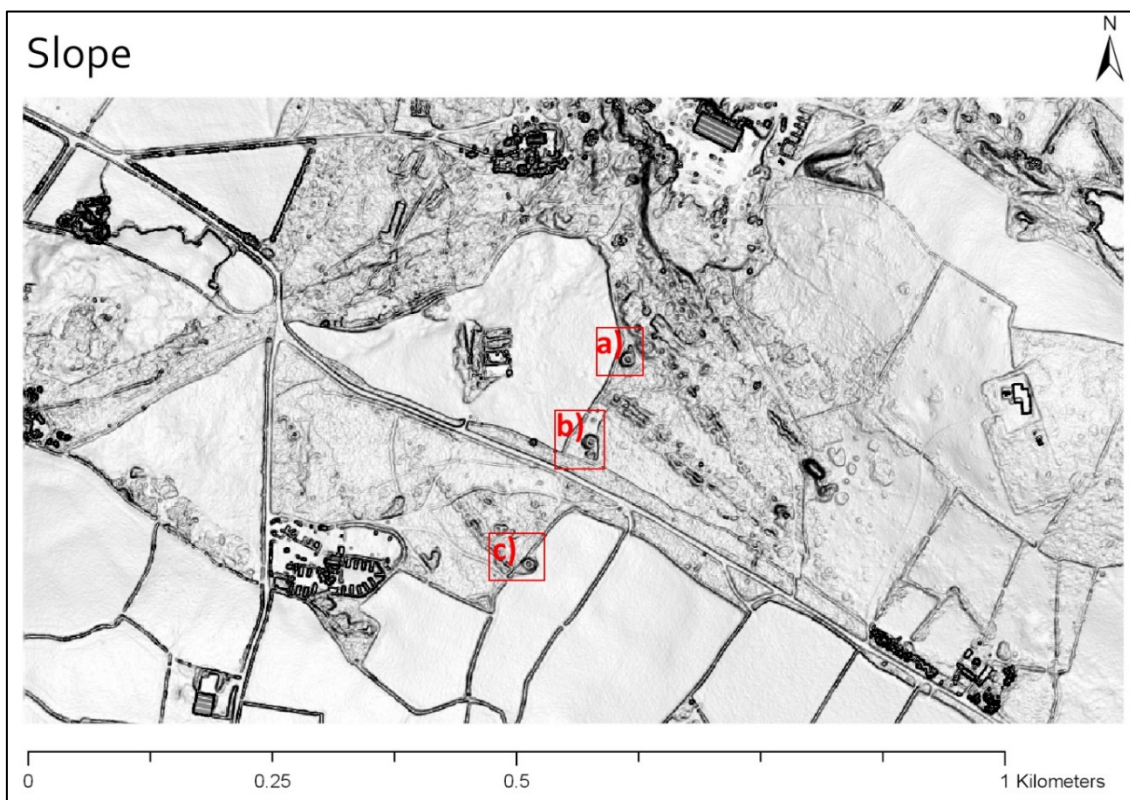
Figure 4-8: Slope representation of the test site. The three largest shafts a), b) and c) are well defined, with the coning around the collars providing a strong signature.
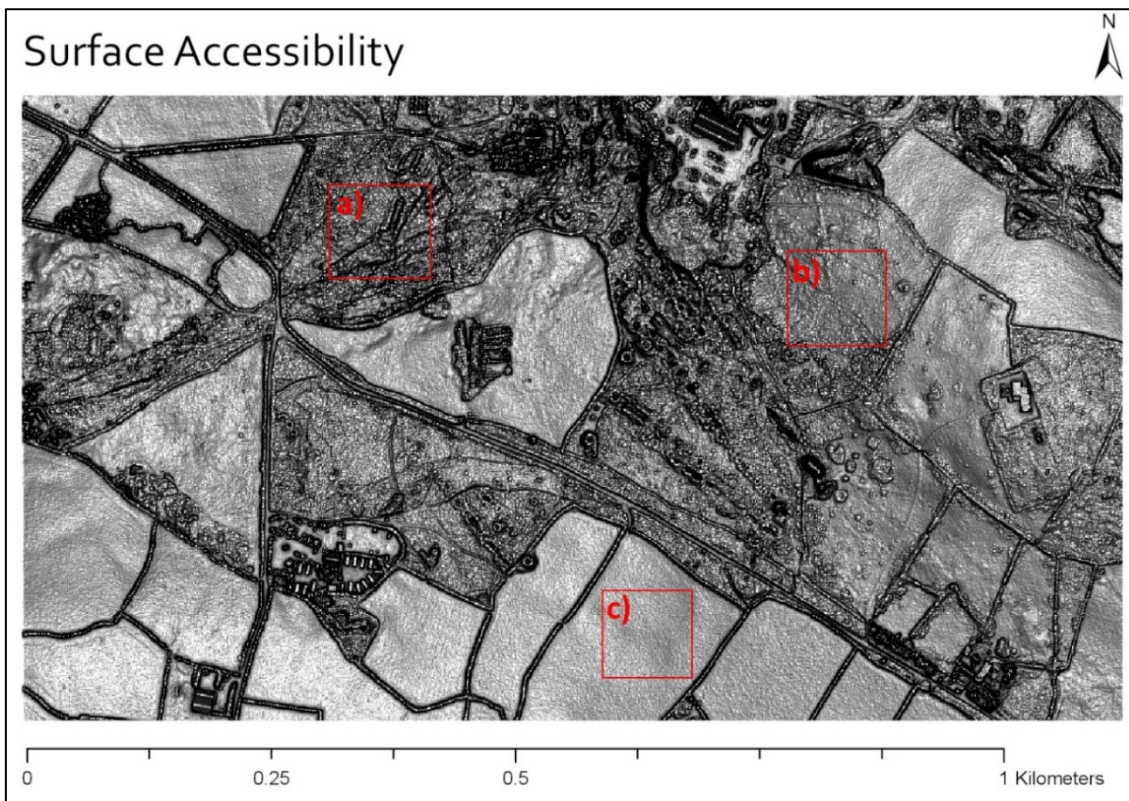
Figure 4-9: Surface accessibility. Vegetation differences are well defined, with a) corresponding to thicker ground vegetation, b) coarse grass and c) cultivated fields.

Another visualisation strategy which examines the view angles from each point is openness. Openness is calculated by measuring the angular size of a sphere either looking up or down from every pixel (Doneus, 2013). It is calculated as either a positive or a negative openness value. As it is calculated in relation to the terrain rather than the sky, features on slopes appear the same as features on horizontal ground. This can make the images harder to interpret manually but may allow it to be a strong input to machine learning algorithms, as features are displayed independent to the angle of terrain (Doneus, 2013). Negative openness is not the inverse of positive openness and highlights deep features instead of protruding features. For the test area little difference was observed in the results from positive and negative openness, with positive openness shown in Figure 4-11. Both the sky view and openness methods are very good at enhancing the appearance of smaller less well defined pits and indicate strong potential for automated mining feature signature detection.
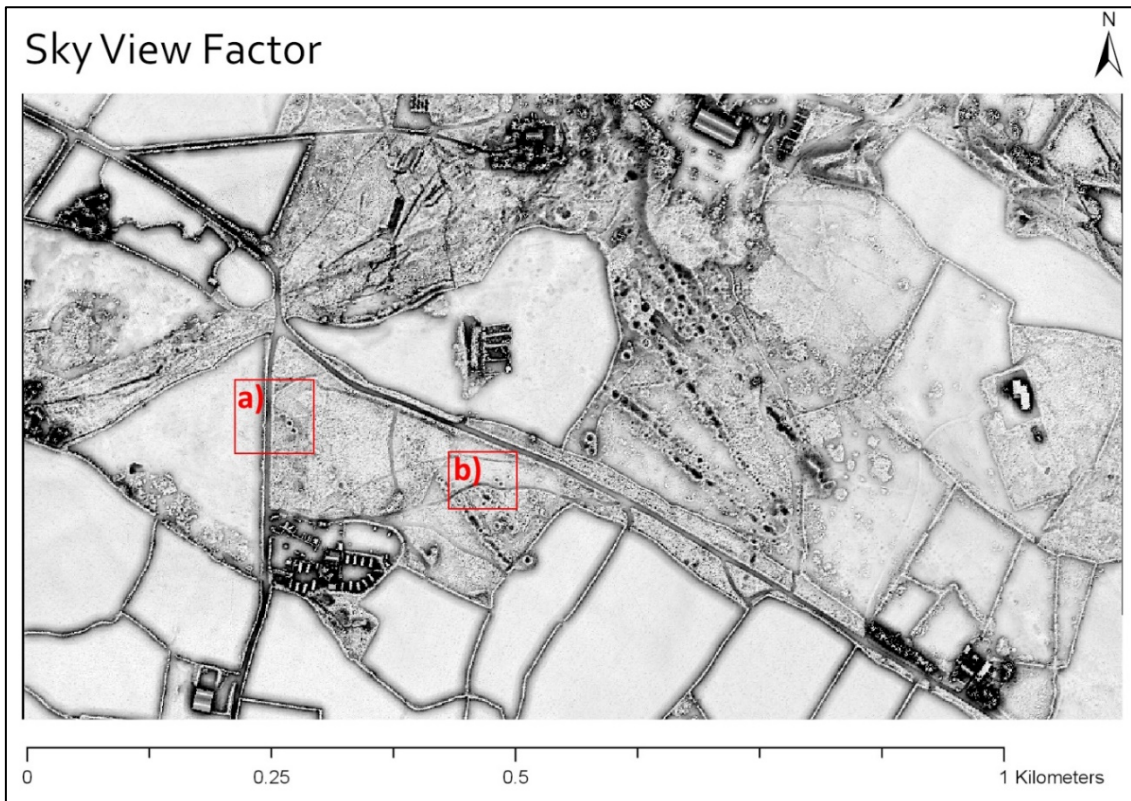
Figure 4-10: Sky view factor. Very small pits shown highlighted by a) and b) are picked up using this visualisation technique.
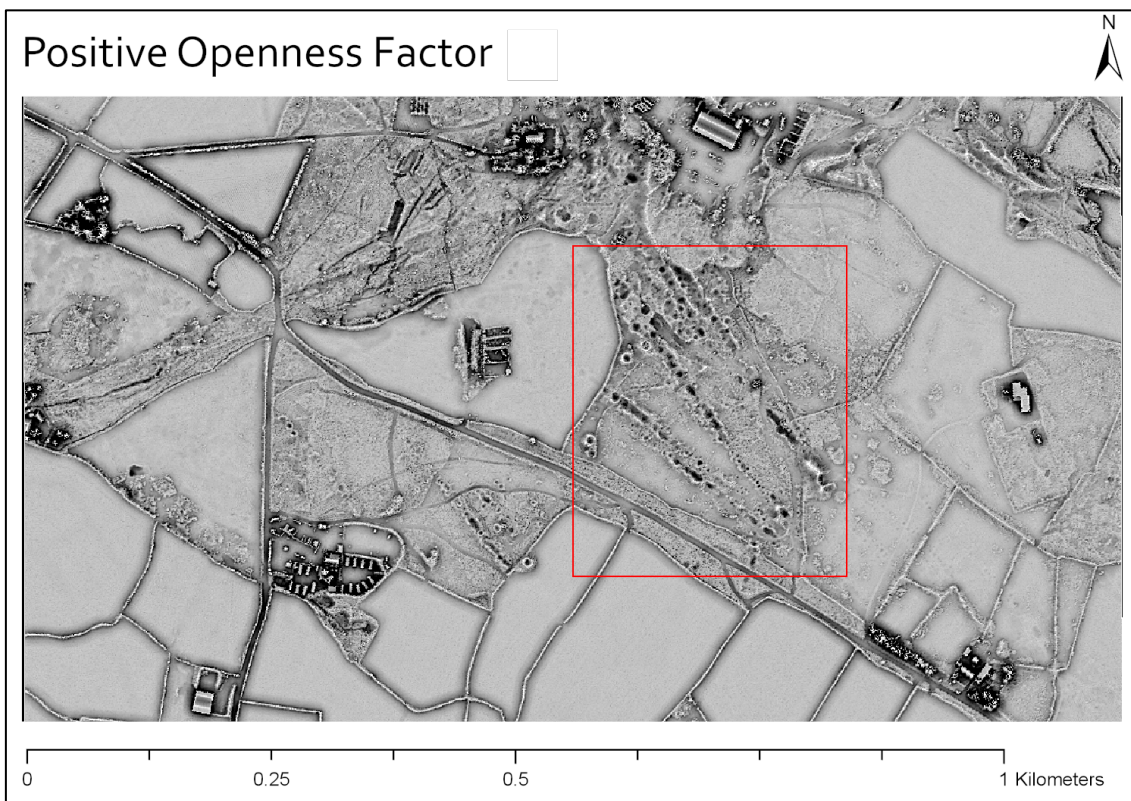
Figure 4-11: Positive openness. The NW-SE orientation of the lodes is readily apparent due to the clear depictions of the connected smaller pits.

Another technique based on the angles of terrain around a point is local dominance, which measures how dominant an observer standing on a point would appear, as seen from surrounding areas (Kokalj and Hesse, 2017). This technique is particularly useful for very subtle relief features, as shown in Figure 4-12, where the faint pits in the centre of the area and even the field topography are visible.
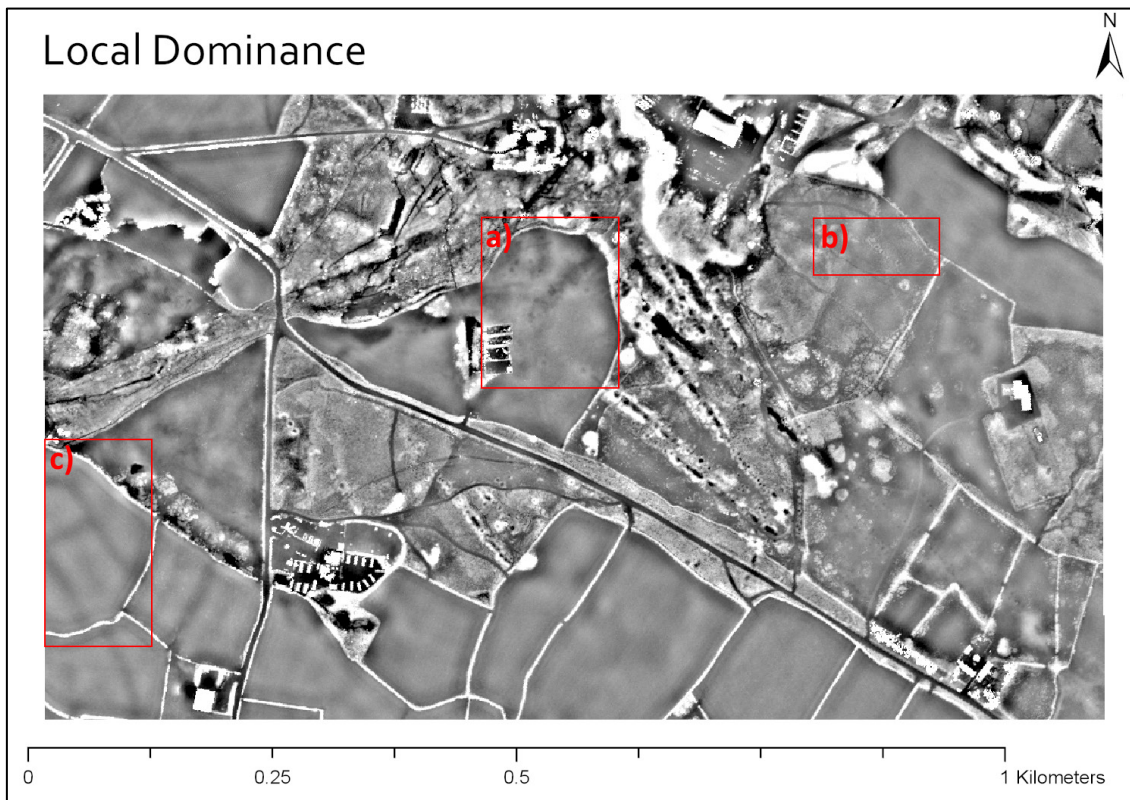
Figure 4-12: Local dominance. Faint mining remains are visible at locations a) and b), whereas c) shows the faint topography within one of the fields.

Two further advanced visualisation strategies borrowed from archaeology are of interest for detecting mining landscape signatures, there are local relief models and multi-scale integral invariants (Kokalj and Hesse, 2017). A Local Relief Model (LRM) is a representation where the major features of the landscape have been removed by a process of detrending. These models are created by first smoothing a DEM so that small features are removed. The smoothed DEM is then compared to the original DEM and areas that are the same in both models are extracted to build the new smoothed DEM. This is finally subtracted from the original to produce the LRM. The LRM enhances small relative changes in the landscape such as those from mining features. Figure 4-13 shows an LRM model which accentuates the elevation changes within the

fields. A weakness of the LRM representation is the difficulty in determining whether a variation is positive or negative, for example a trench or bank will appear similar (Doneus, 2013), in Figure 4-13 the LRM has been overlaid semi-transparently over a standard hillshade to minimise this issue.
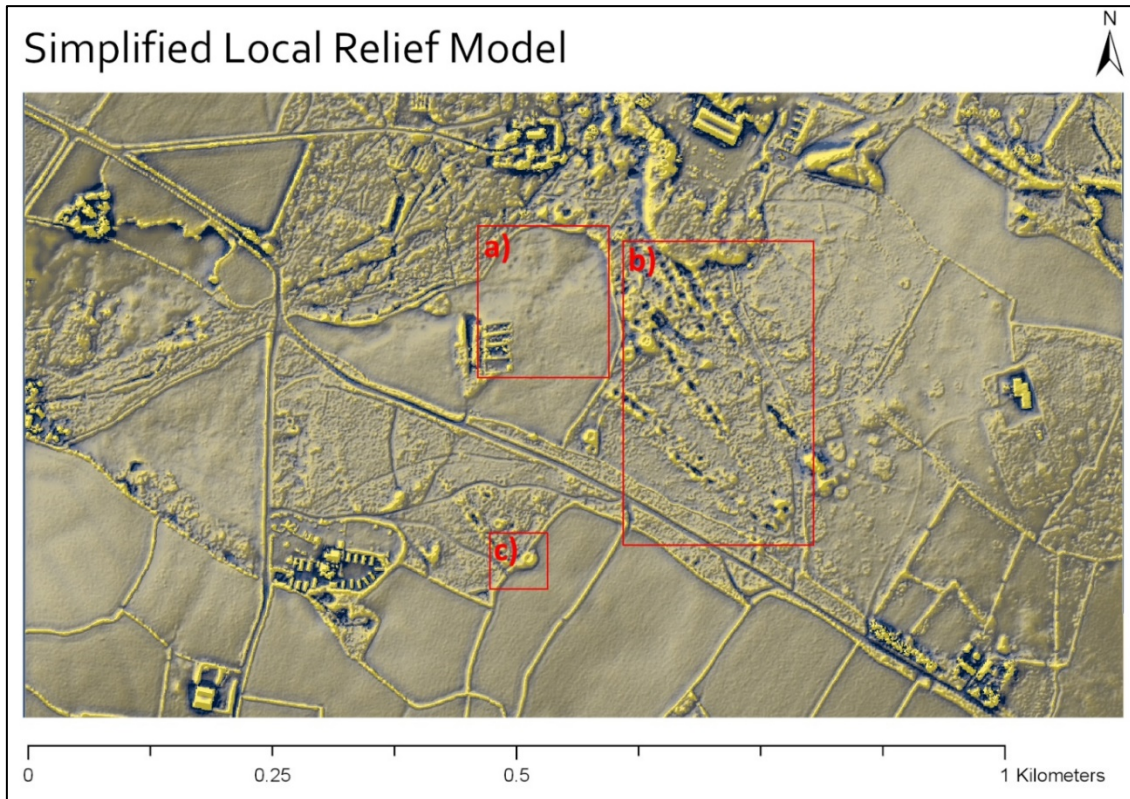
Figure 4-13: Simplified LRM overlaid with a hillshaded model for better depth visualisation. Very faint remains can be seen at a), b) shows the clear multi-pit delineation c) shows the large shaft.

Multi-scale integral invariant visualisation is a technique adapted from methods used to enhance readability of ancient cuneiform tablets (Kokalj and Hesse, 2017). This method allows small variations in surface texture to be accentuated; in Figure 4-14 the faint paths across the landscape can be seen and the definition of the faint pits in the centre of the area is particularly clear.
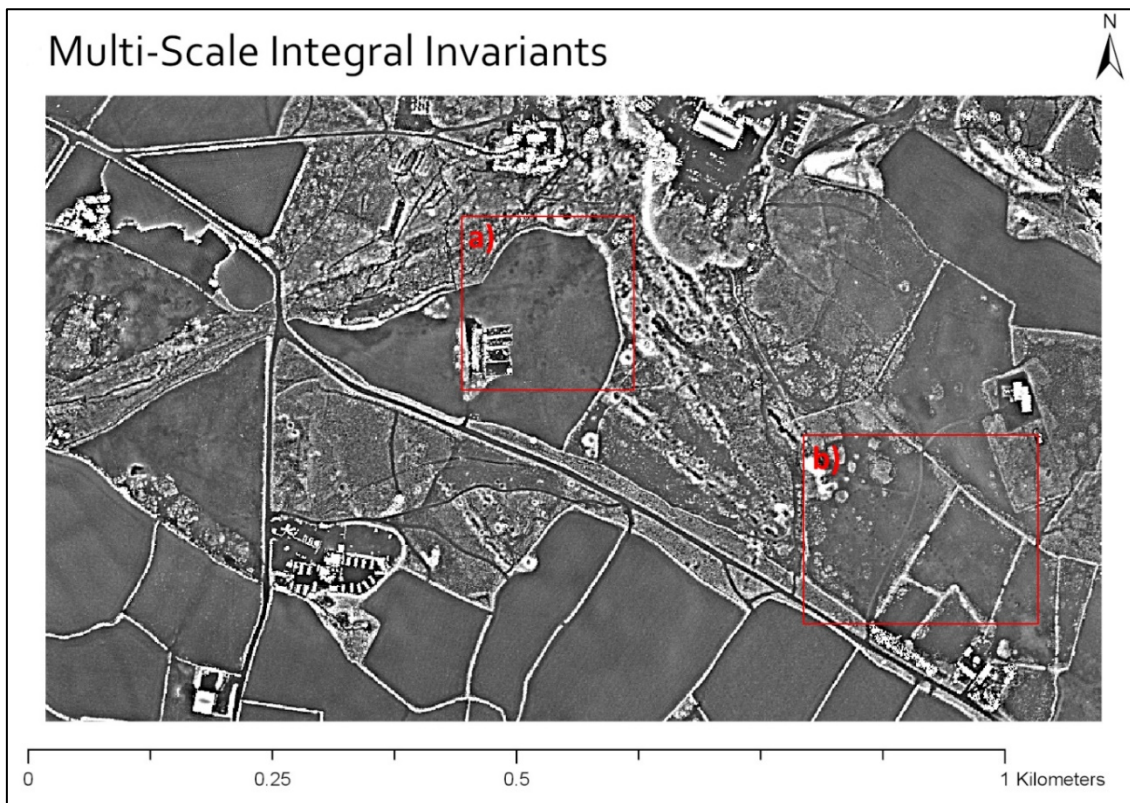
Figure 4-14: Multi-scale integral invariants showing powerful visualisation of indistinct features.

4.3.2 Visualisation summary

The techniques evaluated in this section demonstrate the wide range of ways that a simple grid of measured elevation values can be visualised. The hillshaded models were clearest for human interpretation but lacked detail in the more complex and less defined areas. Of the view angle based methods, the openness technique most clearly delineated the complex interconnected pits and therefore allowed the lode directions to be inferred. Very faint pit remains were most clearly in the multi-scale integral invariant method, whilst the local relief model when combined with a simple hillshade provided a strong overall landscape view. This section has qualitatively evaluated these representations in relation to the ease in which human eyes can perceive landscape signatures; the following sections will address how deep learning techniques can generate their own representations of LiDAR data through convolutional filters, allowing specific landscape signatures to be detected.

4.4 Detection using deep transfer learning – project two

Building on the visualisation algorithms to aid human interpretation, new detection algorithms can be designed which can allow a computer to identify features from LiDAR data in an automated or semi-automated manner. As discussed in Section 4.1 the majority of the published work in this area comes from the field of archaeology as there has been no prior work published on using LiDAR data for mining hazard identification.

Early methods for semi-automated archaeological site identification used template matching (where a predefined template is passed over the scene) or rule-based methods (where rules are applied to determine an object's category). Successful applications of template matching are described by Trier in (2009a) and (2009b). Other proposed methods utilise GEographic Object-Based Image Analysis (GEOBIA), examples of these are described in Sevara et al. (2016) and Freeland et al. (2016). These types of techniques require prior knowledge of the shape and size of the object to be identified and perform well on relatively simple geometries but are less effective at generalizing to unseen or partially occluded examples (Trier et al., 2016). This is because these methods are responding to pre-programmed definitions of the object to be detected rather than 'taught' about the object features.

Machine learning algorithms can provide better results on more complex datasets; recently, very high accuracies have been obtained by combining an advanced visualisation technique based on topographic deviation at multiple scales with a random forest machine learning classifier to identify Neolithic burial mounds (Guyot et al., 2018). In particular, convolutional neural networks (CNNs) have been shown to be capable of solving diverse and complex problems such as visual image question answering (Gao et al., 2015) and real time object detection for over 9000 categories (Redmon and Farhadi, 2017). Considerable research has been carried out in the broader remote sensing community as to how to design and modify similar systems for aerial remote sensing tasks. Primarily this work has involved VHR images as the input to the CNN, either building their own network architecture (Sun et al., 2018) or modifying and fine tuning existing computer vision models (Cheng et al., 2016; Ren et al., 2018). Nogueira et al. (2017) give an overview of the advantages

and disadvantages of these approaches, concluding that fine tuning an existing trained model provides the best results, however, the lack of an appropriate training datasets makes it very difficult to develop a model. Borrowing a similar model and transferring it to the problem at hand is one possible solution (Razavian et al., 2014).

The primary balance that must be addressed when choosing an approach is the applicability of the model versus the availability of training data. If training data and computing power allow, the ideal scenario is to design and train a model from scratch for the required task using the specific data that is required. However, available training datasets for remote sensing data are small and usually not representative of a wide range of environments. Conversely, labelled training datasets in the computer vision community are vast: ImageNet has over 14 million labelled images in 20,000 object categories (Deng et al., 2009) and models trained on these large datasets tend to be less prone to overfitting and can generalise well compared to ones trained on small datasets (Nogueira et al., 2017). However, there are differences in the type of objects they have been trained to detect. For example, in computer vision the objects tend to take up more of the frame and can appear at very different scales, but generally not in many different rotations, whereas for aerial data the scale is relatively constant, but the object can have many rotations (Ren et al., 2018). When using a pretrained model to generalise to images created from a LiDAR DEM the problem is exacerbated, as most existing models have been trained on three channel RGB images and not one channel depth images. This, along with the differing ways that objects appear in a LiDAR DEM versus imagery, can make transfer learning with LiDAR data challenging (Ball et al., 2017).

Two published studies have used CNNs with LiDAR data to identify archaeological objects, with promising results. Trier et al. (2019) found strong positive identifications on one dataset but on their second dataset, which contained more varied objects, their results were less conclusive. Verschoof-van der Vaart and Lambers (2019) employed a similar methodology using variously trained versions of the same pretrained deep learning model to detect multiple classes of archaeological objects, achieving accuracy scores comparable or surpassing those obtained by the other machine learning methods. In both studies a transfer learning technique was used, with the

essential methodology involving the generation of a local relief model (Hesse, 2010) from the LiDAR data and then either converting this generated single channel image into a conventional three channel image stack by triplicating the greyscale channel (Trier et al., 2019) or by modifying the input layer of the CNN (Verschoof-van der Vaart and Lambers, 2019). Both studies used models that had been trained on RGB images of terrestrial scenes such as ImageNet. A recommendation from both studies was to use a model pretrained on data more similar to LiDAR data in the future; however, obtaining such models was determined to be challenging.

Outside of archaeology and even outside of terrestrial remote sensing, large planet scale digital DEM datasets exist from sources such as the Lunar Reconnaissance Orbiter (Zuber et al., 2010) and the Mars Global Surveyor (Albee et al., 2001). Several studies have built and trained CNNs to detect craters from these datasets (Palafox et al., 2017; Silburt et al., 2019; Wang et al., 2018). These models are designed to be highly receptive to elevation changes and to roughly circular patterns observed in single channel DEM images. This makes them a good fit for the problem of mineshaft detection. An example of this type of model was built by Silburt et al. (2019), based on the U-net semantic segmentation model, itself originally designed for medical image segmentation (Ronneberger et al., 2015). This model, named DeepMoon[12] was trained on 30,000 labelled images randomly extracted over the entire surface of the moon combined with the existing catalogues of moon craters. This is a larger and more robust training data set than those available for other LiDAR remote sensing applications, providing a possible solution for the problem of finding applicable transfer learning datasets.

4.4.1 Pre-processing

All three datasets covering the full range of scales were examined in the detection study. The 0.5m Dartmoor dataset was used for training the model which was subsequently tested on the 0.5m Dartmoor dataset, the 1m St Just dataset and the 0.25m Yorkshire dataset. The St Just and Yorkshire datasets

---

[12] Available at https://github.com/silburt/DeepMoon.git

were not used for training as they do not contain enough representative samples to form a large training dataset.

In all cases, the objects to be detected are trial pits, shallow pit workings and shaft heads. Examples of these are shown in Figure 2. Trial pits are dug whilst prospecting for tin lodes. They are usually 2-3m in diameter, of limited depth (up to 1m) and are often silted, water filled and reedy (Newman, 2010). Shallow pit workings are comprised of alignments of deeper pits which are dug to below the soil overburden and mined downwards from there; however, these are not underground mines and there is no lateral development between the pits. The depth of these types of workings would be limited by the ability of the surrounding side-walls to remain intact before collapsing, which is usually less than 3m. These workings present as conical depressions often accompanied by a ring of spoil material, crescentic on the downhill side in sloping ground (Newman, 2010). The final category are shafts for true underground mines. These have mainly been capped or backfilled in Dartmoor for public safety; however, evidence may remain in the form of large conical pits or straight openings. Site inspections may reveal a collar of finished material lining the inside of the shaft, but this is generally not visible from aerial surveys.
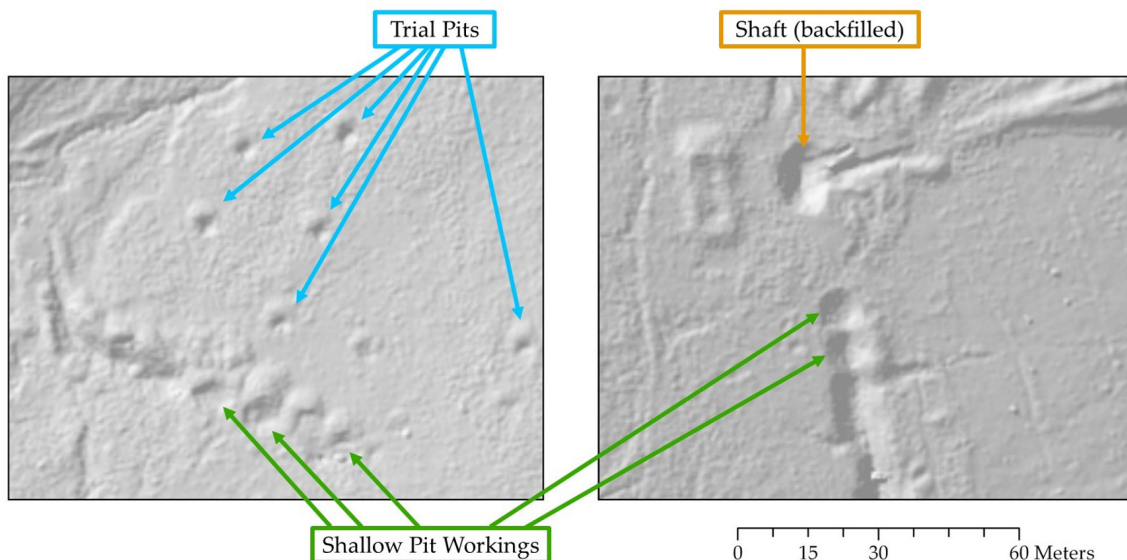
Figure 4-15. Examples of the historic mining objects found in this study displayed on a 315° azimuth 35° sun elevation hillshaded visualization created in ArcGIS from the 05m resolution Dartmoor dataset. Base DSM © Environment Agency 2015

The raw DSM data for each area was imported into ArcGIS Pro (ESRI, 2019) along with several other interpretive layers such as historical maps and aerial images to create a GIS of the study area. Other GIS software could be used for this step, but ArcGIS Pro was chosen as it has a function for automatic exporting of image tiles and training labels, crucial for the later steps of the workflow. The additional GIS layers were only used to add context to the dataset to aid the human operator. To generate training and validation datasets, a desktop survey was carried out to identify features resembling mining pits. The test area datasets were created in the same way, but in order to validate the performance of the model every feature in the test set was later confirmed with a ground survey. This survey involved visiting the test sites with two reference maps, one containing the predictions and one containing the human generated pit locations from the desktop survey. Using these maps in conjunction with a handheld GNSS for site orientation the true existence of pits shown on the maps was confirmed or rejected. The pits were not recorded with the GNSS as in many cases it is not safe to access the ground directly above suspected shafts. A schematic of the methodology is shown in Figure 4-16, precise processing steps for reproducibility are given in **Appendix B**.
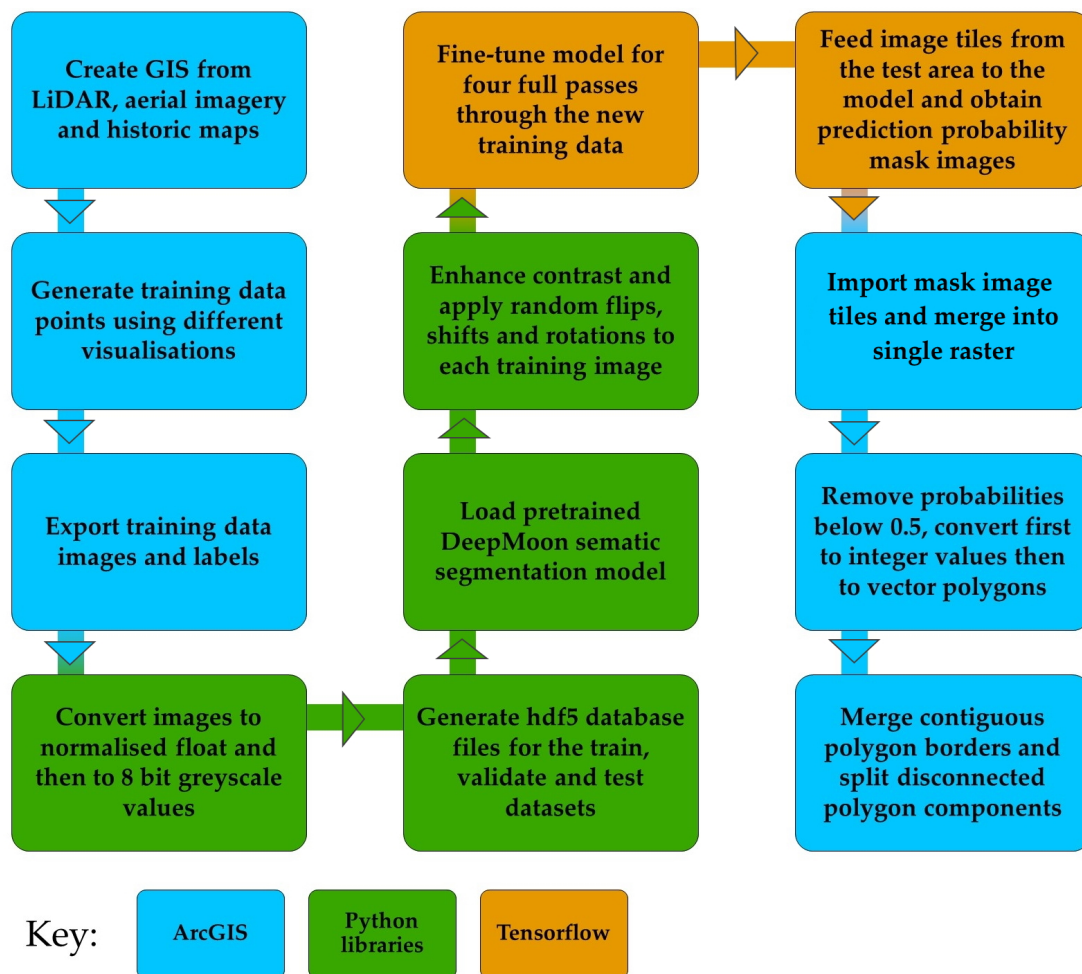
Figure 4-16. Methodology process diagram.

For the model inputs, image tiles of 256x256 pixels were exported along with the pit locations as .xml labels to create image segmentation masks. The overlap between tiles was set to 52% to ensure no pits were split by tile seams. To preserve the fine detail in the DSM image, the image tiles first were exported as 16-bit float images with the values corresponding to the actual ground elevation of the data within that tile. Each tile was then individually rescaled to greyscale values between 0-1 maintaining its original distribution before finally being converted to an 8-bit integer format. To enhance contrast the image tiles were further rescaled linearly prior to model input. This rescaling strategy was appropriate to this dataset as no tiles contained large elevation changes such as cliffs. The image tile preparation process is shown in Figure 4-17. For the training and validation datasets, only image tiles which contain mining pits were exported. These datasets are stored in hdf5 format with the image names used as the database key. Table 4-1 shows the dataset splits, number of pits and pit instances per dataset, along with the minimum, mean and maximum pits per

127

image tile. The pit instances are greater than the number of pits as some pits are present on more than one image tile due to the >50% overlap between tiles.

Table 4-1: Dartmoor and Yorkshire dataset statistics

Dataset	Image tiles	Pit ground	Pit instances	Minimum	Mean	Maximum
Train	542	1568	3649	1	5.96	59
Cross-validate	71	254	423	1	5.96	33
Test Dartmoor	196	193	654	1	5.74	24
Test Yorkshire	900	172[1]	n/a	n/a	n/a	n/a

[1] Only pits within a section of the dataset were ground truthed as shown in Figure 4-24.
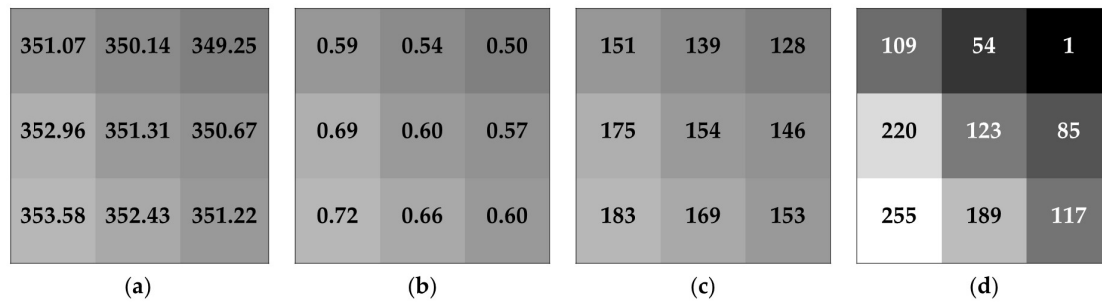
Figure 4-17. Overview of image preprocessing pipeline. (a) shows a selection of original individual pixel values, (b) shows the same pixels rescales between 0 and 1. (c) shows the conversion to greyscale.(d) shows the pixel values after linearly rescaling by tile range.

As discussed in Section 4.3, other visualisations of LiDAR data have been shown to aid in identification of archaeological features by humans. To test whether this holds true for machines, a SLRM model and both positive and negative openness models were generated from the original exported tiles using the Relief Visualization Toolbox (Kokalj and Hesse, 2017). As openness is calculated in relation to terrain rather than the sky, features on slopes appear the same as features on horizontal ground (Doneus, 2013). This is a valuable property for the Dartmoor data as most of the features are situated in rolling moorland terrain. Figure 4-18 illustrates the different visualisation types generated for this study.
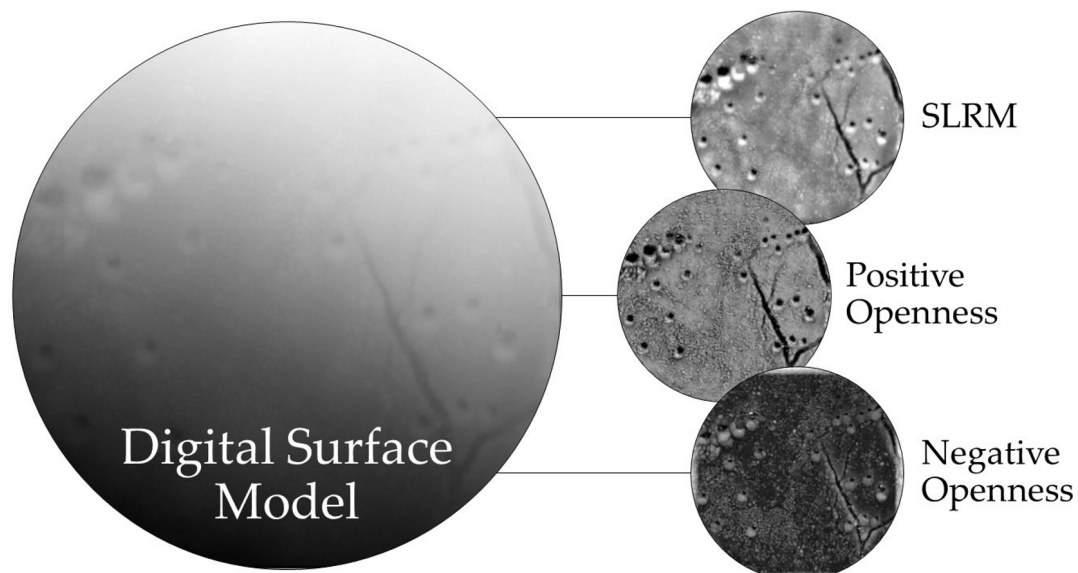
Figure 4-18. Illustration of the different advanced visualisations created from the original LiDAR DSM. Base DSM © Environment Agency 2015, visualisations created using the Relief Visualisation Toolbox

.

4.4.2 Deep learning model

The type of model used in this research is a variant of an Artificial Neural Network (ANN) known as a Convolutional Neural Network (CNN), described in Chapter 2. Initially, an object detection pipeline using the Inception model (Szegedy et al., 2016) pretrained on the Common Objects in Context dataset (Lin et al., 2014) was trialled. The preliminary results from this method showed some correct detections but there appeared to be many mining pits not detected by the model even after 100,000 training epochs. Images from this initial method are shown in Figure 4-19. It is suspected that the mining pits detection task is simply too different from the original task to achieve optimum results. These initial tests showed a detection rate of less than 40%, this result, along with the recommendations from Trier et al. (2019) motivated a search for a transfer learning candidate model that resembles more closely the task at hand instead of continuing to refine the Inception model.

After exploring alternative options such as the Kitti dataset (Geiger et al., 2013), the exact model chosen for this research is a version of the U-net model designed by Ronneberger et al. (2015) and modified by Silburt et al. (2019). The U-net model is an encoder-decoder (see Badrinarayanan et al. (2017)) model with a near symmetrical architecture, designed for biomedical image segmentation. It has no final fully connected layer, replacing it with a 1x1

convolutional layer with a sigmoidal activation function to output pixelwise class probabilities, thus reducing the number of hyperparameters to tune and making it more suitable for small numbers of training data. The original U-net achieved significant accuracy improvements over the next best architecture in the International Symposium on Biomedical Imaging's cell tracking challenge despite the training set only containing 35 images (Ronneberger et al., 2015). Biomedical image analysis shares many challenges with remote sensing LiDAR analysis such as small training sample sizes, single channel images and high resolution data. Therefore, it is more applicable to use a model such as U-net rather than one of the models designed for large datasets of natural images, as shown in Figure 4-19. The U-Net architecture is described in more detail in Chapter 5, Section 5.2.4.

(**a**) COCO image   (**b**) Dartmoor hillshade   (**c**) Dartmoor SLRM

(**d**) Cell image   (**e**) Moon DSM   (**f**) Dartmoor DSM
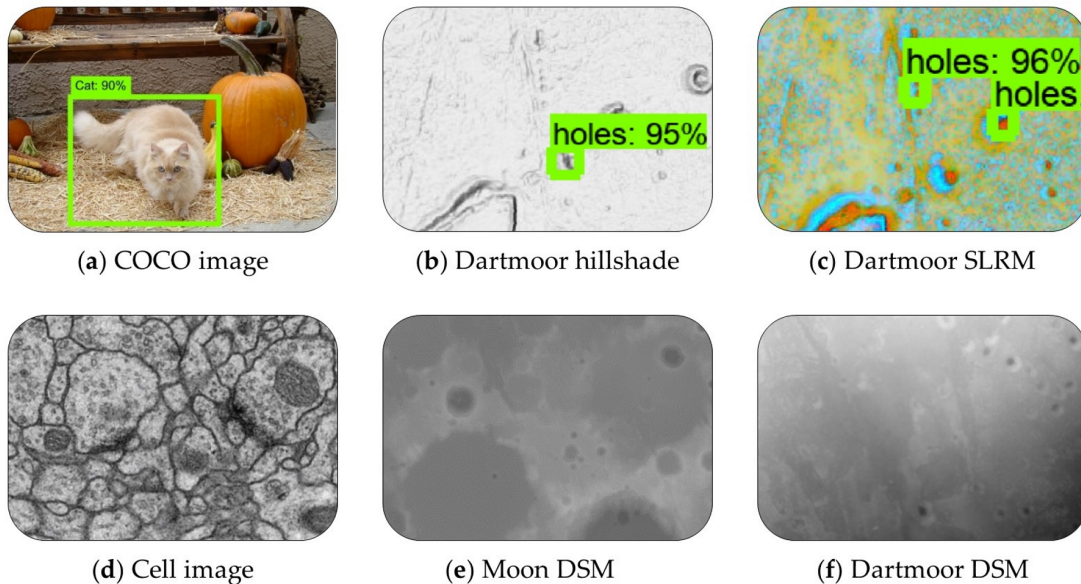
Figure 4-19. Examples of the input data to different pretrained models. (a) is an example from the Common Objects in Context (COCO) (Lin et al., 2014) *(b) and (c) show the results from an object detector pre-trained using the COCO dataset. (d) shows the type of microscopy data which the U-net architecture was designed to segment* (Ronneberger et al., 2015) *and (e) shows data from the lunar DSM which was used to pre-train the model used in this research* (Silburt et al., 2019). *(f) shows the DSM data used in this project. Base DSM in (b), (c) and (f) © Environment Agency 2015.*

4.4.3 Transfer learning

Nogueira et al. (2017) found that for remote sensing problems with limited training data, a transfer learning strategy achieved the most accurate results across all tested datasets. In transfer learning, instead of initialising the model weights from scratch, the weights from another model trained for many epochs on a larger dataset are used. One transfer learning strategy involves removing the last layer of the network and replacing it with a layer to classify the objects

of interest, this is required if the final classification categories are different. Another approach is to fine tune a model by adding new training examples whilst keeping the final output layer the same. All the model weights can be updated, or the lower layers can be frozen and only the weights in the upper layers are updated. For this research, as the classification is the same geometrically if 'crater' is substituted for 'pit' a fine-tuning strategy was employed with all weights unfrozen. As this study utilises a pre-existing model, the same software Python (Python 2019), TensorFlow (Abadi et al., 2016) and Keras (Chollet, 2015) used by the creators of the original DeepMoon model (Silburt et al., 2019) are used throughout. All of these packages are industry standard and available free from their respective websites.

4.4.4 Model training

In a neural network the hyperparameters can be used to control overfitting; for the DeepMoon model, the hyperparameters include weight regularisations for the convolutional layers, dropout layers, filter size, model depth, and learning rate. Full details on these hyperparameters and complete model design can be found in (Silburt et al., 2019). These hyperparameters were chosen after a cross validation check using 60 models, where the hyperparameters were chosen randomly from across their standard ranges. To avoid overfitting on the small project dataset used in this research the hyperparameters chosen in Silburt et al. (2019) have been maintained here, with only minimal fine tuning training. Silbert's base model was trained for 4 epochs (where one epoch equals a full pass through the entire training set). As the lunar dataset contained 30,000 images this training totalled 120,000 training examples. A standard learning rate of 10^{-4} was found to deliver the best results (Silburt et al., 2019). The additional training for transferring the model to its terrestrial context involved 4 more epochs of 520 images, totalling 2,080 new training examples. The number of fine-tuning epochs was varied to determine the most effective fine-tuning strategy, discussed further in Section 4.4.6.1. To further control overfitting, data augmentation is carried out between epochs. In this process all input images are randomly flipped, rotated and shifted prior to model input.

4.4.5 Post-processing

Once the model is trained and verified against the cross validation dataset, individual image tiles to be tested are inputted to the model and probability masks are outputted as .tif files. Using the same naming convention for both input and output files results in correct translation into the original coordinate system. Using ArcGIS, all output probability masks are then mosaiced into one continuous raster covering the entire test area.

For qualitative visual analysis and map creation, a graduated stretch symbology where solid colour depicts probabilities of 1 and fully transparent depicts probabilities of 0 is used for maximum readability. This visualisation scheme maintains information on the confidence of the prediction and allows for the more subtle workings of the model to remain visible. This enhances the model's readability in comparison to a yes/no response as it symbolises uncertainty in the model, allowing an archaeological prospector more freedom to interpret the results using human reasoning. To quantitatively determine the rate of true positives, false negatives and false positives in order to report accuracy metrics, a new binary mask layer was created containing only pixels with prediction probabilities above 0.5. These pixels were then vectorised, merged and filled to create a vector layer of predicted pits to use in spatial queries. A comparison of these post processing methods is shown in Figure 4-20. It can be seen in Figure 4-20 (c) that there are some incomplete rings, this is because some detections are made up of a mixture of pixels above and below 0.4 probability. This further supports the decision to use the full masks rather than the instances for interpretation where possible.

(a)   (b)   (c)
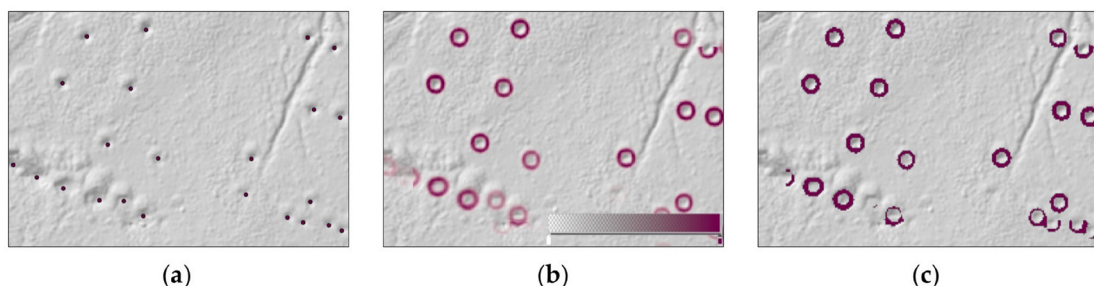
Figure 4-20. Comparison of qualitative and quantitative results representations. (a) shows the ground truth locations of a section of very shallow (30-50cm depth) mining pits in the Hexworthy test area. (b) shows the model's predicted results depicted with a graduated transparency colour scale representing model confidence and (c) shows a binary mask where all prediction pixels above 0.4 are assigned as 'pit' and all others are discarded. DSM © Environment Agency 2015.

4.4.6 Results

During training binary cross-entropy was used as the loss metric as it is the standard loss metric used for problems of this type. For the Dartmoor dataset the training loss began at approximately 0.02 for the DSM and between 0.03-0.04 for the other visualisation types, reducing to an average of 0.0146 for all data types after four epochs. There was negligible variation in the loss by visualisation type. The cross validation loss remained within 0.005 of the training loss for each epoch with the average cross validation loss 0.0145 after four epochs. However, during human examination of the output masks it was observed that because the model is attempting to lower the global loss over every pixel, the numeric values output from the TensorFlow console did not fully describe the real effectiveness of the model for detecting pit objects. This is suspected to be due to the fact that the model loss is a pixel based loss function rather than an object based one. Figure 4-21a displays the losses per epoch; showing that whilst the cross validation loss continues to decrease after four epochs, when compared to the F1 score shown in Figure 4-21b it can be seen that the real detection accuracy degrades after four epochs.

4.4.6.1 Cross validation results

In light of this, a much smaller human cross validation was carried out on five sample tiles from the cross validation dataset. These tiles were chosen after inspecting all tiles in the validation dataset as good representative examples to assess each model's performance at both ends of the difficulty spectrum, from simple cases with several well defined pits to complex cases with multiple ill-defined and overlapping pits or pits within larger trenches. To determine the optimal fine-tuning strategy, the number of epochs for which the model was retrained was varied and the results were examined by counting the detection instances over these tiles.

For each model and each tile, the number of true positives (correctly detected pits), false negatives (undetected pits) and false positives (detections which do not correspond to true pits) were counted. From these numbers the precision (the proportion of the model's pit predictions that were correct) and the recall (the proportion of actual pits that were detected) were calculated. The F1 score

133

(harmonic mean of precision and recall) was also calculated, as it is a useful single valued accuracy metric for a detection problem of this kind (formulas defined in Table 4-2). Due to the variability of deep learning model convergence, training will not produce identical results every time, to account for this each test was run three times and averaged. Figure 4-21b shows how the precision, recall and F1 scores vary as the number of fine-tuning epochs is increased. It should be noted that this figure shows accuracy metrics over only 5 tiles from the validation dataset, chosen for their difficulty to evaluate model generalisation ability. Therefore, it does not represent the general performance obtained by the model on the test datasets (Table 4-2). It can be seen that the best results are found after three to four epochs of training. The degradation of accuracy after four epochs could correspond to overfitting; because each epoch trains the model using the same 520 test images, albeit augmented differently each time. As another test, the DeepMoon model was also run directly on the Dartmoor data without any fine-tuning training, this gave detection rates of approximately 40% with a bias towards large pits more similar in appearance to impact craters.

Once the optimal amount of fine tuning was determined, the four advanced visualisation types were tested against the same five sample images. Each of the visualisation types depicted previously in Figure 4-18 were used as the training data input for fine tuning the model. Using the knowledge from the previous validation test, the models were trained for four epochs; as before, each test was run three times and averaged. Longer training runs of eight epochs were also tested. This is to account for the possibility that due to the greater difference between some of the visualisation styles and the model's original lunar DSM training data more epochs might be required to obtain strong results. However, these tests displayed the same behaviour as that shown in Figure 4-21b. It can be seen from Figure 4-21c that whilst the precision is high for all four data representations, the recall and therefore the F1 score is poorer for the advanced visualisations.

(a) Training and validation loss by epoch (DSM dataset)

(b) Accuracy metrics by training epochs (DSM dataset)

(c) Accuracy metrics by visualisation type
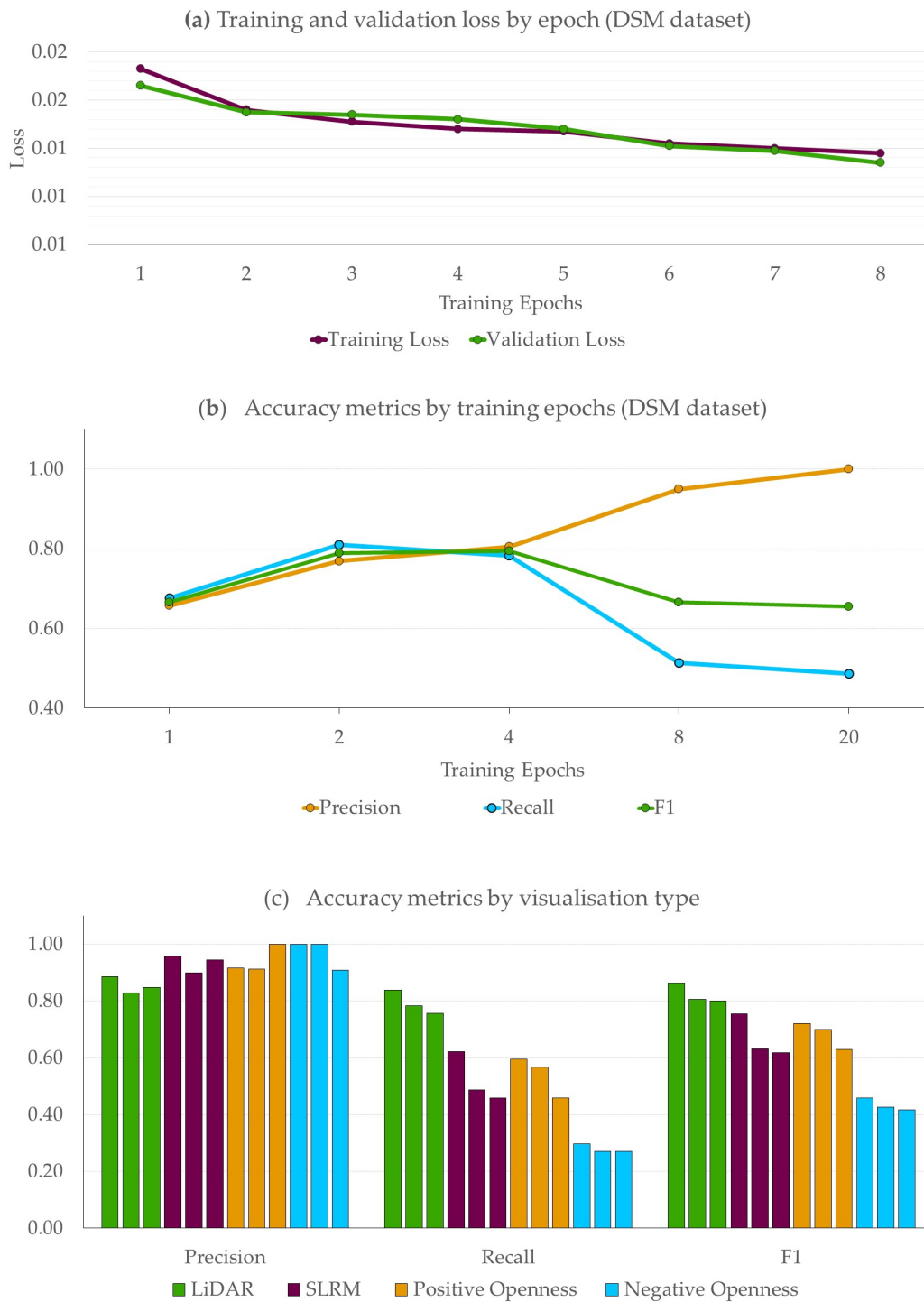
Figure 4-21. Accuracy metrics by training epochs (a) and by visualisation type (b), both evaluated on sample tiles from the validation dataset. Note: this figure shows accuracy metrics over only 5 tiles from the validation dataset chosen for their difficulty to evaluate model generalisation ability. Therefore, it does not represent the accuracy obtained by the model on the test datasets.

4.4.6.2 Test area results

The cross-validation results informed the development of the final model, which was then evaluated on the final unseen test datasets. The model was primarily evaluated on a 1km^2 tile of LiDAR data in Dartmoor approximately 20km away

135

from the training area. Two additional tests were carried out on the Yorkshire dataset more than 500km away from the training data and on the original St Just area used in the visualisation section. The quantitative results obtained from the Dartmoor and Yorkshire datasets are summarised in Table 2. For all results the highest performing model from the Dartmoor validation dataset was used for the predictions. It must be noted that these results have been calculated from the binary results mask. Of the missed detections 23 out of the 38 in Dartmoor and 17 out of 30 in Yorkshire are still visibly predicted in the full transparency results layer. This is because they fall below the 0.5 probability threshold used in the binary masking operation, thereby removing them from the count. Quantitative results were not measured for the St Just data as there were too many confusion objects and pits that were uncertain to the human digitiser. The model predictions for the St Just data are shown in Figure 4-22, for performance reference this figure can be compared with the visualisation results in Figure 4-6 to Figure 4-14.
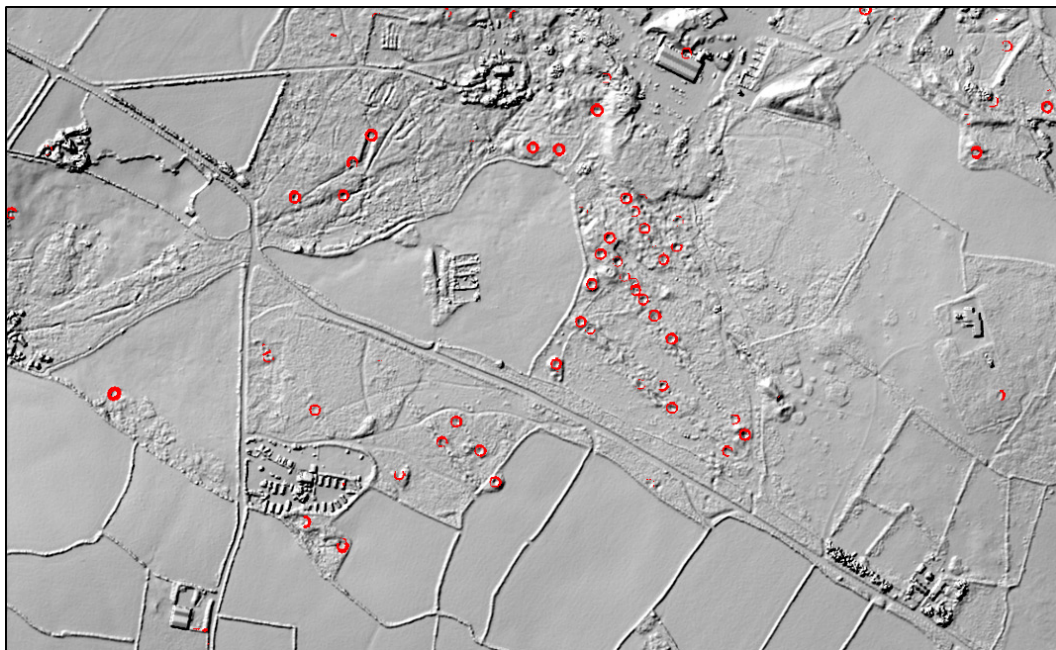
Figure 4-22: St Just area deep learning algorithm qualitative results.

Table 4-2: Full results from Dartmoor and Yorkshire test datasets

Test Area	True Positives	False Positives	False Negatives	Precision[1]	Recall[2]	F1[3]
Dartmoor	155	37	38	0.81	0.80	0.81
Yorkshire	142	13	30	0.92	0.83	0.87

[1] Precision = True Positives / (True Positives + False Positives)
[2] Recall = True Positives / (True Positives + False Negatives)
[3] F1 = 2 * ((Precision * Recall) / (Precision + Recall))

4.4.7 Discussion

The cross-validation results from the different types of LiDAR visualisations indicated that the model performed better when trained on the raw 8-bit DSM height values rather than any of the advanced visualisations. It is suspected that whilst these visualisations are effective for human interpretation of LiDAR data (Kokalj and Somrak, 2019) and also effective for more traditional machine learning techniques (Guyot et al., 2018), because deep CNNs learn their own feature representations during training, it is not desirable to artificially alter the data representation prior to input. However, it also must be taken into account that the CNN chosen in this study was pretrained on 8-bit DSM height values, thereby introducing a bias towards this representation. To fully test which LiDAR visualisation is best suited for CNNs in future, would require a robust CNN trained from scratch on multiple differently visualised representations of the same data; however, such a model has not been made publicly available from any known sources at this time. To attempt to test this theory with the existing datasets experiments were carried out to create a model from scratch using the DeepMoon architecture and the Dartmoor training data with different visualisations. However, no meaningful results were obtained from any visualisation, presumably due to the limited size of the training dataset.

The SLRM and openness visualisations are included in this study as discussion points, to observe how the predictions vary and to provide stimulation for future work including that presented in the final project of this chapter. An example of the predictions on a single challenging tile for each visualisation type is shown in Figure 4-23. It can be seen that the predictions from the raw DSM are the most sensitive, resulting in the least amount of missed detections, and is the only visualisation type that picks up the isolated pit in the lower right corner. The confusion areas of low probability are easily filtered out by setting a probability threshold of 0.5 in the post-processing steps, as discussed in Section 2.6.
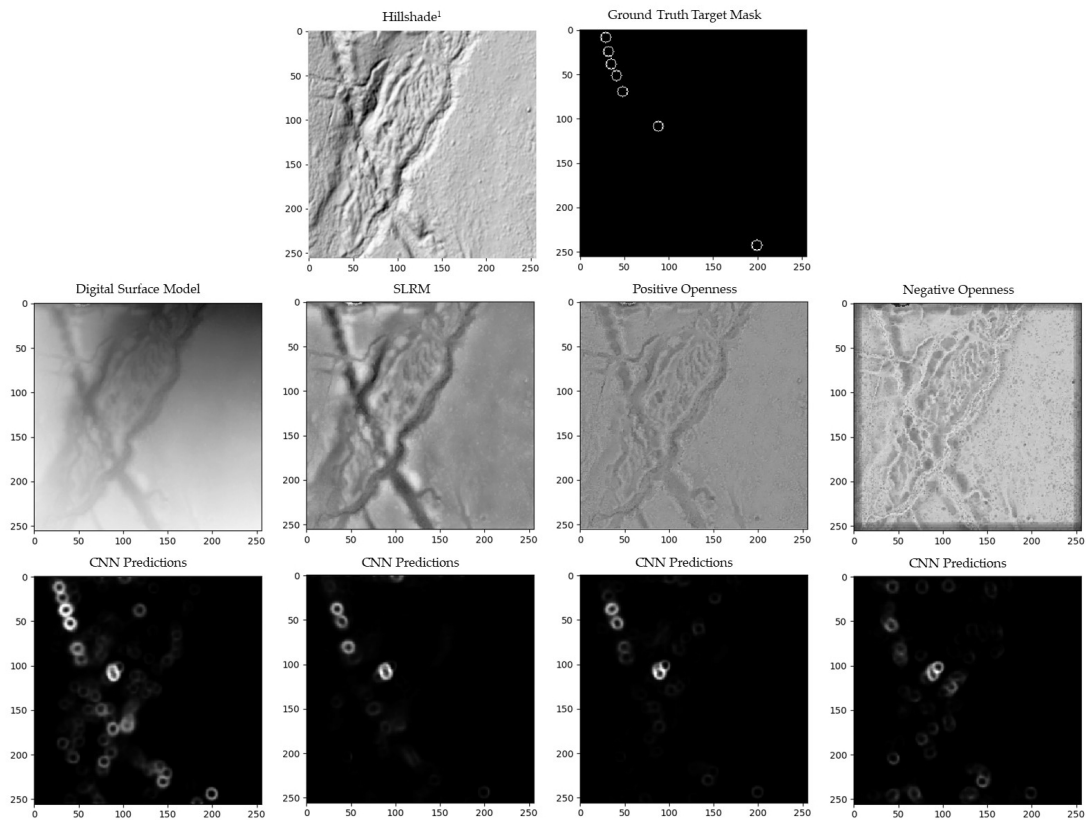
Figure 4-23. Results from a single image tile for each of the different visualisation predictions. [1]Hillshade used for display purposes only and not processed by the CNN model. Coordinate system arbitrary pixel based.

The final test area results demonstrate that the model is highly effective with the correct detections greatly outnumbering the missed and false detections, displaying strong precision and recall simultaneously. Figure 4-24 shows the full transparency results overlaid on the Dartmoor and Yorkshire test datasets. This figure shows that the model is highly capable of discerning mining pits and is not overwhelmed by false positives. It also demonstrates that even if individual detections might not always be correct the greater trends in the landscape are very clearly reproduced by the model. From a management perspective, these automatically generated maps clearly delineate the extents and key structures of these historic mining sites, with limited confusion areas due to model assumptions and landscape morphology. Recorded precision scores in LiDAR based deep learning applications range from 0.12 (Trier et al., 2019), through 0.62 (Trier et al., 2016) to 0.90 (Verschoof-van der Vaart and Lambers, 2019). While these scores are not directly comparable as each study has used a different dataset and detection object, the precision scores of 0.81 and 0.90 obtained in this research can be considered state of the art at this point in time.
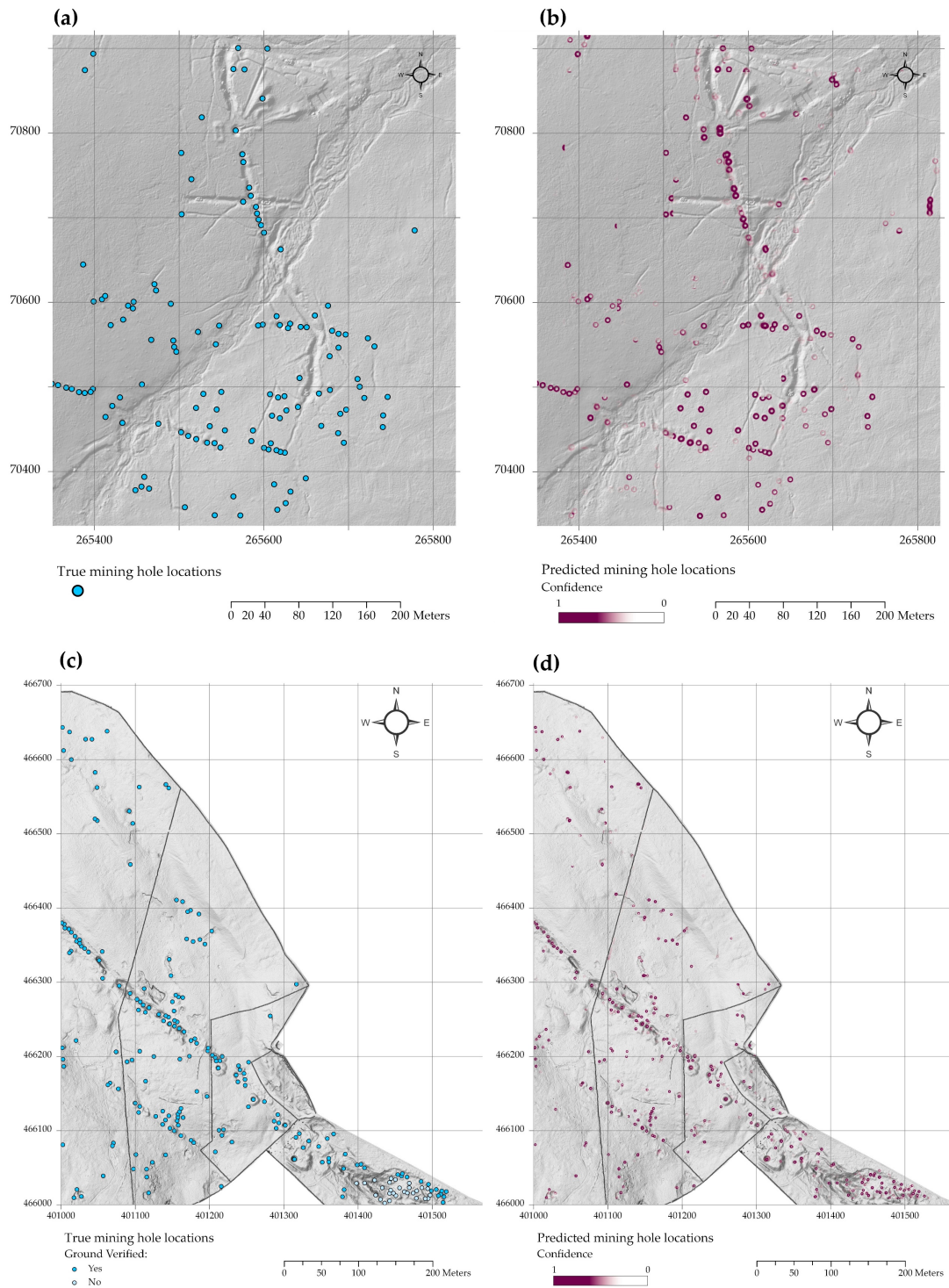
Figure 4-24. Results overlaid on hillshaded LiDAR. (a) and (b) are from the Dartmoor Hexworthy mine test area, Ordnance Survey grid tile SX6570. (a) shows the true mining hole locations in blue and (b) shows the model's predicted mining hole locations in magenta. (c) and (d) show the results from the Yorkshire Yarnbury mine test area, Ordnance Survey grid tile SE0166. (c) shows the true mining hole locations in blue and (d) shows the model's predicted mining hole locations in magenta. Coordinate system British National Grid, DSM © Environment Agency 2015.

In Figure 4-24 (a & b) small confusion areas can be seen around the ends of larger openworked trenches. This is due to the fact that the model is making predictions on cropped image tiles; if only the end of the trench is visible in the

tile, the model's strong generalisation ability works against it and it will predict a semi-circular occluded hole. As the tiles have 52% overlap these false positives are typically removed by the raster mosaic post processing step, however, due to anomalies in position and tile overlap, some remain.

The St Just and Yorkshire tests as shown in Figure 4-22 and Figure 4-24 (c & d) were carried out to examine the model's ability to generalise to different types of mines and different resolution data. In the St Just test area, whilst the algorithm did detect some of the pits and shafts, the low resolution of the LiDAR dataset and the many different buildings and taller vegetation types present in this area posed difficulties. For cases such as this where the data and ground conditions are suboptimal for automated algorithms the resulting map could still be used by a human to add interpretive information in conjunction with the other visualisation types discussed in Section 4.3.

For the Yorkshire dataset, the model surpassed its previous performance on the original Dartmoor dataset, as shown in Table 4-2. As described in Section 4.2 the Yorkshire LiDAR DSM is twice the resolution of the Dartmoor data. However, during the ground truthing exercise it was found to contain more confusion objects such as building remains, stone lined trenches and drainage culverts. The model was capable of discriminating between building foundation remains and excavated platforms from mining pits and made only two false positive detections in these areas. This is an extremely positive result and indicates the model is doing more than just looking for unnatural changes in ground elevation and is searching instead for areas that contain the features which it was trained on.

Of the false detections on this dataset, one drainage culvert was mistaken for a hole, but the geometry was such that it was only discernible as a culvert from a side view under the road unafforded to the LiDAR data. This is a limitation of all overhead remotely sensed data and is not specific to a deep learning model. Two trenches were misidentified as pits but only where dense vegetation masked their linearity causing them to appear as circular depressions on the LiDAR.

For all sites, the site verification visits revealed that many of the detected pits would be difficult to locate either on foot or from aerial photography as they are faint, shallow and reed-filled. Whilst ground truthing, many pits were near-invisible until the surveyor was within a few meters of the model's predicted location; as well, whilst traversing the sites to verify the predictions, no isolated pits were seen that were missed by the model, all missed detections were within larger excavations that had caused confusion. Figure 4-25 shows a photograph taken looking north from the Hexworthy site, aligned with the same view from the LiDAR model overlaid with aerial imagery and predicted hole locations.
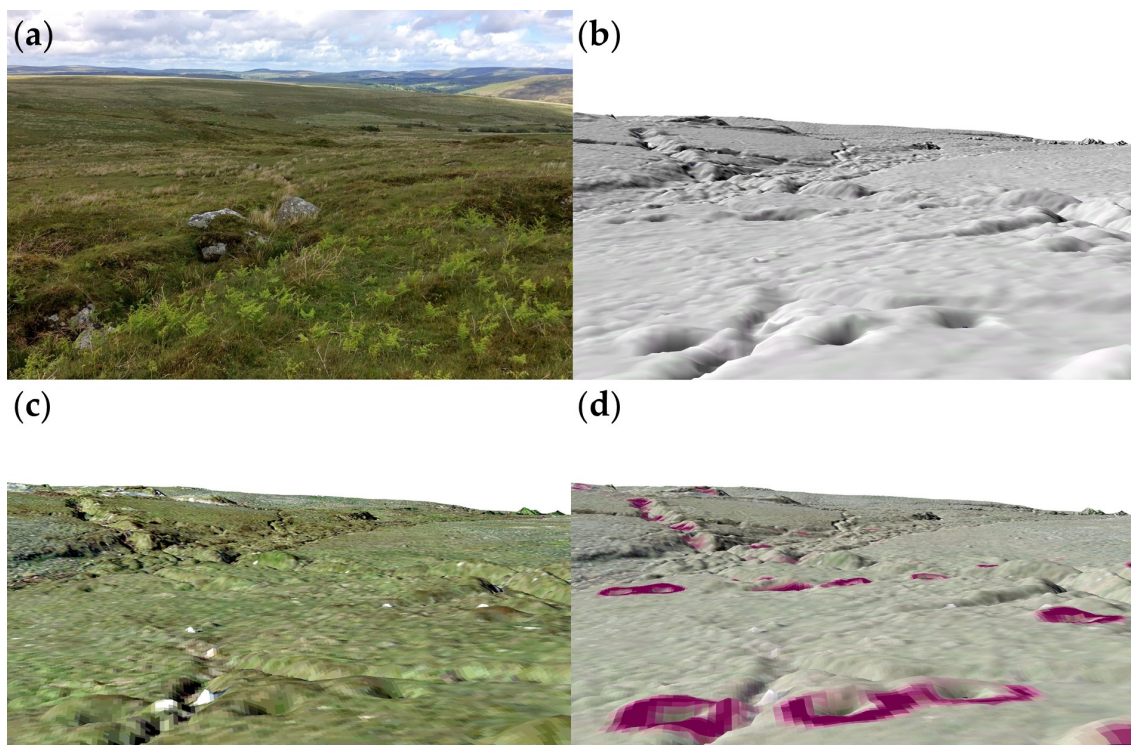
Figure 4-25. Ground level view of the Hexworthy historic mine site. (a) is a photograph taken during the verification survey, (b) shows the same scene in a hillshaded DSM, (c) includes OSGB 2010 aerial imagery and (d) includes the model's predictions. DSM and aerial imagery © Environment Agency 2015 & Digimap Getmapping Plc.

The results from the different resolution tests indicate that this model is able to generalise to new sites provided the resolution is equal to or higher than the training dataset. From these tests it appears that dataset resolution is strongly correlated with classification accuracy. These tests also show that despite being trained on one resolution of data the model is capable of being applied at a different higher resolution without the need for additional training, greatly increasing its applicability for varying quality and resolution general purpose LiDAR datasets. This is crucial as most LiDAR is not flown specifically for historic mining hazard detection purposes; therefore, detection algorithms must

be capable of working with varying accuracy and resolution datasets gathered by many agencies for diverse reasons. It is not certain whether the poorer performance on the 1m dataset was attributable more to the dataset resolution or the dataset quality. A current LiDAR campaign to re-survey the South West to a higher quality 1m resolution is ongoing (Environment Agency, 2020), once complete this new dataset could be tested with the model to determine whether satisfactory results can be obtained from higher quality 1m datasets.

The mining pit detection model created here can be rapidly run on any LiDAR DSM suspected of containing remains of historic mining activity; the approximate time to process a 1km tile including manual ArcGIS post-processing is 5 minutes. This pipeline could be easily automated further, as this research has been concerned with the ultimate performance of the deep learning model the periphery workflow has not yet been streamlined. As an output, simple GIS point layers (with their accuracy specifications of ± 20%) can be supplied to the land managers such as Dartmoor National Park and Yorkshire Dales National Park. These results are usable directly by the land managers to rapidly inform future decisions about safety, preservation and management.

4.4.8 Deep transfer learning summary

The transfer learning model developed in this research shows strong, repeatable results for the task of detecting historic mining pits. It is a novel application of knowledge from the disparate but related field of planetary remote sensing, achieving state of the art results on its allocated task. It is capable of differentiating between natural depressions and manmade ones, even in areas of occlusion and erosion. This is due to the close resemblance between the data on which the base model was pretrained and the data for the problem at hand. Other strengths of this model are its ability to output full pixelwise segmented confidence masks for any size and resolution data, alongside this workflow's integration with existing ArcGIS tools where possible to ensure ease of use and repeatability.

This model can run on large swathes of LiDAR data extremely quickly and produces meaningful results which will aid management of large scale historic mining landscapes. The model is also valuable for detecting outlying smaller

pits away from the main shafts and mineral veins. These are often unrecorded remains of earlier prospecting and information on their location can add to understanding of a site's exploitation history along with management of associated site hazards. It is envisaged that this model would be run as a first step in the prospecting process, vastly reducing the areas to be analysed in fine detail in a desktop search or fieldwork survey by a human analyst. With a false positive rate of less than 20% it does not overwhelm the analyst with incorrect predictions, providing an effective tool for preliminary site investigation and allowing confidence in the use of the model. The workflow and model presented here will allow the scale and magnitude of sites to be rapidly analysed, underpinning better management of historic mining areas.

4.5 Lineament detection – project three

The deep learning model and data processing techniques developed in the previous project can additionally form the foundation for further mining and geological applications. This project described in this final section demonstrates a lineament inference tool which builds upon the deep learning model developed in Section 4.4. Detection of geological lineaments is a significant part of regional geological analysis, providing information on local geological structures. Lineaments are a broad category of features, corresponding to mappable linear surface features which may represent a subsurface phenomenon (O'Leary et al., 1976). Traditionally, lineaments were digitised manually from airborne and spaceborne optical imagery or airborne geophysics, however, these methods are time consuming, subjective and potentially unreliable (Masoud and Koike, 2017). Alongside the time and subjectivity issues, in many climates direct fault mapping is also challenged by a lack of exposed surface rocks across large geographical extents (Yeomans et al., 2019). To address these issues, much research has been focused on developing semi-automatic methods for lineament detection, from early methods using potential field data (Blakely and Simpson, 1986) to modern MATLAB based toolboxes (e.g. TecLines; Rahnama and Gloaguen, 2014). Semi-automated methods historically have had difficulties with roads and field boundaries, along with vegetation obscuring the ground surface in optical imagery. Using LiDAR data instead of optical data can overcome some of these issues, as shown in Grebby et al. (2012).

In many areas of the world, particularly in post-industrialised nations, the marks of historic mining activity are still visible on the landscape. Rather than using the natural geomorphology to map the structural geology to infer the mineralisation, it may also be possible to infer the mineralisation directly from the mining remains. Furthermore, in some cases data on mine workings and mineralised structures may be lost, therefore, methods such as this can add value. This method could also be used to search along strike for potential shafts that may have been covered or undetected.

Primarily, semi-automatic lineament detection approaches follow a processing workflow of data representation, image enhancement, edge extraction and edge

connection (Masoud and Koike, 2017; Šilhavý et al., 2016). The input data format can be an image from an optical multispectral satellite sensor (Rahnama and Gloaguen, 2014; Soto-Pinto et al., 2013), a multiview hillshade from a Digital Elevation Model (DEM) (Masoud and Koike, 2017; Šilhavý et al., 2016), a principal curvature image generated from a DEM (Bonetto et al., 2015) or a tilt derivative image generated from airborne geophysics and LiDAR data (Middleton et al., 2015; Yeomans et al., 2019). The input image is then pre-processed to improve its characteristics for edge detection. The techniques used here vary based on the input raster type. Linear features are detected using either object-based image analysis (Middleton et al., 2015; Yeomans et al., 2019), Canny edge detectors (Mallast et al., 2011), Random Sample Consensus (RANSAC) algorithm (Bonetto et al., 2015) or variants of the Hough Transform. The Hough transform is an image processing method for detecting lines, originally proposed by Hough (1962) and described in the context of lineament detection by Wang and Howarth (1990). It is robust to line gaps and noise, making it the algorithm of choice for lineament detection in many geological toolboxes such as ADALGEO (Soto-Pinto et al., 2013) and TecLines (Rahnama and Gloaguen, 2014). In general, following the line extraction, the approaches employ some form of post-processing to improve segment connectivity and reduce noise. Historic mine workings can cause problems with traditional semi-automated methods due to the anthropogenic modification of the land surface and their lack of linearly connected features. Therefore, the use of a deep learning based method is useful to hone the lineament detection.

4.5.1 Methods

The Dartmoor dataset was chosen to be used for these lineament inference experiments, as it is the most extensive LiDAR dataset examined in this thesis. Geologically, Dartmoor National Park is underlain by the Dartmoor Granite pluton and is the largest granite pluton exposed at surface ($650 km^2$) within the Early Permian Cornubian Batholith (Scrivener, 2006). The granite is characterised by its peraluminous geochemistry and K-feldspar megacrysts (Simons et al., 2016). The area is variably mineralised and southern Dartmoor is known for tin veins of "black tin" or cassiterite (Dines, 1988). The test area for this study is focussed over the Hexworthy Mine (an amalgamation of Hootens Wheals and Hensroost mines) where the main vein structures trend

approximately NNW and subordinate veins course ESE-WNW (Dines, 1988). The area shows demonstrable surface workings and provides an ideal case study site.

The pipeline proposed in this research contains two modules, the first module detects mining pits using deep learning and the second module fits mineralisation trends to these detections using a Hough transform. Figure 4-26 shows the processing pipeline. Module 1 is identical to the workflow described previously in this chapter in Section 4.4 and will not be re-examined here. The only methodological difference is that for the purpose of geological line fitting it is hypothesised that precision should take precedence over recall, as noise from false positives may have greater negative impact than missed detections. To test this theory, the positive openness representation model and the DSM representation models were selected for further processing. As shown in Figure 4-21 the positive openness model has the highest precision and the second highest recall, whereas the DSM model has the second highest precision and the highest recall. The DSM model also exhibits a higher overall F1 score. The lower scoring representations of SLRM and negative openness were not processed further. As described previously, the outputs from the deep learning model are a binary mask image of pit probabilities for every 256 x 256 image tile. These individual image patch masks are then merged back into a single raster layer by taking the mean values. This allows every ground metre to be predicted twice, improving model robustness. This mosaicing is the final step in Module 1.

Module 1

```
┌─────────────────────────────┐
│ Create GIS from LiDAR, aerial│
│  imagery and historic maps   │
└─────────────────────────────┘
              ↓
┌─────────────────────────────┐
│ Generate training data using │
│  different data representations│
└─────────────────────────────┘
              ↓
┌─────────────────────────────┐
│ Export image patches and     │
│  convert to 8-bit values     │
└─────────────────────────────┘
              ↓
┌─────────────────────────────┐
│ Pre-process and augment      │
│    training images           │
└─────────────────────────────┘
              ↓
┌─────────────────────────────┐
│ Fine tune pretrained DeepMoon│
│  model for 4 new epochs      │
└─────────────────────────────┘
              ↓
┌─────────────────────────────┐
│ Generate predictions from test│
│           area               │
└─────────────────────────────┘
              ↓
┌─────────────────────────────┐
│ Merge prediction patches to  │
│  generate full result mask   │
└─────────────────────────────┘
```

Module 2

```
┌─────────────────────────────┐
│ Threshold image to remove all│
│      pixels below 0.6        │
└─────────────────────────────┘
              ↓
┌─────────────────────────────┐
│ Flood fill background to smooth│
│        non pit values        │
└─────────────────────────────┘
              ↓
┌─────────────────────────────┐
│ Inverse flood fill to retain only pit│
│           centres            │
└─────────────────────────────┘
              ↓
┌─────────────────────────────┐
│ Apply interactive Hough      │
│  transform to extract lines  │
└─────────────────────────────┘
              ↓
┌─────────────────────────────┐
│ Transform extracted lines to │
│      world coordinates       │
└─────────────────────────────┘
              ↓
┌─────────────────────────────┐
│ Calculate line bearings and  │
│ compare to known geological data│
└─────────────────────────────┘
```

Figure 4-26: Processing workflow diagram showing both the deep learning pit detection module and the line fitting module.

The full area mask forms the input to Module 2 for the geological line fitting algorithm. In this module the merged raster layer is pre-processed in Python using OpenCV (Bradski, 2000) to improve its characteristics for line fitting. A thresholding algorithm is applied to maintain only the pixels with a probability above 0.6 of belonging to the pit class. This removes some of the artefacts at image boundaries and also limits the amount of incorrect predictions and noise shown in the image. As it is easier to fit lines to dots rather than rings, the

147

background is filled with white using a simple flood filling algorithm, which colours all connected pixels with the specified new colour. This step removes the rings leaving just the centres. For the final pre-processing step, the image is inverted back to a black background to maintain consistency. These pre-processing steps are shown in Figure 4-27.
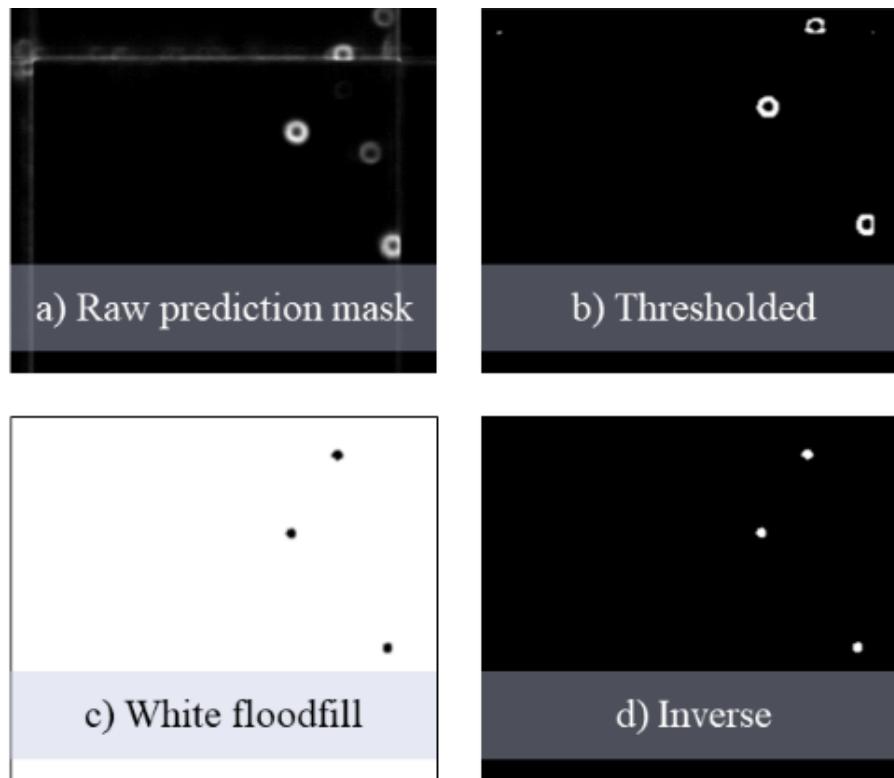
Figure 4-27: Image pre-processing to optimise the prediction result mask prior to the line fitting operation

To fit the lines, an interactive Hough transform program was created to allow the user to control the parameters of the transformation whilst viewing the fitted lines. This allows for suitable settings for the Gaussian blur filter, the edge enhancement filter and the Hough transform itself to be varied and their effects visualised. The Hough transform is sensitive to the specific geometry of an dataset, therefore, rather than set the parameters for the test dataset based on empirical assessment for each test image, as described in Rahnama and Gloaguen (2014) the interactive step allows the method to be easily used with multiple datasets of varying properties. This choice introduces compromises related to higher subjectivity and lower automation; however, it improves generalisation and usability at the proof of concept stage. As can be seen in Figure 4-28, the essential trends do not change despite different settings, only

the number and density of the extracted segments. This allows the user to adjust the detection to noise ratio appropriately.
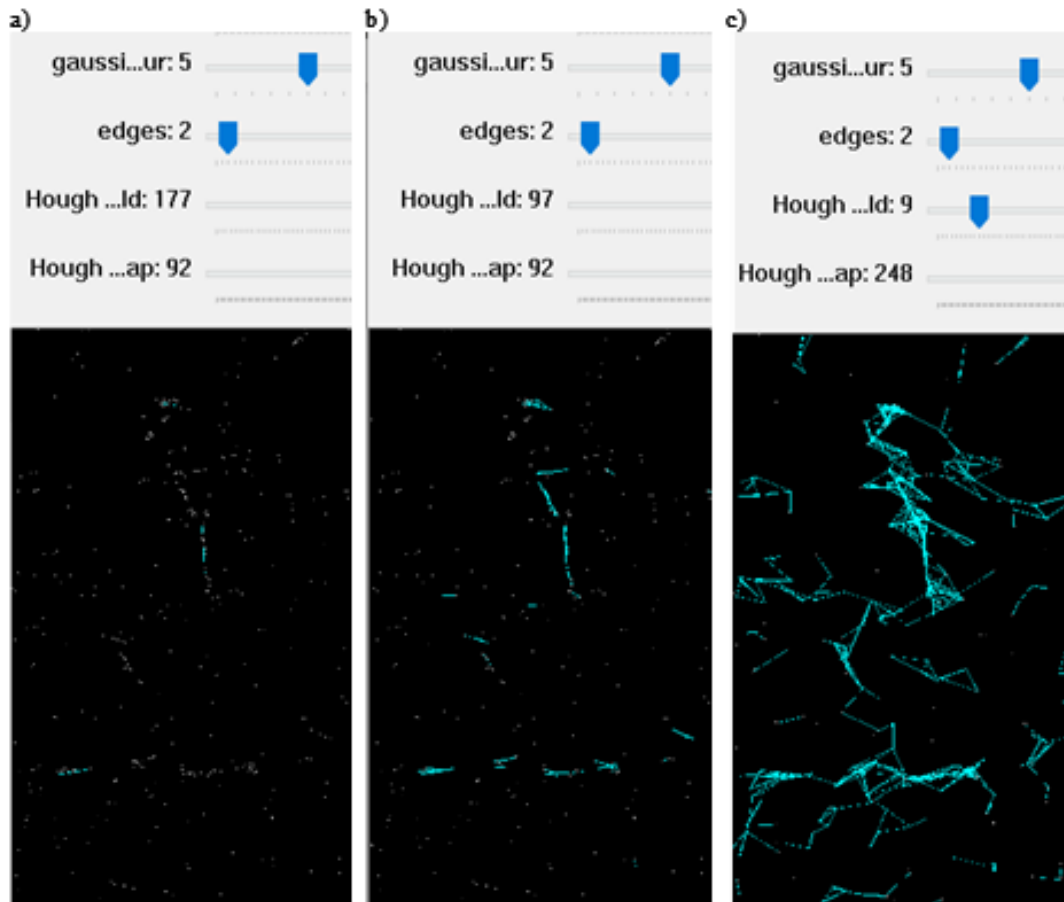
Figure 4-28: Sample of the interactive Hough transform program showing settings that are (a) too low, (b) suitable and (c) too high

After visually appropriate settings are chosen, the lines are converted from image to map coordinates and exported as georeferenced coordinate pairs. The lines can then be imported into a GIS software package for further visualisation and analysis such as bearing calculations.

4.5.2 Results and discussion

To evaluate the results of the line fitting module, the angles of the polylines generated from the Hough transform for both data representations were compared to those published in Yeomans et al. (2019), shown in Figure 4-29. The general trends show good agreement; however, direct comparison is challenging due to the differing scales of the datasets. The lineaments generated by Yeomans et al. (2019) are for the entire south west of England while those generated here are only for a 16 km^2 area of Dartmoor National

149

Park. It can be seen that the DSM shows clearer correlations with the lineaments extracted in Yeomans et al. (2019).
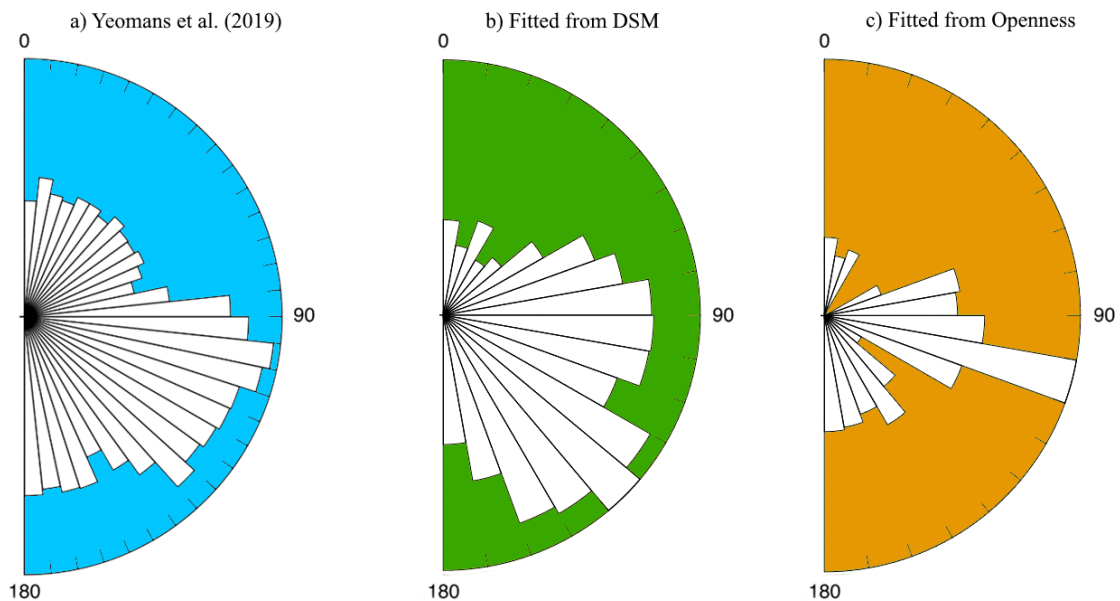
Figure 4-29: Comparison of half rose plots showing dominant ESE trends. (a) shows regional polyline angles observed by Yeomans et al. (2019), (b) shows polyline angles from lines fitted in this study using the model trained on the DSM representation and (c) shows the polyline angles from lines fitted using the positive openness representation.

Alongside the lineaments from Yeomans et al. (2019), the generated lines were also compared to those provided by the British Geological Survey (BGS) in their 1:50,000 linear geology vector map layer (BGS, 2016), shown in Figure 4-30.
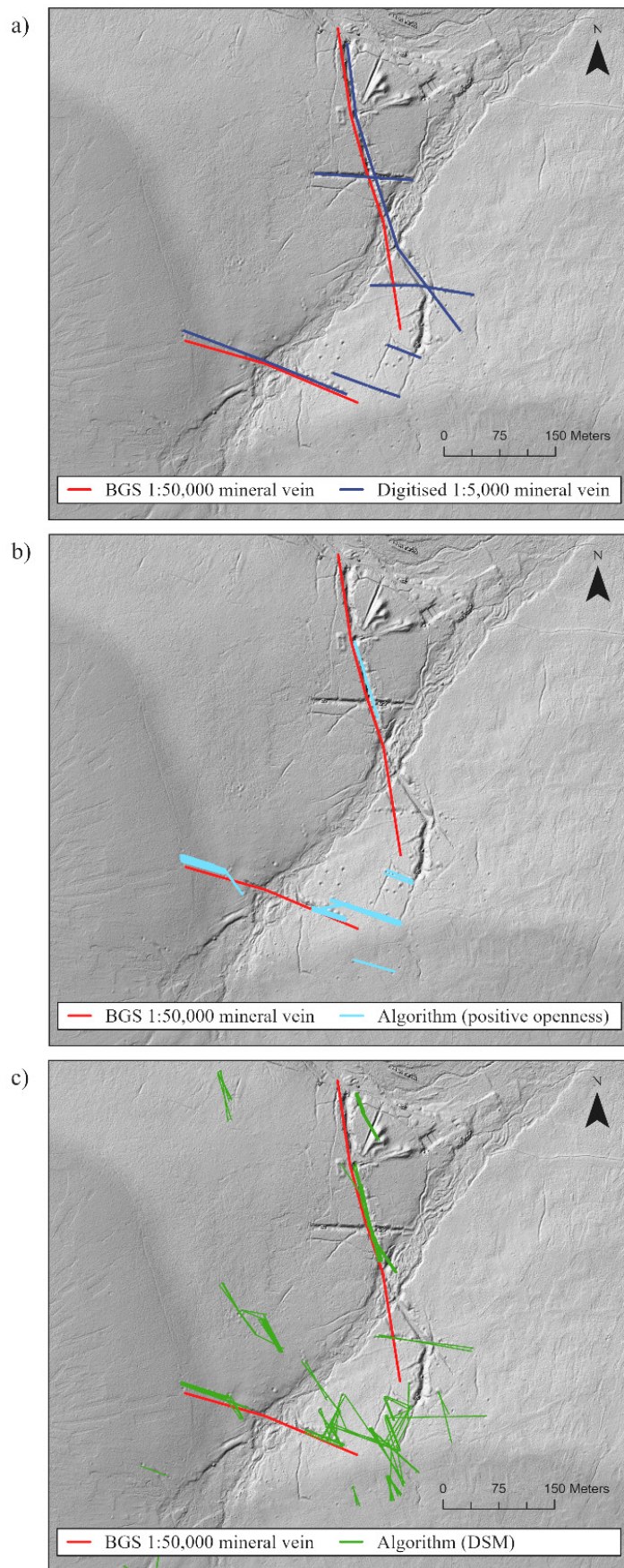
Figure 4-30: Results showing the British Geological Survey's 1:50,000 linear geology layer compared to (a) manually digitised high resolution lines and (b) the lines generated by the automated algorithm. Geological Map Data BGS 2020, base DSM © Environment Agency 2009

Again, the differing data scales proved challenging, with scale related imprecisions noticeable in the BGS data when viewed at 1:5,000 due to a resolution of 50m at 1mm line thickness. Figure 4-30a shows the BGS data alongside higher resolution probable mineral vein locations, digitised manually from the LiDAR data. It can be seen that several smaller linear features are not present on the BGS layer, along with a deviation in angle on the southern end of the main north-south vein. Figure 4-30b shows the automatically extracted lines from the positive openness representation plotted against the BGS data. It can be seen that the general trends are positive, with the algorithm picking up several line angles more precisely than the 1:50,000 layer, but that it does not extend far enough in many instances. For the additional mineral vein locations inferred in Figure 4-30a, two were picked up by the algorithm, and two were missed. It is hypothesised that as the algorithm is fitting lines to densities of detected pits, the shorter line segments are due to the CNN not detecting a large enough cluster of points at the extremities of the lines, leading to missed sections. This can be attributed to the lower recall of the positive openness predictions. Another factor is that neither CNN model was not trained to detect trenches that do not contain pits; the two missed east-west veins are primarily trenches containing very few pits, likely the cause of the missed line detections. Figure 4-30c shows the results from the lines automatically extracted using the predictions from the DSM representation. There are many more detected lines and the result appears noisier than that shown in Figure 4-30b, though the more southerly missed east-west trench has been picked up.

4.5.3 Lineament detection summary

The geological lines generated using this technique correlate with the trends of the well-known lineaments in the Dartmoor area, both those semi-automatically extracted from LiDAR data by Yeomans et al. (2019) (Figure 4-29) and those published by the British Geological Survey (BGS) in their 1:50,000 mapping products (BGS, 2016)(Figure 4-30). The results using the positive openness representation provide cleaner results when viewed on a map, however, the results from the DSM representation are more successful at detecting missed lines at high resolutions and show greater directional agreement on the half-rose plot. Further work to incorporate trench identification into the deep learning model would improve detection accuracy, alongside further refinements of the

Hough transform parameter selection process. This preliminary work demonstrates that the lines produced from this technique can aid geological interpretation in regions of historic mining activity, particularly where records have been lost or are incomplete.

4.6 Summary

The three projects which make up the research in this chapter examined different but interconnected ways that LiDAR data can be used to add to the knowledge of a legacy mining environment. LiDAR is particularly useful for detecting landscape signatures which may be overgrown or eroded, making them difficult to distinguish with optical imagery. The first project introduced the concept of multiple data visualisations and investigated how these can be used to aid human interpretation and to extract additional details from the gridded LiDAR elevation values. Eleven different visualisation representations were generated for an area featuring many of the typical Cornish historic mining landscape features and the strengths and weaknesses of each representation type were observed.

The second project developed a deep transfer learning method for successfully detecting trial pits, shallow workings and shafts. This method was trialled on multiple LiDAR datasets of different resolutions and landscape types. This project revealed that the deep transfer learning technique was highly capable of detecting these mining remains from 0.5m and 0.25m LiDAR datasets; however, the accuracy was notably reduced when using the 1m dataset. This result highlights the importance of good quality high resolution datasets for future deep learning projects.

The final project explored how further subsurface geological knowledge can be inferred from the results of the deep learning model. This project demonstrates that the value of deep learning models does not end with object identification, but rather is one step on the path towards intelligent landscape modelling both on surface and underground.

Overall, the results from these interconnected projects demonstrate that LiDAR data can be used to aid understanding of past mining landscapes, both by enhancement for human interpretation and by applying automated detection techniques. These techniques can be used to rediscover sites, monitor legacy risks, add information for geological exploration and expand the historic record. With the increasing global availability of high resolution LiDAR datasets, the methods described here have wide ranging applicability for many countries with poorly documented legacy mine sites.

Using satellite imagery and deep convolutional neural networks to detect land cover changes

Chapter overview

This chapter investigates how deep learning models can be created for multispectral satellite imagery where no suitable transfer learning model can be found. The application focused on in this chapter is detecting Artisanal Scale Mining (ASM) in Ghana from Sentinel-2 multispectral satellite imagery. This chapter develops techniques for modifying CNN architectures to ingest an arbitrary number of image channels and could be used for many multispectral and hyperspectral applications. This chapter is based on the paper **'A Sentinel-2 based Multispectral Convolutional Neural Network for Detecting Artisanal Small-scale Mining in Ghana: Applying Deep Learning to Shallow Mining'** published in the journal Remote Sensing and Environment[13]. In addition to the development of the deep learning model, this chapter also investigates how the results from this model can be used to monitor small scale mining land use changes across the study area.

5.1 Introduction

Artisanal and small-scale mining (ASM) is a rapidly expanding source of livelihood for many in the Global South, particularly in rural areas. It is estimated that over 40 million people are directly working in ASM across 80 countries, with a further 150 million people dependent on ASM indirectly (Intergovernmental Forum on Mining Minerals Metals and Sustainable Development, 2017). ASM activities can be considered to exist on a spectrum of formality (McQuilken and Garvin, 2016), from highly illegal mining either within nature preserves (Boadi et al., 2016) to fully licensed, environmentally compliant and formalised small-scale mining (Hilson, 2002).

Whilst ASM generates vital socioeconomic benefits for communities and countries it is also associated with environmental and social problems such as

---

[13] The candidate is the first author of this paper and the authorship contribution statement is as follows: **Jane Gallwey**: Methodology, Conceptualization, Software, Investigation, Writing – Original Draft, Visualization. **Carlo Robiati**: Investigation, Conceptualization, Validation, Data Curation. **John Coggan**: Supervision, Project Administration, Resources. **Declan Vogt**: Supervision, Writing – Review and Editing. **Matthew Eyre**: Conceptualization, Writing – Review and Editing, Supervision, Resources.

land degradation, deforestation, water pollution, illegal immigration, low worker safety and child labour (Corbett et al., 2017; Hilson, 2002; Intergovernmental Forum on Mining Minerals Metals and Sustainable Development, 2017). The excavation of ground for mining causes widespread deforestation and leaves dangerous unstable pits that fill with standing water, creating breeding grounds for malarial mosquitoes (Bansah et al., 2018). The unregulated mineral processing also leads to heavy metal pollution, especially increased mercury levels (Bansah et al., 2018; Telmer and Stapper, 2007).

Appropriate management of ASM activities is critical, with its benefits directly linked to 9 of the 17 United Nations Sustainable Development Goals (Hilson and Maconachie, 2020). However, effective ASM management is inhibited by factors such as the informal and undocumented nature of the sector, legacy of inappropriate policies, limited government resources and the remote locations of mine sites (Corbett et al., 2017; Hilson and Gatsinzi, 2014). Detailed, accurate and inexpensive geoinformation about ASM activities would aid legislative pathways by providing rapid mapping resources to support small scale licensing claims, one of the major barriers to effective legislation (McQuilken and Garvin, 2016). These datasets would also support the timely tackling of environmental problems by focusing enforcement and remediation efforts where they are most needed.

Remote sensing data sourced from earth observation satellites could generate this geoinformation, although in the past their applicability for mapping ASM activity has been hindered by the resolution of the sensors and the accuracy of the classification techniques (Asner et al., 2013). Landsat and MODIS, the most prominent satellite imagers of the 1990s and 2000s, have resolutions of 30m and 250m respectively, making identification of the often small ASM alluvial mine workings problematic as workings can cover less than half a pixel, even at Landsat resolution. ASM activity visible from space is primarily surface mining, usually alluvial in nature and presenting as ribbons of clustered pits (Snapir et al., 2017). They are normally located on or near watercourses, along with dry pits and bare earth waste piles (Owusu-Nimo et al., 2018), ranging in size from less than half a hectare to several hundred hectares for the larger clusters.

The past five years have seen the launch of advanced satellite imaging systems with greatly increased spatial resolutions. These include the European Space Agency's (ESA) Sentinel-2 platform (Berger et al., 2012) which provides multispectral imagery with 10m resolution in four visible and near infrared bands and 20m resolution in four further infrared bands (Drusch et al., 2012; European Space Agency, 2015). Spectrally, the alluvial workings are highly reflective in the near infrared bands between 700nm and 875nm and display the greatest contrast from vegetation, rooftops and open ground in the short-wave infrared bands between 1600nm and 2500nm.

Delineating ASM related deforestation is a subset of Land Use/Land Cover (LULC) mapping. The most common approach to LULC mapping performs pixel-based classification based on the spectral signatures of the classes of interest, utilising machine learning classifiers. Pixel based classifiers have been used for mapping ASM activity in Ghana (Boakye et al., 2020; Kusimi, 2008; Obodai et al., 2019), Burkina Faso (Leroux et al., 2018) and Brazil (Lobo et al., 2018). The reported omission/commission errors range from 8-40% for the mining class, indicating a large variability in the accuracy of this method (Boakye et al., 2020; Obodai et al., 2019).

A subset of pixel-based methods, used primarily in deforestation studies, leverages spectral mixture analysis to detect sub pixel changes. Multiple studies have used this technique to detect ASM related deforestation (Asner et al., 2013; Asner and Tupayachi, 2016; Caballero Espejo et al., 2018). Typical omission and commission errors from these methods are in the region of 10-25% (Asner and Tupayachi, 2016; Caballero Espejo et al., 2018), with some difficulties encountered in correctly classifying the water pools within the alluvial mine sites (Caballero Espejo et al., 2018). The greatest weakness of the pixel-based classifiers is their lack of spatial context. Each pixel is considered individually, leading to a 'speckled' effect in the resulting classification maps (Blaschke et al., 2014) which can be mitigated to some degree by postprocessing but not wholly removed (Kelly et al., 2011).

Object based approaches attempt to address this lack of spatial context by first segmenting the image into regions based on spectral similarity then classifying the regions rather than the individual pixels (Blaschke et al., 2014). These

approaches have seen relatively little application for detecting ASM, but a successful implementation for general rainforest LULC change analysis (including mining) is found in Souza-Filho et al. (2018) with commission/omission errors in the range of 10-30%. However, finding adequate parameters for the initial object segmentation can be labour intensive and relies on good domain knowledge (Nuijten et al., 2019). All methods to date have required a significant level of human input, including feature and segmentation parameter design, manual cluster selection and manual relabelling of incorrect pixels.

Outside the field of traditional remote sensing, techniques from computer science could provide a solution: since 2010, advances in the available processing power of Graphics Processing Units (GPUs) have allowed a type of deep artificial neural network known as a Convolutional Neural Network (CNN) (LeCun et al., 1989b) to emerge dominant for most image processing tasks (Gu et al., 2018). CNNs are inspired by biological visual cortexes and work by adaptively and automatically learning spatial dependencies and hierarchies of features from gridded data. They can approximate highly nonlinear functions whilst maintaining spatial connectivity between pixels (Goodfellow et al., 2016).

Considerable research has been carried out in the broader remote sensing community as to how to modify CNNs for LULC tasks; a review is given in Ball et al. (2017). The main identified challenges to adoption were found to be the limited availability of large amounts of pre-labelled training data and the multiple channels found in multispectral and hyperspectral imaging systems (Nogueira et al., 2017; Signoroni et al., 2019). Deep learning based LULC applications have been successfully implemented using Sentinel-2 data, but the images are most commonly clipped to contain only the first three or four bands, as seen in Kroupi et al. (2019) and Wurm et al. (2019), to maintain compatibility with non-multispectral deep learning models. Very few studies have modified deep learning models to ingest true multispectral data, with Kemker et al. (2018) describing the most comprehensive and promising study to date. Their method tested two different deep architectures for learning 18 class segmentations on 8 channel images, including testing pretraining with synthetic imagery. Their approach achieved high per class accuracies on the larger classes from their dataset; however, some smaller classes proved more challenging, although

their proposed method significantly outperformed their other tested methods on those classes. No published work has described the application of true multispectral deep learning for classifying ASM and deforestation.

LULC mapping allows delineation of ASM activity, but to extend from delineation to monitoring, the model must include a temporal change detection element. Post-classification change detection is the method used by most ASM mapping studies to date (Asner and Tupayachi, 2016; Boakye et al., 2020; Kusimi, 2008), but it is sensitive to the errors in the original classification maps and can produce imprecise results especially for smaller land cover classes such as mining or urban. The improvements in classification accuracy achievable by using a CNN could circumvent the weaknesses of post-classification change detection methods and lead to simple and effective change mapping.

By bringing together recent advances in deep learning and satellite sensor technology, this chapter proposes a new method to automatically map the extent of alluvial ASM activities at a hitherto unprecedented level of accuracy and detail for minimal cost, allowing for effective ASM monitoring. The objectives of this chapter are to: i) design a multispectral CNN model capable of distinguishing between mining, built/developed and vegetation land use classes from freely available Sentinel-2 imagery, ii) benchmark the model's performance against other classification methods, iii) demonstrate the performance of this model for mining area detection across a large spatial and temporal range of images, with minimal human input and iv) provide maps which could be used to analyse the impacts of ASM policies over the studied time period. The novel contributions of this chapter include: the development of a multispectral deep learning model which is significantly more accurate than existing techniques for detecting ASM, a full processing pipeline for monitoring ASM and urbanisation via spatial and temporal mosaicing, and the production of a large scale dataset showing the extent and expansion of ASM in the Ghanaian case study area from 2015-2019.

5.2 Methods

This research work for this chapter built a multispectral CNN designed to detect mining and built environments from Sentinel-2 satellite data. Alongside this,

several traditional pixel-based machine learning classifiers were employed to benchmark the CNN method against current practices. Investigation was then carried out to assess how the design of the CNN impacts performance across multiple validation patches. Once the model design was finalised, an area of more than 6 million hectares was analysed over 4 years, with a temporal sampling of once per year. To improve the classification reliability of the yearly maps, ensemble methods using probability averaging were used to generate the most probable class from multiple images per season.

5.2.1 Case study area

Ghana was chosen as the case study area as it is at the front line of the ongoing land use changes happening throughout mineral endowed areas of the Global South. Ghanaian rainforest is being deforested at the world's fastest rate, with over 60% more forest lost in 2018 than 2017, initially reported by Global Forest Watch (Weisse and Goldman, 2019) and verified using Sentinel-2 data by Dekker (2019). This is due to mining, agriculture, logging, fires, and urban expansion (Boadi et al., 2016).

Alongside deforestation, ecological issues arising from Artisanal Scale Gold Mining (ASGM) in Ghana include mercury, arsenic and other heavy metal contaminations, decreased water quality and land disturbances (Rajaee et al., 2015). Ghana's economy is heavily reliant on gold and cocoa exports, which often compete for the same land (Snapir et al., 2017). Mining comprised 35% of all national exports in 2014, of which 65% came from large scale mines and 35% from ASM (McQuilken and Garvin, 2016). This is one of the highest percentages attributable to ASM worldwide, largely related to the influx of Chinese miners since the mid-2000s, bringing with them heavy machinery and causing an accelerated rate of land degradation (Botchwey et al., 2019).

Government policy towards ASM has been varied. Initially Ghana was a global leader in ASM formalisation with the Small-Scale Gold Mining Law in 1989 (McQuilken and Garvin, 2016); however, the licensing process was bureaucratically challenging and precluded access to the system for most poverty driven artisanal miners, propelling the sector into increasing informality (Hilson, 2001). To address the issues associated with ASM, in 2013 the government set up a National Task Force to curb operations; however, it was

perceived by many as a façade to placate the media and the public (Hilson et al., 2014). All small-scale mining including legal operations was banned in April 2017 as part of the Task Force's Operation Vanguard (Botchwey et al., 2019; Ministry of Lands and Natural Resources, 2019), with the ban partially lifted in December 2018 (Bansah et al., 2018).

The study area boundary chosen is the Ghanaian territory corresponding to the Precambrian West African Craton (Labou et al., 2020), shown in Figure 5-1. This area covers the major gold belts of Ghana and corresponds to all of the Western and Central provinces and large parts of the Brong Ahafo, Ashanti and Eastern Regions. Additionally, the area contains the ten districts inspected by Owusu-Nimo et al. (2018) in a fieldwork based study, allowing results from our model to be examined within the context of existing ground truth data. The findings from Owusu-Nimo et al. (2018) can provide valuable supplementary information for interrogating the results of our model, despite the scales and study types not being directly comparable.
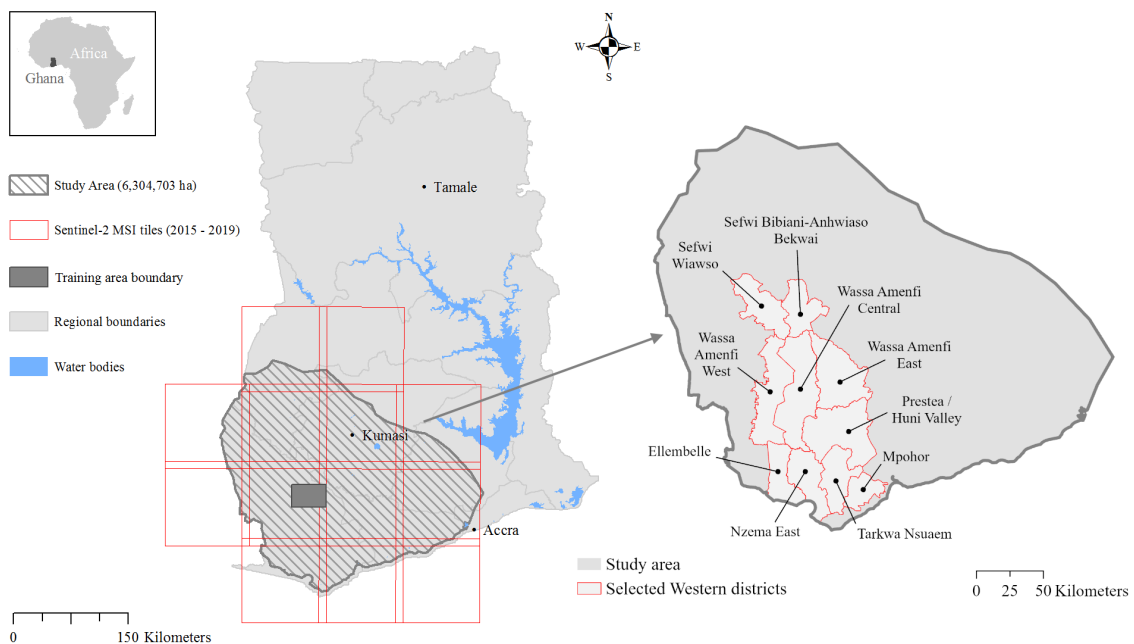
Figure 5-1: Overview map of study area, Sentinel-2 tile coverage and training data boundary. Righthand sub-image shows the ten districts common to the Owusu-Nimo et al. (2018) study.

5.2.2 Datasets

The imagery chosen for this research was sourced from the ESA's Sentinel-2 MultiSpectral Instrument (MSI). This platform was selected because: i) it offers spatial resolutions of 10m or 20m per pixel (depending on band) in the visible

and infrared, adequate to detect most alluvial mining sites; ii) at these spatial resolutions it provides ten spectral bands, ranging from 490nm to 2200nm, delivering sufficient spectral resolution, particularly in the short wave infrared range; iii) it is 'future-proof' with operation planned until 2025 with potential for extension to 2030; and most importantly iv) it is open access data (European Space Agency, 2015; van der Meer et al., 2014). Forkuor et al. (2020) state that open access satellite data can be valuable in assisting data-scarce developing countries to measure progress towards Sustainable Development Goal targets. Whilst other sensors such as Planet's PlanetScope satellites offer faster revisit times and higher spatial resolutions, as explored by Shendryk et al. (2019), it was deemed important to avoid potentially costly commitments to private companies when designing a tool to be used for sustainable development purposes. This, alongside its higher spectral resolution informed our choice of Sentinel-2 data.

The Sentinel-2 data was downloaded as 100km x 100km UTM registered orthorectified tiles, known as 'granules' in the Sentinel-2 nomenclature (European Space Agency, 2015) in either L1C (top of atmosphere) or L2A (bottom of atmosphere) formats. The total study area partially or wholly intersected 13 different tile footprints. Between three and five granules containing less than 20% cloud coverage were downloaded for every dry season in Ghana (November - March), hereafter known as the temporal period. Three tile footprints intersected the study area by less than 4%. To reduce dataset size only one granule per temporal period was downloaded for these areas, as it was possible to manually select a granule with completely clear skies over the small area of interest. In total 211 granules were downloaded, details of which can be found in supplementary file S1. The ESA command line program sen2cor was used to convert the L1C tiles to L2A (Gascon et al., 2017). The six bands with spatial resolutions of 20m were upsampled to 10m using bicubic interpolation (Vaiopoulos and Karantzalos, 2016). For clarity, pixels corresponding to lakes, known large scale mining operations and a 5km buffer along the Atlantic coast were manually masked out and do not contribute further to the analysis.

To provide training data to the model, a 30km x 45km area centred at 2°12'W 5°55'N in the Western province was digitised, the location is shown in Figure

5-2. This area has high mining activity, helping to mitigate the issues arising from imbalanced classes, where results can be biased due to the larger number of examples in one class in relation to another (Peng et al., 2019; Shendryk et al., 2019; Wurm et al., 2019). Initially, the Sentinel-2 granule from the 23[rd of] January 2016 was manually digitised into the three macroclasses of mines, built areas[14] and vegetation. This granule was chosen as it was the earliest cloud free granule available, allowing high resolution RGB imagery (0.31-0.5m pixel size) imaged by DigitalGlobe's WorldView satellites in 2015 to be used for additional reference (DigitalGlobe, 2015). In the Sentinel-2 imagery, the visibility of mines and settlements generally was sufficient to allow confident human digitisation; where confusion areas occurred the higher resolution imagery was consulted, with due consideration given to the temporal shift between the datasets. The initial 2016 digitisation was used as a starting point for digitising two further training tiles covering the same geographic area but imaged in on the 11th of February and the 2nd of April 2019. These additional temporal tiles were included to obtain training examples under differing radiometric and atmospheric conditions, improving the final model's generalisation ability. Training area labels were digitised by one operator and checked by another. This check indicated high reliability of generated labels, with only small areas and class boundaries displaying disagreements, mainly arising from the subjective nature of delineating precise mine boundaries. When generating the training boundaries, the emphasis was placed on delineating land use rather than pixel level land class, therefore, small patches of vegetation within mined areas were included in the mining class rather than the vegetation class. This strategy more closely represents the actual land area degraded by mining and also generates smoother class boundaries for later interpretation.

For validation, two 5km x 5km areas were digitised using the same methodology. Validation Tile A was generated to the north of the training area using the same 2016 granule, while Validation Tile B was generated from a 2018 granule not used for training. Figure 5-2 shows the training and validation areas. These two validation tiles allow performance to be examined under both known and unknown radiometric conditions. For testing, eight additional 5km x

---

[14] The built class includes areas of bare earth along roads and around buildings but does not generally include fallow fields unless contiguous with other dwellings.

5km tiles were generated using the same workflow. Three of these tiles were extracted from the same larger 2016 Sentinel-2 granule used for training; these make up the known radiometric test set. Another three were extracted from different years and locations within the study area, these make up the unseen Ghana test set. Lastly, two tiles, from Suriname and Indonesia, were created to investigate the model's global generalisation ability. All tiles are representative of the type of mining landscape found across the region, containing clear examples of all classes and minimal cloud occlusion. Full details of the test tiles are found in **Appendix C-1**.
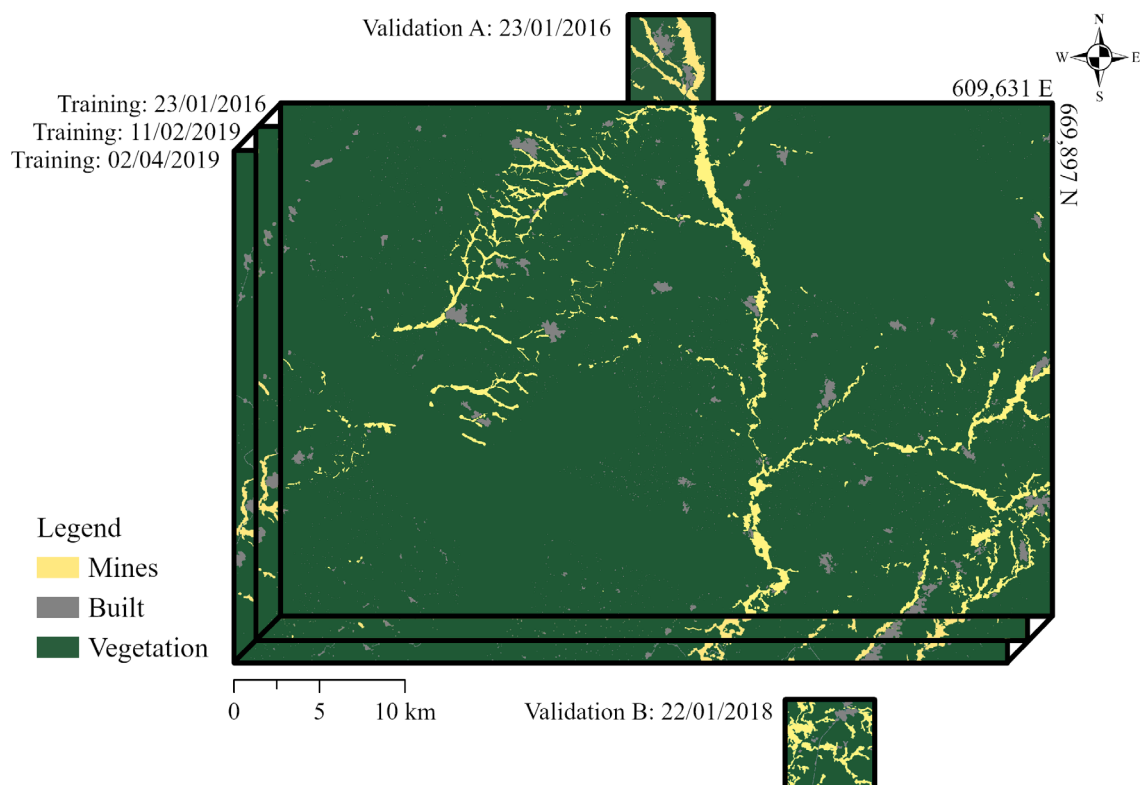
Figure 5-2: Locations of the manually digitised training and validation areas. Dates refer to the acquisition date of the imagery, the training area boundary is identical for all years, the offset is used to indicate temporal change.

The contiguous training and validation coverages were then converted to a patch-based structure, similar to that described by Wang et al. (2019). This structure was chosen for its ability to ingest large scenes in a memory efficient manner, whilst decreasing overfitting tendencies. From the full labelled training scene of 9,000 x 4,500 pixels, 16,000 images of size 256 x 256 pixels were randomly extracted. During training, these images were augmented with random X and Y reflections and random rotations from 0° to 90° every epoch. From this augmentation and patching strategy, it is possible to feed the model

480,000 individual training patches generated from the initial 9,000 x 4,500 reference image, leading to robust generalisation from limited human annotated source data.

5.2.3 Cloud removal

In areas of the tropics where much of the world's ASM activities are located, it is not uncommon to obtain only one completely cloud free image per year, despite weekly satellite revisit times. The area of southern Ghana chosen for this study is one of the cloudiest regions of the world (Coulter et al., 2016); therefore, utilising partially cloudy images is essential for any robust ASM monitoring system. Sentinel-2 software comes with in-built cloud removal tools; however, these are based on the high water reflectance of clouds in certain bands (European Space Agency, 2019) which also occur in alluvial ASM ponds. When tested, the Sentinel-2 cloud removal tools were unable to distinguish between clouds and ASM, removing large areas of valid pixels. To replace the Sentinel-2 cloud removal tool, pixels exceeding a set reflectance threshold in all three of the visible wavelength bands were masked out, as the main differentiable feature between the clouds and the ponds is their hue in the visual spectrum. This method was effective on the majority of the isolated cumulus clouds found during the dry season. Cloud edge boundaries and high cirrus clouds were not removed by this technique but eliminated later in the ensemble averaging stage, allowing potential detections to be made under light cloud obscuration. Figure 5-3 shows a comparison of the different cloud removal techniques on a typical cloudy Sentinel-2 image.
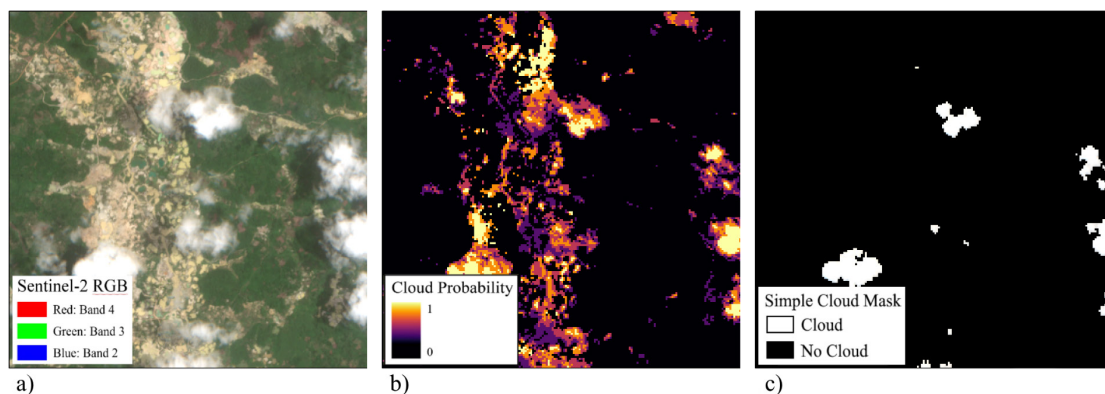
Figure 5-3: Comparison of cloud removal techniques over a sample cloudy image tile (a) using Sentinel-Toolbox (b) and our novel removal strategy (c).

5.2.4 Model development

The model chosen for this research is a type of U-Net, a successful encoder-decoder network designed by Ronneberger et al. (2015) for biomedical image segmentation. It has an encoder path consisting of multiple convolutions, ReLU activations and maxpooling operations, followed by a decoder path which upsamples the lower level stacks with the aid of skip connections from the higher resolution layers, combined with further convolutions and ReLU activations. It has been proven to be a highly effective architecture, particularly for problems with limited training data (Ronneberger et al., 2015) and has been the model of choice for multiple remote sensing applications (Iglovikov et al., 2017; Peng et al., 2019; Zhao et al., 2019). The implementation used here is adapted from a multispectral U-Net available online[15] in MATLAB format (MathWorks, 2019). Our version, which is modified to ingest Sentinel-2 data is illustrated in Figure 5-4**.** The model was created and trained with MATLAB on a desktop computer with a single NVIDIA Titan X GPU and 64 GB of RAM. Using this hardware setup, training time was approximately 16 hours.
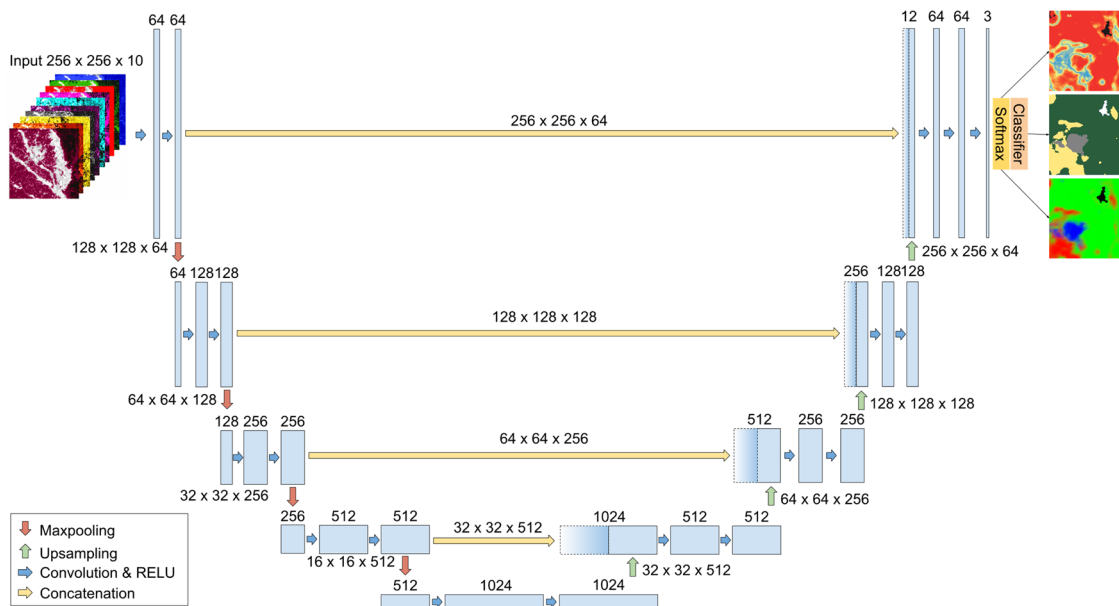
Figure 5-4: Schematic of the U-Net architecture used for the model. The input is a 10 channel multispectral image of 256 x 256 pixels and the outputs include: a 1 channel prediction mask, a 1 channel prediction probability layer for the highest class probability and a 3 channel prediction probability layer for all classes, all at the same resolution as the input image.

---

[15]https://uk.mathworks.com/help/images/multispectral-semantic-segmentation-using-deep-learning.html

Due to the small training set size, strong regularisation strategies were employed to minimise overfitting. Alongside the data augmentation discussed in Section 5.2.2, two 50% dropout layers (Srivastava et al., 2014) were added, one at the end of the encoder path and another at the model's midpoint. Stochastic Gradient Descent with Momentum (SGDM) was chosen as the optimiser as it has been shown to provide better generalisation abilities than adaptive optimisation methods (Wilson et al., 2017). The momentum hyperparameter was fixed at 0.9; this value is typical and effective in practice (Géron, 2017; Goodfellow et al., 2016; Hinton, 2012). Further hyperparameters were determined by training multiple models; during training the validation patch loss was monitored and after training a manual inspection of the output prediction masks for the two validation tiles was carried out. An adaptive learning rate was used, initiated at 5e-2 and reduced by a factor of 10 every 10 epochs. This was chosen iteratively by examining loss curves during training. Weight decay was set to 1e-4. Increasing this value to 5e-4 did not improve model performance. With the hardware configuration described above, it was found that a field of view of 256 x 256 and a mini-batch size of 16 fitted into memory and provided high quality results. Neither increasing the field of view to 512 x 512 nor increasing the mini-batch size to 128 improved model accuracy, with both requiring longer training times. The mini-batches were shuffled after each epoch to increase convergence and improve accuracy (Bengio, 2012). Automated early stopping was not used, however, human monitoring of validation loss during training showed a divergence in training versus validation accuracy after 30 epochs; therefore, the number of training epochs was set to 30, using the principals described in Chapter 2, Section 2.4.6.4.

Loss was calculated using binary cross-entropy, as used in the original U-Net implementation (Ronneberger et al., 2015). As the dataset contains highly unbalanced classes, experiments were carried out to obtain the best class weighting strategy for the loss function. Initially, inverse proportional weighting was trialled (Huang et al., 2016); however, this resulted in undue importance being given to minimising omission errors in the rarer classes, resulting in notably decreased accuracy. Halving the weight of the majority vegetation class provided the most balanced results and was the strategy used for the final trained model.

5.2.5 Pixel based method comparison

To compare the results obtained from the U-Net model, benchmarking was carried out against several pixel-based machine learning classifiers. The first classifier tested is Spectral Angle Mapping (SAM), implemented in the Supervised Classification Plugin (SCP) (Congedo, 2016) of the opensource QGIS software. Processing was carried out using the 2016 training data following the workflows described in Congedo (2016), Boakye et al. (2019) and Obodai et al. (2019). Alongside the SCP workflow, another experiment was undertaken to determine whether stronger machine learning classifiers such as a Multi-Layer Perceptron (MLP) or a Random Forest (Breiman, 2001) were better able to model the pixelwise relationships between the spectral reflectances and the land cover classes.

To test this hypothesis, MLP and Random Forest classifiers were built with the Scikit-learn libraries in Python (Pedregosa et al., 2011). The MLP used is a simple shallow model with two fully connected hidden layers of 10 and 5 nodes respectively, ReLU activations, an adaptive learning rate and Adam solver. For the Random Forest classifier, 100 trees were used for the model. For both these models, class rebalancing was carried out to reduce the vegetation class to five times the mining class. A hyperparameter search was undertaken using 4 fold cross-validation on the training data; however, negligible differences were observed. Processing involved unrolling each image into a single *n x m* vector for input to the classifiers, where: *n = image width x image height, m = 10*. Post-classification, the prediction masks were reshaped back to the original image dimensions to assess their accuracy against the human generated masks.

5.2.6 Post processing

Following the CNN model prediction stage, a result stacking strategy was developed to combine predictions across temporal periods. Ensemble methods such as stacking have been shown to improve the performance of most machine learning classifiers with the greatest gains inversely proportional to model correlation (Dieterich, 2000). Prediction stacking was also necessary for this application in order to mosaic together multiple partially cloud-occluded prediction maps. The U-Net model was modified to output both the classifications from the final layer and the per class prediction probabilities from

the softmax layer. In Figure 5-5, (a) shows the Sentinel-2 RGB data as a reference, (b) the single predicted class labels, (c) the model's confidence in its predictions and (d) the prediction probabilities for every class mapped to the RGB colour channels. As there are only three land cover classes for this application, added value was gained by mapping the per class probabilities to each of the colour channels, allowing intuitive visual interrogation of the model's predictions. This probability visualisation map goes some way towards addressing a criticism often levelled at deep learning models that they are a 'black box' solution that does not reveal the processes which led to the results. The probability mapping provides insights into how the model thinks and why it makes mistakes. Muddied colours represent areas of confusion, where the model has predicted similar probabilities for multiple classes, as seen in Figure 5-5 (d).
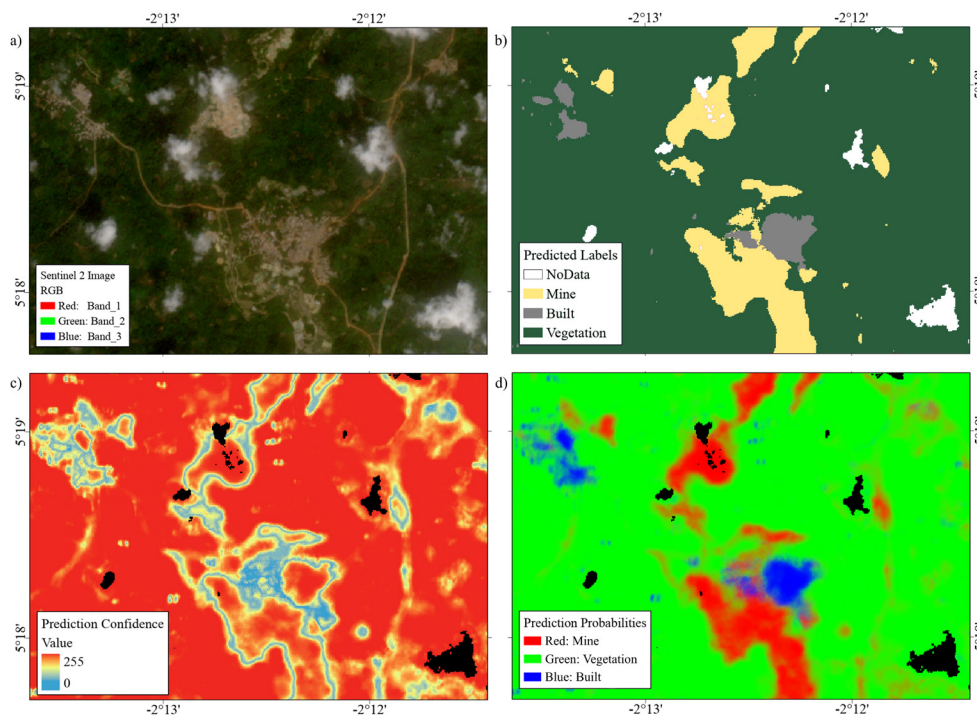
Figure 5-5: Model outputs showing a) the Sentinel-2 RGB data as a reference, b) the single predicted class labels, c) the model's confidence in its predictions and d) the prediction probabilities for every class mapped to the RGB colour channels.

Whilst the simplest model ensemble would combine the final classified images for each temporal period, utilising the full prediction probability images allows greater weight to be given to the more confident predictions. The models were combined as shown in Equations 5-1 to 5-2. For a dataset of *l* number of images, each with 3 probability classes mapped to the channels *r, g, b* and pixels *i, j as follows*:

169

$$result\ (i,j) = max(r_{(i,j)},\ g_{(i,j)},\ b_{(i,j)})\qquad (5\text{-}1)$$

$$where\quad r_{(i,j)} = \frac{\sum_1^l dataset(i,j,r,l)}{valid_{(i,j)}}\quad g_{(i,j)} = \frac{\sum_1^l dataset(i,j,g,l)}{valid_{(i,j)}}\quad b_{(i,j)} = \frac{\sum_1^l dataset(i,j,b,l)}{valid_{(i,j)}}\qquad (5\text{-}2)$$

Where valid$_{(i,j)}$ = number of valid pixels at location *(i, j)* across dataset *l*.

The final pixel class is determined by averaging the probabilities from each prediction image per class over every pixel, then taking the class with the highest mean probability score. Null values from occluded pixels did not contribute to the averages. This method rewards predictions with high confidence and minimises the contribution of pixels with high confusion. It is particularly effective at removing noise around the boundaries of clouds, as the confused pixel is only present in one of the temporal period images. After this temporal ensembling, the image tiles are geographically mosaiced using the maximum class value in overlap areas to prioritise vegetation and minimise clouds. The final model outputs are four images of predicted class probabilities over the entire study area, corresponding to each dry season of the study period (winter 2015/16 – winter 2018/19).

5.2.7 ASM monitoring

From the prediction maps, several datasets were generated to further the geospatial understanding of ASM patterns in Ghana. Firstly, yearly change maps are generated by subtracting the overall prediction maps using a simple integer change mapping schema, detailed in **Appendix C-3**. Secondly, to visualise the mining related changes across the entire study area, a kernel density heatmap was generated showing the density of new ASM pixels ($100m^2$) per square kilometre for each year. Lastly, to study how illegal mining is encroaching on Ghana's forest reserves, the total number of new mining pixels detected inside or adjacent to protected forests over the course of the study was calculated. The geospatial data for the forest reserve boundaries was sourced from the Ghana Open Data Initiative portal (Forestry Commission, 2010).

5.2.8 Model evaluation

Accuracy assessment techniques from both the computer vision and remote sensing communities were used to evaluate the model. Pixel masks commonly

used in computer vision applications were used in the first two assessments. These assessments compared the raw prediction masks output by the model to a manually digitised reference mask. The first assessment used tiles extracted from the same larger 2016 Sentinel-2 granule used for model training, as the SCP classification methodology is not designed to generalise to different satellite images acquired on different days due to the changes in radiometric values (Congedo, 2016). Validation Tile A was used by all classifiers for model tuning, Test Tiles B-D are fully unseen. All classifiers in the first test were trained on only the 2016 training dataset. For the second assessment, three further test tiles from within the Ghanaian case study area but acquired in different years were used, along with two international test tiles chosen to investigate the model's global applicability to other spectrally similar ASM sites. The classifiers used in the second assessment were trained on both the 2016 and 2019 training datasets.

The metrics used for evaluating these assessments are mean class accuracy, mine class accuracy, mean class Intersection Over Union (IoU) and mine class IoU. These metrics are derived from the relationships between True Positives (correctly detected pixels, TP), False Positives (incorrectly detected pixels, FP) and False Negatives (undetected pixels, FN). Mean class accuracy is defined as the mean of the per class accuracy, where the per class accuracy is the number of true positives per class divided by the total number of pixels per class (Equations 5-4 and 5-5).

$$Acc_{mines} = \frac{TP_{mines}}{(TP_{mines} + FN_{mines})} \quad Acc_{built} = \frac{TP_{built}}{(TP_{built} + FN_{built})} \quad Acc_{veg} = \frac{TP_{veg}}{(TP_{veg} + FN_{veg})} \quad (5\text{-}4)$$

$$MeanClassAccuracy = \frac{Acc_{mines} + Acc_{built} + Acc_{veg}}{3} \quad (5\text{-}5)$$

Mean class accuracy, though intuitive, can create misleading results, especially in the presence of many false positives. The IoU score is a robust and commonly used metric for semantic segmentation problems as it provides a statistical accuracy metric which penalises both false positives and false negatives. Each class IoU is calculated as the number of true positives divided by the number of true positives, false positives, and false negatives, shown in Equation 5-6. The mean class IoU is the average of the IoU scores for each class.

$$IoU_{class\_n} = \frac{TP_{class\_n}}{\left(TP_{class\ n} + FN_{class\ n} + FP_{class\ n}\right)} \qquad (5\text{-}6)$$

These pixel level assessments are standard in computer science image segmentation problems and provide a challenging test of the model's abilities. However, they are not directly comparable to the point based stratified random sampling accuracy assessments described in Olofsson et al. (2014) and used in related remote sensing works such as Snapir et al. (2017), Obodai et al. (2019) and Espejo et al. (2018). In order to more closely compare our CNN method to prior work, a third assessment, based on stratified random sampling, was carried out over the entire study area for each year using the post processed prediction maps. The sample points were chosen according to Equation 5-7 (Cochran, 1977) described in Olofsson et al. (2014):

$$\mathbf{n} = \frac{(\sum W_i S_i)^2}{(S(O))^2 + \left(\frac{1}{N}\right)\sum W_i S_i^2} \approx \left(\frac{\sum W_i S_i}{S(O)}\right)^2 \qquad (5\text{-}7)$$

where N = total number of pixels, $S(O)$ is the standard error of desired accuracy estimate, W_i is the proportional area of each class and S_i is the standard deviation of each class, calculated from the user's accuracy U_i as U_i as $S_i = \sqrt{U_i(1 - U_{i)}}$. Using the formula from Equation (5-7) and a desired standard error of 0.01 gives a suggested value of 509 sample points. For point class allocation, the points were divided into 50% for the largest class and 25% each for the smaller classes, giving totals of 254, 127 and 127. These numbers were increased to 300, 150 and 150 to allow for loss due to falling within masked areas and pixels which were unidentifiable to a human from the available Sentinel-2 data. All 600 points were manually compared against the corresponding year's Sentinel-2 RGB images for every temporal epoch of the study. Higher resolution data was consulted when available for the precise temporal period. The metrics used to report the results from this accuracy assessment are overall accuracy, class accuracy, commission errors and omission errors.

5.3 Results

5.3.1 Accuracy assessments

The first assessment compares the results from the initial CNN model against three other machine learning classification methods. The test tiles for this

assessment are radiometrically similar to the training data as they are extracted from the same larger Sentinel-2 granule. The results from this assessment are shown in Table 5-1. The second assessment evaluates the model's generalisation ability against radiometrically unseen test data from both within the study area and from ASM operations in Suriname and Indonesia. The results from the second assessment are shown in Table 5-2 and Figure 5-6. The SAM classifier was not included in the second assessment as the SCP workflow is not designed for processing radiometrically unseen granules.

Table 5-1: Results from the first assessment. This assessment evaluated the results from each classifier on the radiometrically similar test tiles. Full explanation of accuracy metrics is provided in Section 5.2.8.

	Mean Class Accuracy	Mine Class Accuracy	Mean IoU	Mine Class IoU
Spectral Angle Mapping				
Validation Tile A	0.71	0.82	0.65	0.73
Test Tile B	0.67	0.75	0.62	0.68
Test Tile C	0.69	0.73	0.63	0.64
Test Tile D	0.75	0.75	0.66	0.56
Average	***0.72***	***0.79***	***0.65***	***0.68***
Random Forest				
Validation Tile A	0.87	0.90	0.67	0.63
Test Tile B	0.85	0.93	0.81	0.86
Test Tile C	0.85	0.90	0.73	0.73
Test Tile D	0.88	0.85	0.75	0.63
Average	***0.86***	***0.90***	***0.74***	***0.71***
Multi-Layer Perceptron				
Validation Tile A	0.86	0.89	0.67	0.64
Test Tile B	0.84	0.93	0.80	0.87
Test Tile C	0.86	0.90	0.75	0.75
Test Tile D	0.88	0.85	0.76	0.67
Average	***0.86***	***0.89***	***0.75***	***0.73***
Convolutional Neural Network				
Validation Tile A	0.92	0.93	0.88	0.89
Test Tile B	0.93	0.95	0.89	0.90
Test Tile C	0.92	0.95	0.86	0.88
Test Tile D	0.94	0.91	0.83	0.78
Average	***0.93***	***0.94***	***0.87***	***0.86***

Table 5-2: Results from the second assessment. This assessment evaluated the results from each classifier on the radiometrically unseen test tiles. All test tiles are from the different granules than those used for training. Full explanation of accuracy metrics is provided in Section 2.8.

		Mean Class Accuracy	Mine Class Accuracy	Mean IoU	Mine Class IoU
Random Forest					
Eastern	2017/03/28	0.72	0.97	0.49	0.31
Kumasi	2017/01/27	0.82	0.79	0.75	0.73
Obuasi	2019/01/22	0.70	0.70	0.60	0.65
Ghana Average		***0.75***	***0.82***	***0.61***	***0.56***
Surinam	2019/12/01	0.65	0.53	0.52	0.50
Indonesia	2019/09/04	0.65	0.58	0.49	0.53
Multi-Layer Perceptron					
Eastern	2017/03/28	0.78	0.96	0.56	0.36
Kumasi	2017/01/27	0.83	0.80	0.76	0.75
Obuasi	2019/01/22	0.78	0.96	0.58	0.61
Ghana Average		***0.80***	***0.91***	***0.63***	***0.57***
Surinam	2019/12/01	0.67	0.55	0.53	0.51
Indonesia	2019/09/04	0.66	0.43	0.44	0.40
Convolutional Neural Network					
Eastern	2017/03/28	0.92	0.96	0.86	0.83
Kumasi	2017/01/27	0.94	0.97	0.83	0.78
Obuasi	2019/01/22	0.87	0.96	0.79	0.84
Ghana Average		***0.91***	***0.96***	***0.83***	***0.82***
Surinam	2019/12/01	0.73	0.83	0.61	0.71
Indonesia	2019/09/04	0.84	0.82	0.57	0.70
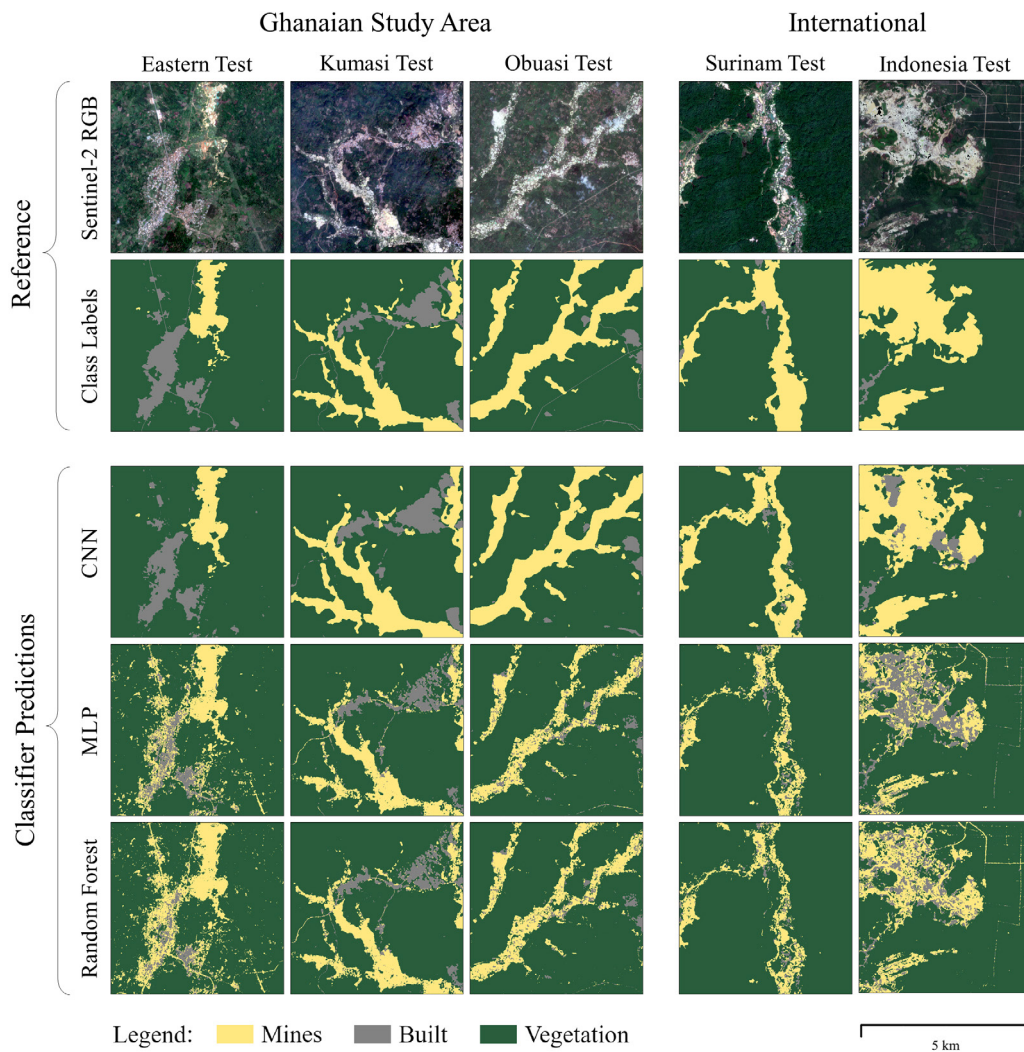
Figure 5-6: Qualitative view of the results from assessment two. Reference information for all five unseen test tiles is displayed alongside the corresponding prediction masks from the various machine learning classifiers.

The first two accuracy assessments (Table 5-1 and Table 5-2) were carried out using the raw prediction results from the CNN model on small individual test tiles, allowing the prediction results to be compared to the labels generated by an operator from a single Sentinel-2 reference image for every pixel. To evaluate the performance over the entire nationwide study area for each temporal period, a third accuracy assessment was carried out, this time based on point based random sampling. This assessment used the post processed class prediction maps, as described in Section 5.2.8. The results assessment three are given in Table 5-3**.** The class accuracies obtained here are higher than the class accuracies from the pixelwise assessments, due to both the probability voting at the ensembling stage and the lower chances of the randomly sampled points coinciding with confusion areas along class boundaries.

175

Table 5-3: Confusion matrices and accuracy metrics for the results from the point based stratified random sampling assessment. The metrics used are class accuracy, overall accuracy, omission error and commission error.

Winter 2015-16	Predicted			Class	Omission	Commission
	Mines	Built	Vegetation	Accuracy	Error	Error
Mines	117	1	0	0.99	1%	6%
Built	2	126	1	0.98	2%	13%
Veg	5	18	299	0.93	7%	0%
				Overall Accuracy		0.95

Winter 2016-17	Predicted			Class	Omission	Commission
	Mines	Built	Vegetation	Accuracy	Error	Error
Mines	107	0	2	0.98	2%	4%
Built	4	119	5	0.93	7%	3%
Veg	0	4	327	0.99	1%	2%
				Overall Accuracy		0.97

Winter 2017-18	Predicted			Class	Omission	Commission
	Mines	Built	Vegetation	Accuracy	Error	Error
Mines	101	2	4	0.94	6%	0%
Built	0	125	3	0.98	2%	5%
Veg	0	4	331	0.99	1%	2%
				Overall Accuracy		0.98

Winter 2018-19	Predicted			Class	Omission	Commission
	Mines	Built	Vegetation	Accuracy	Error	Error
Mines	107	0	0	1.00	0%	2%
Built	1	122	7	0.94	6%	2%
Veg	1	3	329	0.99	1%	2%
				Overall Accuracy		0.98

5.3.2 Applied results

The results obtained in Section 5.3.1 demonstrate that the CNN model is reliably able to locate ASM activity from the Sentinel-2 data to a very high degree of accuracy. Using the full classified area prediction maps and change maps as described in Section 5.2.7, the extents and changes in ASM activities are mapped over time. Figure 5-7 shows the overall land use classes across the whole study area at the start of the study period in winter 2015-16, with insets from the district of Wassa Amenfi East, one of the districts of highest ASM concentrations. The insets show the changes during 2016 alongside the baseline values. The largest scale inset clearly shows the increase in ASM activities both from expansion of existing operations and creation of new ones. Built areas have also increased around the mines, likely due to workforce migration. Figure 5-8 plots the changes in land use categories over the four years of the study. Mining areas increased in 2016 and 2018 but decreased in 2017, built areas increased except in 2016, and vegetation declined steadily each year.
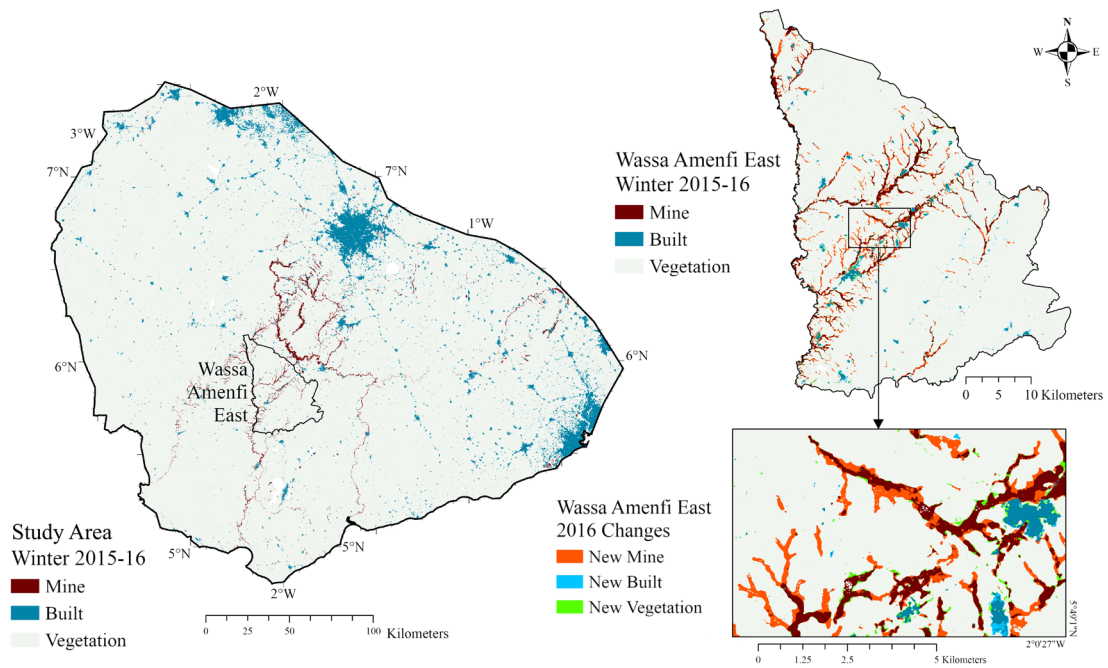
Figure 5-7: Prediction map results displayed for the first year of the study period. The main image shows the full study area at the end of the first temporal period. The upper sub-image shows the district of Wassa Amenfi East with both the winter 2015-16 baseline data and the 2016 changes. The lower sub-image displays a close-up of some of the changes within Wassa Amenfi East detected during 2016.
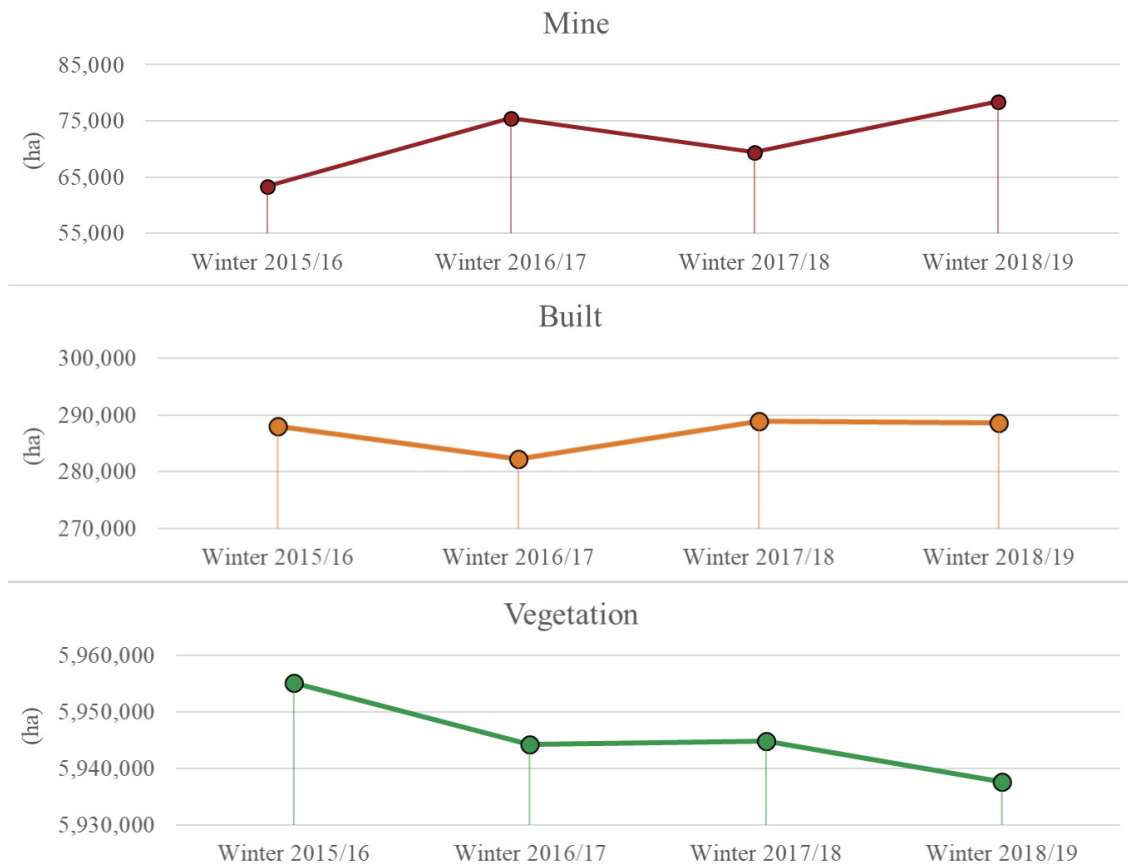
Figure 5-8: Total land cover changes per class observed over the entire study area. Y-axis scale is constant between sub-graphs.

177

To examine the changes in ASM activity in more detail, the total mined area and the yearly mined area changes were computed for ten districts in the Western region. These districts were chosen as they correspond with those studied by Owusu-Nimo et al. (2018), allowing our results to be viewed within the context of their work. The total mined areas and the changes in mined areas are plotted in Figure 5-9, showing a substantial increase in mining in 2016, followed by a decrease across most areas in 2017, followed by a final smaller increase in 2018.
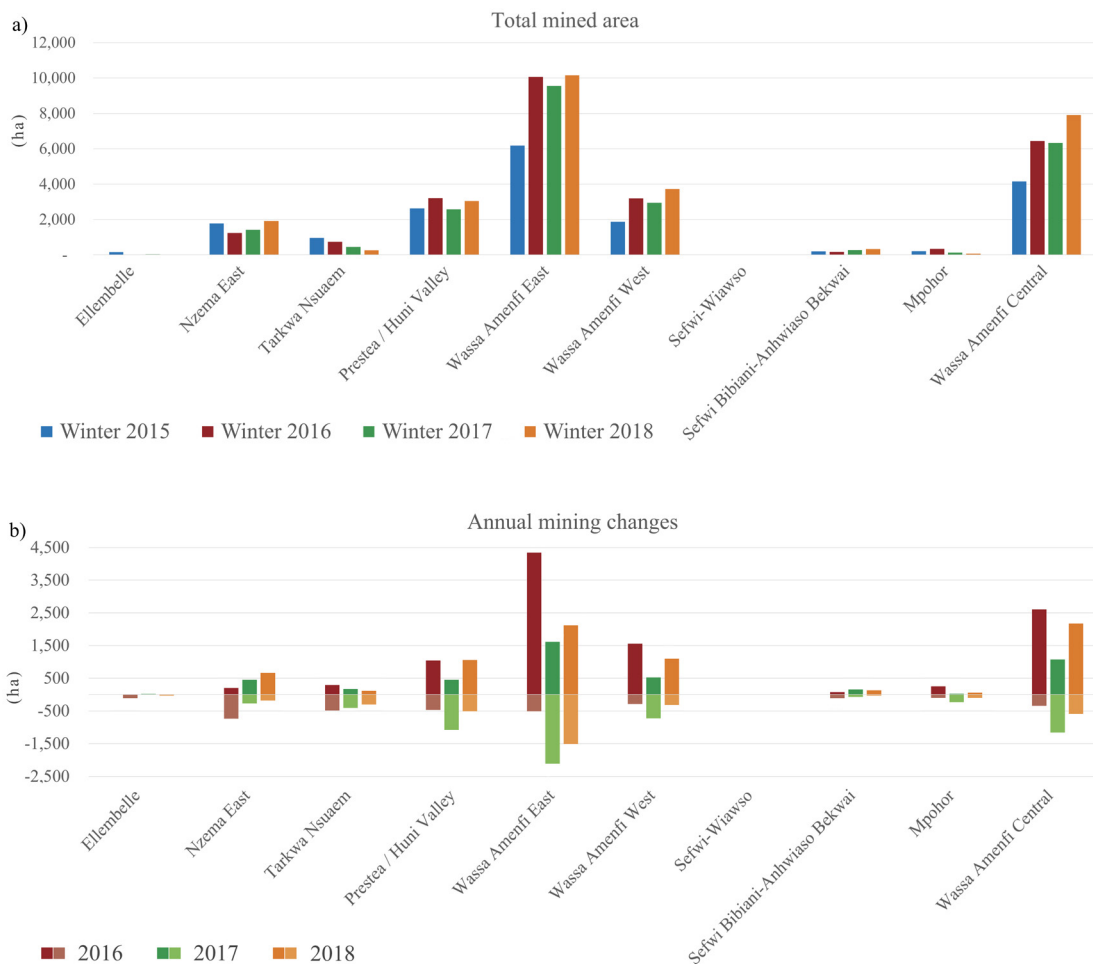
Figure 5-9: Cumulative mined area changes a) and yearly mined area changes b) measured across ten Western districts.

The kernel density heatmap, shown in Figure 5-10, allows observation of the distribution of new ASM pixels each year. It can be seen that in 2016 the greatest focus was in the Western Region, with other hotspots around Accra, due to sand mining and stone quarrying. In 2017 there is an overall decrease, alongside a general migration north-eastward into the Ashanti Region from the Western Region. There was also an increase in the northern Dormaa districts of

Brong Ahafo. In 2018 the intensity resumes in the Western district though at a lower level than seen in 2016. Figure 5-11 indicates the effect that ASM is having on Ghana's protected forests. This shows the locations of all the forest reserves within the study area. The different colours refer to the hectares of new mining detected within their borders from 2015-2019. The inset shows the yearly mining expansion in the most affected reserve of Upper Wassaw.
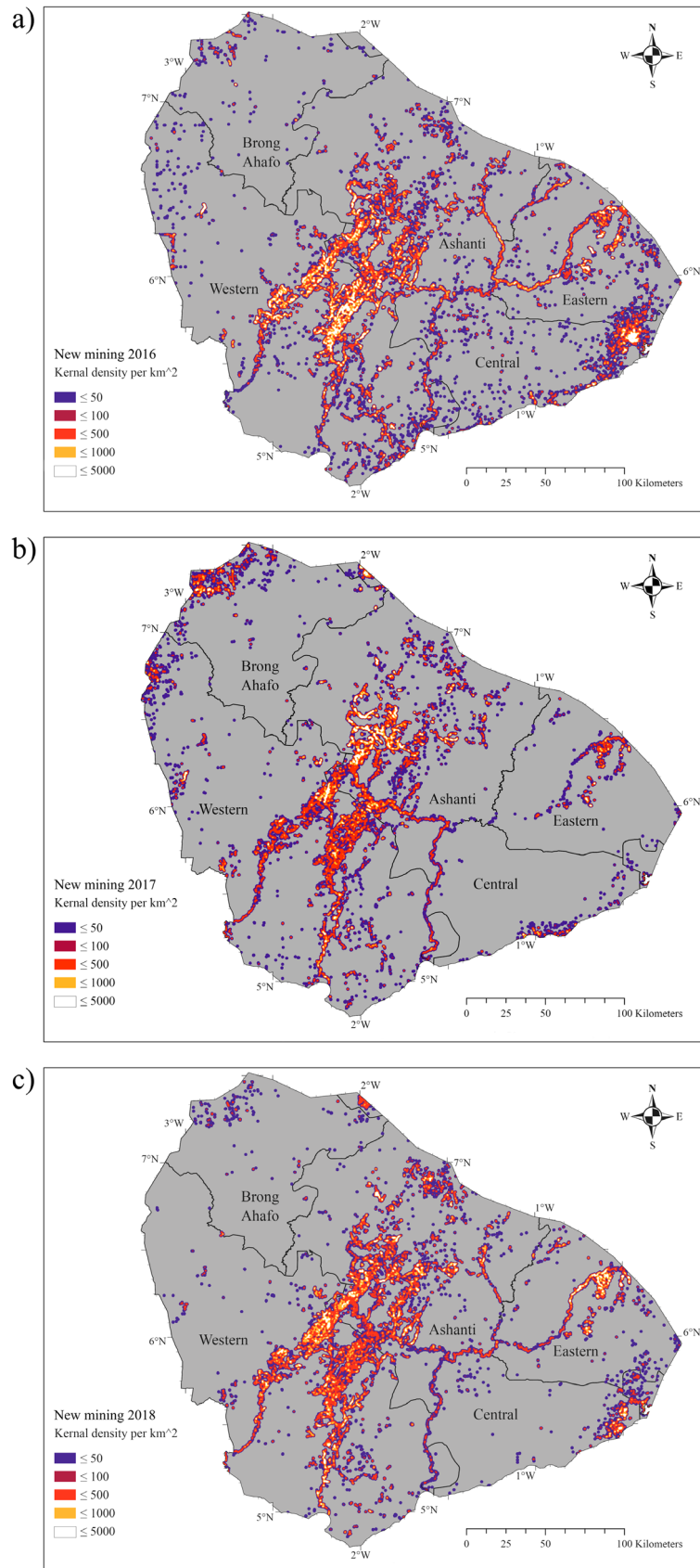
Figure 5-10: Kernel density heatmap showing the concentrations of new ASM activity over the study area over the four years.

180
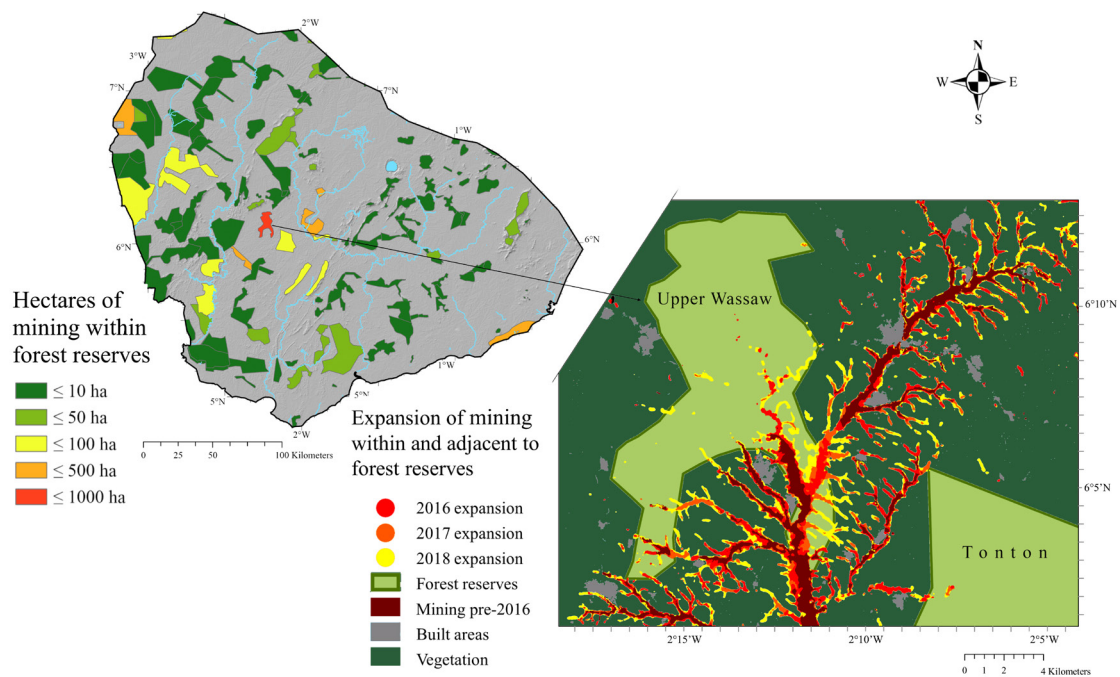
Figure 5-11: Maps showing the hectares of mining within Ghana's national protected forests. The inset shows expansion within the Upper Wassaw reserve. Forest reserve boundaries from Forestry Commission (2010), background GMTED2010 from U.S. Geological Survey.

5.4 Discussion

The results from the series of accuracy assessments demonstrate that using a CNN is a highly effective method for detecting ASM and its related deforestation. The CNN model substantially outperforms the other machine learning classifiers when both false negative and false positives are considered (Table 5-1 and Table 5-2) and does not suffer from the pixel level noise associated with the other methods (Figure 5-6). The CNN is particularly strong at generalisation; the mean IoU score of the CNN changes from 0.87 for the radiometrically similar data to 0.83 for the radiometrically unseen data. In comparison, the mean IoU drops from 0.75 to 0.63 for the MLP and 0.74 to 0.61 for the Random Forest when moving to the unseen granules. It is likely that the greater generalisation ability observed from the CNN is due to its knowledge of pixel context; even under differing radiometric conditions the patterns between the spectral signatures of adjacent pixels can be recognised. The results from the two global test tiles show promise for generalising to alluvial ASM sites globally, despite different vegetation types and mining styles. This indicates that the model could be applied successfully in other countries with the addition of appropriate local training data.

181

The stratified random sampling assessment (Table 5-3) confirms the high accuracy of the classification maps, with omission and commission errors consistently below 8% except for a solitary instance. These errors are lower than others previously reported in the literature. It is suspected that the single high commission error of 13% for the built class in 2015-16 is due to the El Nino related drought in the furthest north region of the study area during that year (Owusu et al., 2019). This resulted in image granules with areas of drought affected fields, which the model was not exposed to during training. These were misclassified as built areas due to their similarity to dirt roads and bare earth around dwellings. If those northern points are removed, the commission error drops to 5%, in line with the other results.

During this assessment it was found that the model predicted some sand mining and quarrying as the mining class, despite not being exposed to these land use classes in training. Only one sampling point from the assessment fell within this confusion area; for clarity, this point was removed from the error matrix table as it is not obvious whether it should be considered a correct or incorrect class for this point. Another anomaly detected from visual inspection of the prediction maps is that polluted rivers with a high sediment load are classified as mines; this is due to having only three possible land use classes in the model. Under this categorisation scheme, polluted rivers most closely resemble alluvial mines and are therefore categorised as such. Whilst the sand mining, drought and the waterway anomalies can be considered errors, they also show that the model is successfully learning from what it has been taught: both land cover types have been assigned to the most similar class based on the extent of the model's prior knowledge. Future studies could incorporate these rarer classes into the training data to avoid such anomalies; however, machine learning algorithms will always be susceptible to new classes that arise after training has been carried out.

The distribution and magnitude of ASM activities observed across the selected Western districts (Figure 5-9) broadly follow those observed during site visits by Owusu-Nimo et al. (2018). The districts of greatest activity are primarily in agreement, with the only substantial difference found in the district of Tarkwa Nsuaem. This is likely due to both the large number of underground illegal mining operations in this district that are not visible to satellite monitoring

systems and also because a large portion of the illegal mining observed by Owusu-Nimo et al. (2018) was within or adjacent to Tarkwa Gold Fields large scale mine which has been masked from this study's results. This agreement between the results from the CNN model and the results from a fieldwork-based study are another indicator that the model is capable of correctly detecting ASM.

The trends observed over the entire study area (Figure 5-8) and the selected Western districts (Figure 5-9) show that whilst mining is increasing overall there was a marked decrease in 2017. This corresponds to the launch of the government's Operation Vanguard (See Section 5.1). This decrease can also be seen in Figure 5-10, where the greatest hotspots from the Tarkwa area in 2016 show significant dispersal by 2017. The Upper Pra and Birim Rivers also show minimal new ASM activity in 2017, though the Dormaa area in Brong Ahafo Region shows an increase, possibly due to migration of miners from areas of higher enforcement elsewhere. From analysis of the intersection of illegal mining activities and Ghana's protected forest reserves (Figure 5-11) it is clear that significant areas of natural rainforest are either at risk from mining or have already been destroyed. In total, over 3,500 hectares of forest preserve had been deforested by mining in the final land cover map from winter 2018-19. Upper Wassaw is the most affected reserve, with 1,300 hectares of mining deforestation; this is more than 11% of its total area.

The applied results from Section 3.2 give a snapshot of the potential applications for which stakeholders could use this model. These datasets also could be used for other applications such as improving the formalisation pathways for small-scale miners, a vital part of improving ASM management as described in Bansah et al. (2018). They also could be used to verify licenses when combined with a layer detailing the locations of licensed small-scale mines. Additionally, the prediction maps could be used to measure progress towards sustainable development targets such as the Reducing Emissions from Deforestation and Degradation target (REDD+). These potential benefits to the formalisation process are amplified by significant savings in terms of capital and workforce resources.

Moving from research to implementation, it is envisioned that the method described here could be re-coded with fully open-source libraries and a user interface to create a software package accessible to GIS professionals worldwide. Generating training and validation patches is straightforward and can be carried out locally by a GIS technician. The workflow is heavily front-loaded in terms of human and computing resources. The initial model training requires deep learning knowledge and a high-powered computer; however, after this stage the processing of large area prediction maps can be accomplished in a single step by a local GIS technician using a standard workstation. A relatively small manually digitised training area of 140,000 hectares can be used to create a model capable of making predictions on a national scale, provided the essential spectral morphologies remain consistent. Most importantly, the resulting model is capable of making generalised predictions on any past or future image from the same sensor, eliminating the need to retrain the classifier for new images, a major time cost for traditional methods. Furthermore, as it uses open source satellite data, there is zero data purchasing cost, making it an attractive alternative to drone-based methods, particularly in its ability to cover very large areas.

5.5 Summary

This study explored the recommendations from Espejo et al. (2018) that higher resolution imagery and artificial intelligence-based methods would be the key to reducing misclassifications and improving the accuracy of automated methods of ASM detection. Utilising powerful deep convolutional neural networks and high-resolution Sentinel-2 data it was possible to robustly learn the spatial and spectral characteristics of alluvial small-scale mining. This research developed the first published multispectral CNN model for this task. It is highly capable of detecting both mined and built areas from Sentinel-2 open source multispectral satellite imagery, alongside clearly distinguishing between the two categories, a task that has proved problematic in the past (Snapir et al., 2017).

The CNN model has been subjected to a series of accuracy assessments to evaluate its abilities. The performance of the model is state of the art, surpassing previously published accuracy figures. The model is able to generalise well, with minimal accuracy loss observed between seen and unseen

radiometric data. The future incorporation of additional training data from varying climates and land use classes could be used to further extend its generalisation abilities. The model is robust to radiometric noise and accurately follows the boundaries of the mined areas, whilst also reducing the issues arising from extensive cloud cover. The processing pipeline developed here rapidly evaluated an area of over 6 million hectares, proving that this methodology can be scaled up to national level for countries to remotely map and monitor small-scale mining.

The datasets produced in this study show how using CNN satellite-based monitoring could provide governments with rapid and detailed knowledge of small-scale mining changes within their jurisdictions. The temporal resolution achievable is dependent on cloud cover, with near real-time weekly updates possible during the dry season. Improved understanding of the spatiotemporal patterns of small-scale mining could be used to track the effectiveness of a range of mitigation strategies employed across different districts. In the Ghanaian study area, these results show a clear link between the establishment of the anti-illegal mining taskforce Operation Vanguard and a decrease in mining activities in 2017. The datasets also enable monitoring of mining-related deforestation within Ghana's protected forests, allowing stakeholders to identify the reserves most at risk and to prioritise conservation. As well as acting as a deterrent for illegal mining, the methodology also has potential benefits in supporting environmentally sound small-scale mining, by identifying sites with good practices and earmarking them for support. Overall, the methodology has been shown to provide accessible, accurate and inexpensive data on ASM regionally, which can be used to create more sustainable mining practices.

Discussion and integration

This chapter brings together the knowledge gained from the three case studies into an overall implementation framework of recommendations for applying neural network based object detection techniques to mining sector remote sensing data, illustrated in Figure 6-1. The framework is made up of three parts: identifying promising applications, developing deep learning models and practical aspects of applying these models to real mining sector problems. After examining the framework, this chapter discusses the scope of this research and situates it within the larger scientific area of remote sensing and artificial intelligence. It then looks to the future, economically and scientifically, concluding with recommendations for further research.

Although each Chapter provides a tangible and effective neural network based model for specific cases, the primary aim of this research is to provide a structure to guide further researchers, both academic and industrial in their application of these emerging algorithms for their own challenges. By choosing to focus on three very different examples, this thesis investigates both the similarities and differences found when applying these approaches to different data structures, scales, tasks and end users.
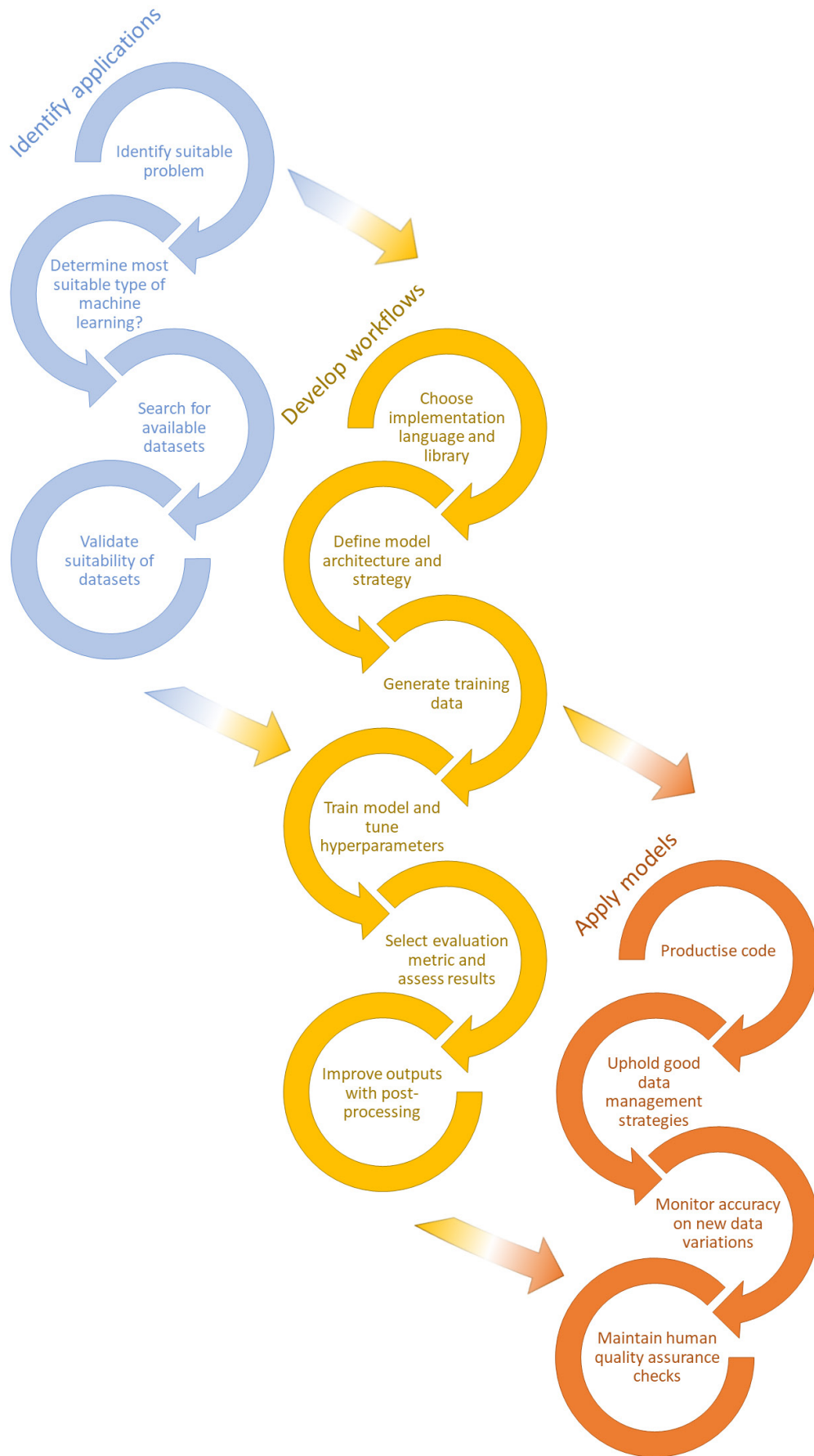
Figure 6-1: Implementation framework key stages.

6.1 Identifying applications

The first section of the framework addresses how to identify specific mining sector challenges for which neural network based object detection may be able to provide effective solutions. These algorithms are a tool, and like all tools they are not the solution to every problem. The early key to successful projects is identifying the types of problems which can most likely benefit from these tools, and conversely, the types of problems which are better suited to other approaches or are not yet solvable with current technologies (Goodfellow et al., 2016). A final consideration at this stage is that this field is moving extremely quickly. In some cases, tasks which the framework indicates cannot be effectively solved today may be achievable in the near future. Therefore, it is advisable to record the conditions which make a task impractical and return to the task if these conditions change.

6.1.1 Machine learning type

As discussed in Chapter 2, neural network based object detectors are only one branch of a large family of machine learning and deep learning techniques. Figure 6-2 proposes a flowchart for guiding algorithm choice, this chart is not intended to be exhaustive or concrete; rather it provides an indicator of how the areas of machine learning and deep learning explored in this research are situated within the wider field. The family of algorithms explored in this research (MLPs and CNNs) are underlined in bold. For simplicity, in this chapter the term 'object detection' is used collectively to refer to the image processing tasks of detection, classification and segmentation; for the full task definitions see Section 2.4.4.
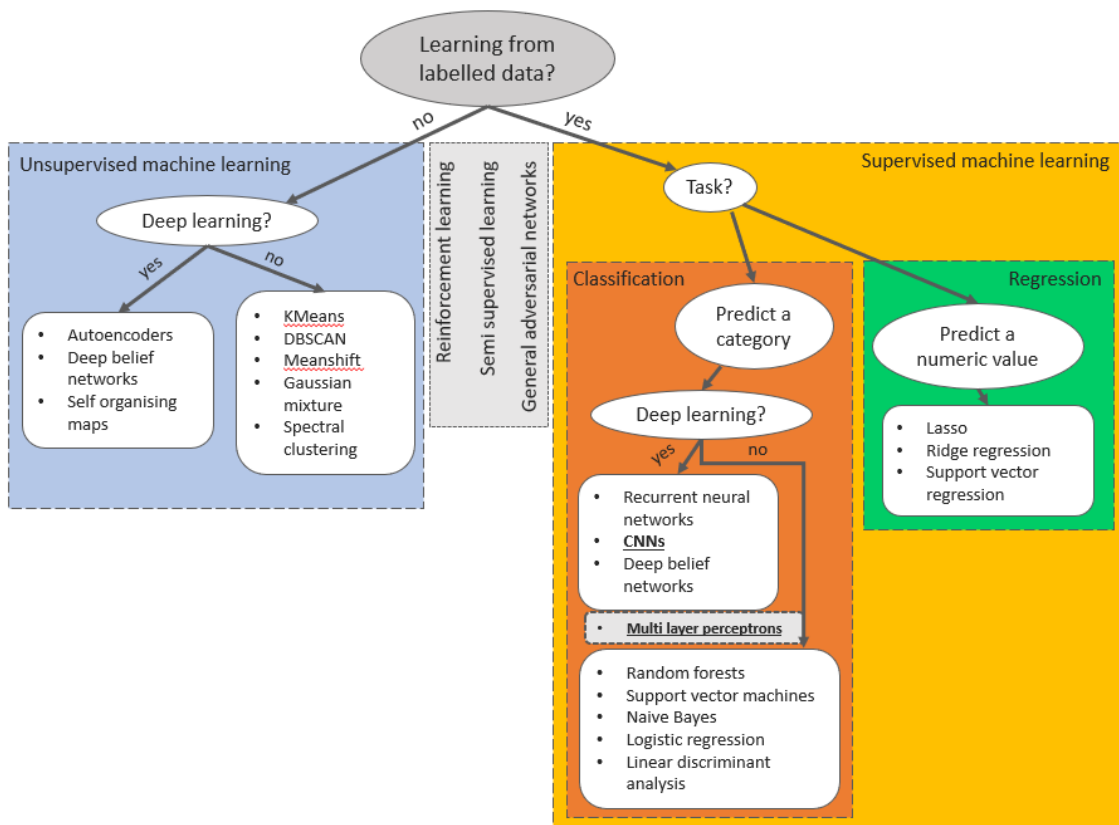
Figure 6-2: Broad overview of machine learning algorithm types. All research carried out in this thesis is situated within the orange 'Classification' box and are shown in underlined bold type. Grey boxes indicate algorithms that are not fully attributable to either category.

Within the broad area of machine learning classification, there is a further consideration as to whether traditional machine learning or deep learning is the most appropriate technique. In general, deep learning has stronger generalisation abilities for more complex problems; however, deep learning algorithms require larger amounts of training data and computing power, making them unsuitable for certain applications such as low power real time learning. Additionally, deep learning algorithms are not yet mature for specialised data types such as 3D point clouds, as discussed in section 2.4.5. Figure 6-3 provides an indicator of how the task characteristics can influence this choice. In this thesis, the applications developed in Chapters 4 and 5 use CNN deep learning algorithms. The application in Chapter 3 uses a neural network with handcrafted features; this architecture can be considered halfway between traditional machine learning and deep learning, as it contains the multiple hidden layers found in deep learning models but does not perform full end to end feature detection and classification. Collectively, the algorithms used in this thesis are henceforth referred to as neural network approaches.

Task characteristics	Traditional machine learning	Deep learning
Supervised classification task?	✓	✓
Requires complex generalisation?	✗	✓
More than 10,000 training examples?	✓	✓
Access to high performance hardware?	✓	✓
High dimensional dataspace?	✗	✓
Less than 1,000 training examples?	✓	✗
Requires human readable model?	✓	✗

Figure 6-3: Task characteristics and their applicability for either traditional machine learning or deep learning.

6.1.2 Task type

Primarily, this research identified that 'needle in a haystack' type object detection tasks are well suited to a neural network approach. To extrapolate the metaphor, for a human to find a needle in a haystack is a proverbially impossible task. However, for a human to find a needle in 10 stalks of hay is not particularly difficult; the challenge of the task lies in the enormity of it. At their current stage of development, many of these algorithms are not superior to humans at a given task when faced with real world imperfect data; their strength lies in their processing speed and that they never lose their attention span. Therefore, tasks which were hitherto impossible due to their scale are particularly promising for neural network solutions.

Additionally, applications will have higher added value if the object to be located is not recorded on any existing mapping. In these situations, creating an initial fully human annotated database would be exorbitantly time intensive, thus negating any possible cost benefit from the further analysis of the data. Such is the case for all three applications developed in this thesis; whilst there is clear demonstrable value in recording the locations of underground rock bolts, historic mining workings and artisanal gold mines, the scale of these tasks and the lack of existing records made them impractical at sitewide, regional or national scales respectively, unless addressed with automated tools.

In general, when determining the suitability of these approaches for any given task, if the problem to be solved involves finding a large number of individually identifiable discrete objects in a very large dataset, then neural networks are expected to be an appropriate tool. For example, in addition to the three applications explored in this thesis, other likely successful applications would include locating roadway potholes, deforestation, traffic cones, coal stockpiles or tailings dams. Identifying tasks which can be readily automated is the first step towards designing effective neural network solutions.

6.1.3 Training data availability

If the task appears to be a good fit for automated detection, the next step is to examine in more detail the availability of the datasets required to train the model. The primary limitation for these types of algorithms is training dataset size. Complex supervised machine learning algorithms need a very large amount of training data in order to learn how to model the correct answers from the input data provided. In the case of deep learning methods such as CNNs, more complex tasks can be solved by deeper architectures, which in turn require larger amounts of training data (Szegedy et al., 2015).

In the case of mining sector problems, often no pre-existing training dataset exists; in these cases, the training data must be labelled manually or adapted from additional sources such as geological maps. The volume of training data required for a successful application depends on the complexity of the model architecture and whether a transfer learning strategy is applied. If a similar problem can be identified from another domain, transfer learning can be used to reduce the amount of new training data needed (Zhuang et al., 2021). However, transfer learning can be unsuccessful if a sufficiently similar previously published algorithm is not found (Zhuang et al., 2021). In this case, the full model must be initialised and trained from random weights, requiring more training data and training time to converge successfully.

If it is not possible to generate the required quantities of training data by human annotation, semi-supervised deep learning methods can be employed. However, these can be more complex to train due to issues such as confirmation bias arising from the semi-supervised pseudo-labelling (Arazo et al., 2020). Other techniques use deep Generative Adversarial Networks (GANs)

to teach an algorithm how to approximate its own labelled examples (Kumar et al., 2017), however, these techniques are still emerging and were shown to not yet outperform transfer learning methods by Majurski et al. (2019). Due to the large training data requirements of deep learning models, traditional machine learning algorithms such as support vector machines, random forests and shallow fully connected neural networks are likely to be a better choice for applications where it is not possible to obtain large numbers of training examples (Liu et al., 2017).

The research carried out in Chapters 4 and 5 allows a comparison to be drawn between the amount of training data required for successful implementation for both random initialisation and transfer learning strategies, due to the similar U-Net architectures used in both chapters. It was found that successful results when creating an entirely new model could be obtained with 16,000 unique training examples, increased with augmentation to 480,000 examples (Section 5.2.2). In comparison, when fine tuning a similar model using transfer learning, successful results were obtained with 520 training examples, increased with augmentation to 2,080 examples (Section 4.4.4).

These results showcase the advantages of transfer learning and indicate how it can reduce the demand on human resources caused by training dataset generation. To aid in identifying potential transfer learning models, it is beneficial to think of the required task not by its specifics but by its data characteristics. This strategy led to the identification of the successful Lunar LiDAR model in Chapter 4. In this case, whilst the specific applications and even scientific fields were very different, the input data structure and the geometric characteristics of the objects to be detected were very similar.

6.1.4 Dataset considerations

The structure of the remote sensing data also plays an important part in identifying the machine learning strategy required and the type of algorithm to use. The vast majority of existing image processing deep learning research has been carried out on standard colour photographs (Tajbakhsh et al., 2020). All data types studied in this research differed from this data type to some degree; however, the most significant difference is in whether or not the data is in a 2D gridded format. Data in this format can be easily read by standard CNN

architectures, which are designed for image processing and are capable of successfully solving complex problems in this domain (Hoeser and Kuenzer, 2020). Some adaptations are needed for the remote sensing data types used in this research, detailed in Chapter 4 for single channel LiDAR and Chapter 5 for multichannel satellite imagery; however, the essential mathematics of the algorithms remain unchanged. Conversely, 3D point clouds are unstructured 3D data and cannot be used with traditional CNN algorithms. Modern deep learning solutions are emerging for this data type (Bello et al., 2020) and are expected to mature in the next few years; however, until this comes to pass handcrafted feature selection methods combined with fully connected neural networks are an effective choice, as described in Xie et al. (2020) and demonstrated in Chapter 3 of this thesis.

Another consideration is the appropriateness of the dataset to the task. Most importantly, the sensor must be capable of acquiring data at a high enough resolution to clearly identify the object of interest and also to differentiate it from the background data. This thesis developed applications for data at 1cm (Chapter 3), 50cm (Chapter 4) and 10m (Chapter 5) spatial resolutions, each appropriate to the particular case studies' application. Chapter 4 also examined 25cm and 1m data, finding that while increasing the resolution produced improved results using the 50cm trained algorithm, decreasing the resolution caused the algorithm's accuracy to degrade considerably. Additionally, the sensor and data type should be appropriate to the type of object to be detected, for example multispectral imagery for vegetation health and LiDAR for topographic variations. The multichannel U-Net architecture developed in Chapter 5 could be used with fused multisource data such as LiDAR and aerial imagery, however, consideration must be given to the balance between increasing model complexity and increasing discrimination capabilities. For example, in Chapter 4, the mining remains to be detected are not significantly visible in colour aerial imagery; therefore, adding three more channels to the input data would add considerably to the model complexity while contributing little to the detection challenge.

Alongside data resolution and applicability of data type, another factor in successful automated detection tasks is data quality. CNN architectures are capable of performing well on noisy datasets (Borodinov et al., 2019; Howarth

et al., 2019); however, transfer learning from a low noise initial training dataset to a noisy fine tuning dataset can considerably degrade accuracy (Rodner et al., 2016). To mitigate against this, noise can be added to the training dataset in a similar way to the pre-training augmentations described in Section 2.4.6.2. Deep learning techniques are also a powerful method for denoising data (Jain and Seung, 2008); for very noisy applications, a multi stage deep learning pipeline could be explored with an initial denoising CNN followed by an object detection CNN.

6.2 Developing workflows

This section of the framework is concerned with how to build effective machine learning models for mining related remote sensing data. The type of model to be built depends on the data and problem type identified by the first stage of the framework, and also depends on whether or not a suitable transfer learning dataset can be found.

6.2.1 Programming language

For building a model, there exists a large choice of language libraries capable of implementing neural network based machine learning algorithms. The choice will depend on the user's own fluency, the type of problem to be solved and whether a suitable transfer learning model exists. If transfer learning is to be applied it is generally recommended to use the same language and libraries that were used to build the original model. If the model is being generated from initialisation, popular libraries would be Python (with Scikit-learn, TensorFlow or Keras libraries) or Matlab with the Deep Learning Toolbox. For a full overview and comparison of current deep learning libraries and languages see Table 2-1 in Section 2.4.6.7.

The conclusions and recommendations from this research in respect to implementation libraries are as follows. For the traditional feature-based machine learning such as random forests and fully connected neural networks used in Chapter 3, Scikit-learn with Python is a robust, easy to use library with well documented examples and a mature codebase. The modular design allows pipelines to be built with interchangeable components, enabling very flexible workflows to be developed. However, Scikit-learn is not designed for deep learning beyond fully connected multi-layer perceptrons.

For modern deep learning, Python with TensorFlow and/or Keras provides some of the most powerful modern deep learning tools for building applications. It is highly customisable, fully open source and is the implementation of choice for the majority of published deep learning research, particularly in the computer science domain (Chollet, 2017). As an alternative to Python, MATLAB's deep learning toolboxes have been expanded rapidly over the last three years and offer a solid platform for research scale deep learning projects. MATLAB's built-in data handling functions simplify training data organisation, augmentation and input, allowing for rapid prototyping.

6.2.2 Model architecture and strategy

For development strategy, if a suitable transfer learning model can be found, it is recommended to use it as a starting point, both to save time spent generating training data and also to begin from an architecture which has been proven to work well on a similar problem. If no transfer learning model can be found, the choice of architecture is determined by the data structure, training dataset size, task complexity and final output required. A common theme regardless of the model architecture used is the need for context to support general inference. Individual pixels or point cloud points are not descriptive enough to predict an object or category without information from their neighbouring pixels or points. This context is added either though the convolutional filter in the 2D CNNs in Chapters 4 and 5 or by using a point neighbourhood as shown in Chapter 3. Further information on contextual elements of model architectures is found in Sections 2.4.3 and 2.4.5.

Reviewing the literature for tasks with similar characteristics can provide a good starting point for selecting an architecture. For image based segmentation tasks with limited training data, U-Nets have been proven to be a reliable choice both in Chapters 4 and 5 of this research and in other literature published during the research period, for applications such as detection of coal stockpiles (Chowdhary et al., 2019), building footprint extraction (Schuegraf and Bittner, 2019) and retinal vessel segmentation (Wang et al., 2019). Architectures will continue to evolve, an adapted U-Net (U-Net++) provided excellent results in Peng et al. (2019) for satellite based change detection. In a current deep learning context U-Nets are a relatively old architecture, many newer

architectures now outperform them on the computer science benchmark datasets. However, they are still a dependable choice when the training dataset is small. More experimentation is needed to determine how well the very deep architectures used extensively in the computer science community perform when used with far smaller training datasets, with an excellent review of how computer science architectures are applied in earth observation given in Hoeser and Kuenzer (2020).

6.2.3 Training data creation

Once a strategy, software library and basic architecture have been chosen, the training data must be generated. The research carried out in this thesis showed that in all 3 cases good results could be obtained with less than 40 hours of human labelling time; however, the applications developed here were either single class or small multiclass problems with reasonably heterogeneous datasets. More complex or varied problems would likely require larger amounts of training data (Burkov, 2019). To determine if the training dataset size is adequate for the problem to be solved, it is necessary to first examine whether the test dataset performance is much lower than the training dataset performance. If this is the case, it can be inferred that the model has not fully learned the parameters of the problem from the provided training data (Goodfellow et al., 2016). This indicates that either the complexity of the model has to be decreased, or the training dataset size has to be increased. The choice depends on the cost and feasibility of gathering more data (Goodfellow et al., 2016). Additionally, if the training data instances are not varied enough the model will not accurately represent the problem and will fail to generalise to other slightly different instances (Géron, 2017). However, if the training data is highly varied without a sufficient number of examples for each object, the model will struggle to identify the essential characteristics of the object which it has been tasked to find and may not perform well. The training dataset should therefore encompass the expected variability across the scene (Géron, 2017).

Section 3.5 discussed the significance of labelling errors in the training data in the context of the Cornish bolt detection dataset. At the scale required for these applications, incorrectly labelled training data below 10% had no appreciable effect on the accuracy of the model (Folleco et al., 2008; Pelletier et al., 2017).

196

In contrast, imbalanced classes did greatly affect the accuracy of the final model and must be addressed early in the workflow development stage. Multiple actions can be taken to mitigate against this. The easiest is to choose a training area where the rare class to be detected is abundant. Next, the training samples need to be rebalanced to an appropriate level for the detection task. As discussed in Section 3.3.2, the optimal class rebalancing strategy may not be 50:50. Determining the best class ratio is generally achieved empirically by monitoring the results on the cross validation dataset as the class ratio is varied.

Augmentation is used to increase the size of image based training datasets. This consists of randomly rotating, flipping and scaling the input image to decrease overfitting tendencies, as described in Section 2.4.6.2. Augmentation was used successfully in Chapters 4 and 5; however, the point cloud data from Chapter 3 was not augmented. This was because it is more complex to generate valid 3D augmented data due to the importance of real world position and orientation. In this case augmentation was not needed as the dataset was sufficiently large for the scope of the problem.

6.2.4 Training

Several techniques were implemented in this research which allowed the models to be trained successfully on standard high powered workstations without the need for supercomputing clusters. PC specifications and model characteristics are given in Table 6-1. The substantial processing speed improvement enabled through transfer learning can be clearly seen in the difference between training times for the two U-net models.

Table 6-1: Computing specifications and selected model characteristics.

Algorithm	Chapter	Transfer Learning	RAM	GPU	Training time (approx.)
Multi-layer perceptron	3	No	32GB	NVIDIA GTX960	100[16] minutes
Single-channel U-Net	4	Yes	32GB	NVIDIA GTX960	12 minutes
Multi-channel U-Net	5	No	64GB	NVIDIA Titan X	960 minutes

---

[16] Total algorithm run time. 90 minutes for feature creation, 10 minutes for training the MLP.

The models developed in Chapters 3 and 4 were trained on the candidate's office PC. The model developed in Chapter 5 required more processing power as it was a full deep learning model trained from initialisation. This model was trained on the department's high powered modelling workstation. The Chapter 5 model would have trained on the smaller PC; however, training time would have increased from ~16hrs to ~100 hours. Mini batching with shuffling between epochs was used for both deep learning models (Sections 4.4.4 & 5.2.4). The input image size and the convolution filter size were also empirically set to values which would fit in memory.

In this research, hyperparameters were set empirically, by taking literature values for the model in question as a starting point and then changing them one by one whilst observing the changes to the results on the cross validation dataset. Alternative methods for hyperparameter choices are outlined in Section 2.4.6.4. While observing the cross validation dataset, it was found that monitoring only the numeric results from the cross validation dataset did not give a full picture of how the model was training and the validity of the hyperparameter choices. Human qualitative interpretation of the results on the cross validation data was a valuable tool in understanding some of the black box type behaviours of the models. For this, several cross validation image tiles which were deemed to be 'difficult' were manually examined after each time the model was trained. This gives an insight into how hyperparameter choice affects the final results and what types of mistakes can be attributed to which incorrect setting.

6.2.5 Assessing results

The accuracy metric chosen also has an impact on how the model learns, how it is adjusted and how the results are understood. Different metrics are appropriate for different applications, though most are based on some combination of the numbers of true positives, false positives and false negatives. Overall accuracy is a poor metric unless the classes are well balanced (Chollet, 2017). Chapters 3 and 4 used the metrics precision, recall, and F1, while Chapter 5 used per class accuracy, intersection over union and omission-commission errors. Depending on the final application, different

metrics can be prioritised during hyperparameter tuning. Generally, there is a precision-recall trade-off, where improving one decreases the other.

The same model architecture can be trained in different ways for different applications, as implemented in Chapter 4. The primary application to detect historic mining pits in Section 4.4 required a balanced precision and recall model, while for the geological mineralisation fitting application from Section 4.5 it was preferable to maximise precision over recall, as for that application, noise was more detrimental than missed holes. The different LiDAR data representations gave different results with the precision and recall values and ratios changing depending on the visualisation type, as shown in Figure 4-21 and Figure 4-23. Based on these results, the model trained on the raw DSM representation was used for the historic pit detection application as it had the most balanced precision and recall, while the model trained on the positive openness representation was used for the mineralisation line fitting application.

Another training data strategy for improving results is to use a confusion dataset, as described in Section 3.5. This ensures the model can focus on its mistakes before encountering the final unseen test data. A confusion dataset can be generated by carrying out multi-fold cross-validation and saving only the misclassified points from each fold to a separate confusion dataset. This dataset can then be used as an additional training data file with either a higher weight or lower resampling, depending on the application's training strategy.

6.2.6 Power of consensus

The last item to consider when developing workflows is whether to add an ensembling or clustering post processing stage. Ensembling improves the results and acts as a noise canceller, minimising the effects of random incorrect predictions. As described in Section 4.4, machine learning methods are good at predicting broad trends but can have errors at an individual instance level. Ensembling multiple predictions therefore makes the final output more robust. In this thesis, Chapter 4 used mean value ensembling and Chapter 5 used a novel ensembling technique based on prediction confidence developed in this research and described in Section 5.2.6.

Clustering works on the same principle as ensembling in that it uses the power of majority to remove noise and improve the confidence of predictions. Clustering is powerful for 3D point cloud data types as the locations of the predicted points in 3D space can be clustered using their real valued distances. In Chapter 3, DBSCAN clustering was used to separate candidate bolt points from general noise, allowing the data to be fed to the algorithm in its uncleaned state, with all confusion objects often present in an underground mine left in the data. This greatly increases the real world applicability of this algorithm, as it can work directly on the raw point cloud in an automated manner, without requiring human data cleansing preprocessing.

6.3 Practical model application

This final part of the framework is concerned with applying the models identified and developed in the previous sections. In the field of deep learning this stage is known as inference, which is where the trained model makes predictions on entirely new data in a real world environment. This section describes some universal considerations and good practice guidelines that have been observed during this research, regardless of the application end goal and the model architecture.

6.3.1 Data management

Both in training and inference these models rely on large datasets; therefore, appropriate data management strategies must be in place at all stages of a project (Munappy et al., 2019). For example, the Ghanaian application ingested 168 GB of Sentinel-2 data and produced 63 GB of prediction results; if extended to yearly monitoring this figure would grow by 25% each year, creating a challenge for organisation and storage. Another aspect of good data management is maintaining the separation of the training, validation and testing datasets, ensuring the model is not biased by viewing the final test dataset during development (Chollet, 2017). As the test dataset can only be used once, it is recommended to separate out several potential test areas after labelling and reserve the unused ones for future research.

6.3.2 Divergence between training data and inference data

Another consideration is the range and variability of the training data in relation to the inference data. For example, the model trained for ASM detection in Chapter 5 was only trained using examples of alluvial ASM in a tropical area. At the inference stage, the model made incorrect predictions in the far northeast of the study area where the landscape changed from tropical rainforest to savannah. To address these types of issues, the model could either be retrained using a large dataset of both tropical and savannah examples, or a new arid landscape model could be trained and the application areas split depending on their climate type with a different model applied to each one.

6.3.3 Code packaging

All models and applications developed in this PhD were written in prototyping code which has not yet been optimised for professional or commercial use. With code optimisation, the training and inference times would reduce, and the ease of use would be improved. With further code optimisation and GUI development, the inference stage of all 3 applications could be deployed for use by non-programmers for each respective real world application.

6.3.4 Human oversight

The final consideration when applying these models to real world mining data is that the models are not infallible and should not be followed blindly. After deployment, continuing random spot checks are recommended, alongside more thorough monitoring whenever conditions in the inference dataset change, for example if a new satellite sensor comes online or a new study region is added to the area. In many cases, the generalisation ability of these models is adequate to adapt to new (albeit similar) sensors or study regions; however, the results should be subject to greater scrutiny when conditions change, to determine whether the change remains within the capabilities of the existing model. It is also recommended to query and be mindful of the potential biases in the model, as any biases present in the training data will persist within the final model (Alvi et al., 2019). Retraining or fine tuning as new data and knowledge becomes available is recommended to maintain the integrity of the model over longer timescales.

6.4 Framework summary

The first section of the framework identified the conditions which lead to successful machine learning remote sensing projects. Prior to commencing model development, the precise task must be defined, the existence of training datasets must be determined, and the data structure and quality must be ascertained. The second section describes developing the models, including choosing an implementation library, training the model and deciding an evaluation protocol. This section also describes some techniques for improving the end prediction accuracy, valuable for real world industrial projects. The final section details some practitioner guidelines for applying these models to real environments post development.

6.5 Scope and future recommendations

The framework described in Sections 6.1 - 6.4 demonstrates how a thorough approach to identifying, developing and applying deep learning and remote sensing technologies ensures optimal use in mining contexts. This PhD showcased multiple areas of the mining sector where deep learning and remote sensing can be successfully applied. The three publications based on the case studies of this research deliver new knowledge and techniques in and of themselves and also can be used to establish how similar data types and mining problems can benefit from this approach in future. The triple application structure was chosen to best demonstrate the breadth of potential for neural networks to solve mining sector tasks. The types of machine learning tasks addressed in this work are all classic computer vision problems such as detection, location and segmentation. Other more specialised tasks such as time series analysis and unsupervised deep learning were deemed outside the scope of this research.

As discussed in Section 2.4, deep learning algorithm development is proceeding at an unprecedented speed. The framework proposed in this thesis for solving mining sector object detection challenges is capable of adapting to future algorithm developments, as it is a general framework and does not depend on any one specific algorithm or technique. This thesis chose algorithms which have reached a proven level of maturity and stability over the last 5 years. This has enabled the research to focus on the adaptation of these

technologies to new mining sector challenges, rather than troubleshooting newly released code which might be updated or reissued regularly. As other algorithms mature, the model architectures used in this thesis can be replaced when newer models stabilise, whilst maintaining the foundational framework and adaptations learned from this research.

Looking to the future, several areas of interest for further research were identified during the course of this project. These range from harnessing technological improvements to further investigation of phenomena observed in this research, as discussed in the following sections.

6.5.1 Deep learning for point clouds

At the beginning of this PhD research in 2017, deep learning for unstructured 3D point cloud data was in its infancy, with the seminal PointNet paper published in July of 2017. Over the last 3 years this area has advanced rapidly (Bello et al., 2020) as discussed in Section 2.4.5 of the literature review. These algorithms are now approaching the preferred level of stability to be implemented in real world applications. Future research for underground applications such as the bolt detector developed in Chapter 3 will undoubtedly leverage these new technologies.

6.5.2 Multiple LiDAR data representations

When using LiDAR data for model input in Chapter 4, different representations such as positive and negative openness and simplified local relief models gave different results, indicating that this form of 'image enhancement' had an impact on how the model made predictions. As this chapter used transfer learning, it was not possible to separate whether this result was due to inherent differences in how the models learned or whether it was due to similarities or differences between the data type of the originally trained dataset and the new fine tuning dataset. To fully examine this phenomenon, it would be first necessary to create a training dataset large enough to train a model from initialisation. Without the bias introduced by the LiDAR representation of the transfer learning dataset, it would be possible to determine which representations were more effective on different dataset types by training and testing multiple models from initialisation. An interesting area for future research would involve determining which LiDAR

data representation out of the many described in Section 4.3 would provide the best results when used with different CNN architectures.

6.5.3 Multi-biome ASM models

As previously discussed, the ASM detection model built in Chapter 5 was trained only on alluvial ASM from tropical areas, as that is the primary type of ASM found in the study area. The model developed in this research could be expanded to multiple ASM and landscape types, allowing it to be applied globally for effective low cost ASM monitoring. Further research is needed to determine whether multiple models combined with geographic area categorisation or one large model trained on many landscapes and ASM types would be the best solution.

6.5.4 Hyperspectral models for mining pollution

While carrying out the ASM project it was hypothesised that vegetation indices such as Red Edge and Normalised Difference Vegetation Index (NDVI) could be generated from the satellite data and used to further identify ASM related pollution in not yet deforested areas. NDVIs have been used in previous non deep learning based land cover studies in the region (Abdoulatif et al., 2019; Barenblitt et al., 2020; Basommi et al., 2015), primarily as a land cover classification input for already deforested areas. After some initial experimentation it was found that vegetation changes in close proximity to deforested pools were indistinguishable to a human analyst.

Hyperspectral sensors with a greater spectral resolution could potentially be capable of detecting these changes; however, it was not possible to acquire a high resolution hyperspectral dataset over the area of interest. Future research could investigate if hyperspectral data can detect ASM at an earlier stage or detect the downstream effects of it. If successful, this could be combined with CNN models similar to those developed in this research to detect both pollution and ASM deforestation.

6.6 The future of deep learning for mining

Advances in the broader fields of AI and automation are continuing at an extremely rapid rate both within the mining industry and the larger tech industry. Currently, AI methods such as deep learning can be used for specific tasks;

however, the future goal is artificial general intelligence and real time decision making. Equipment manufacturers are edging closer to the fully autonomous mine, with advances such as Sandvik's self-routing loader (Wired Magazine, 2020) and Orica and Epiroc's first steps towards an autonomous explosives delivery system

(Valler and Andersson, 2020). The solution to automated real time intelligence in a dynamic environment will likely arise from the self-driving cars of Google and Uber. Adaption of these technologies to mining will be easier, as mining environments are more controlled with less scope for random interactions between humans and machines. Deep learning could be used by machines to make complex semantically labelled models of their environments, allowing higher level autonomous onboard risk assessments and decision making. As labelled digital mining datasets grow, more complex deep learning models can be trained to in turn solve more complex tasks.

Moving out from a mine site scale, ever greater volumes of satellite remote sensing data can be used to monitor the environmental impact of mining even in difficult to reach locations. A recent study counted millions of individual trees in the Sahara (Brandt et al., 2020), leading to a Nature op-ed proposing that every tree on earth could be mapped by satellites in the not too distant future (Hanan and Anchang, 2020). Technologies such as this would allow for rapid mapping of mining related deforestation at a hitherto unprecedented scale. Technologies such as this can influence policy, as they enable countries' commitments to programmes such as the United Nations' 'Reducing emissions from deforestation and forest degradation' (REDD+) schemes to be accurately quantified.

The examples above are only some of the vast range of mining sector applications which may benefit from deep learning in future. Potentially the most change inducing new technologies have not even been envisaged yet. What is certain is that the industry needs to keep pace with technological advances and maintain a critical view of both the possibilities and limitations. This can be achieved by encouraging collaboration between big tech and the mining sector, by attracting more students from a computer science background into mining

programs and by further research into how these technologies can be adapted and applied for maximum benefit within the mining sector.

Conclusion

The purpose of this thesis was to demonstrate that deep learning and remote sensing tools can be successfully used to solve real world mining sector problems. The research developed applications for the three most common remote sensing data structures, showing that these technologies can be applied across a wide range of data types and scales. Each of the individual applications contribute scientific knowledge to their area, from intelligent underground machines to governmental mining policies.

Each of the case studies used a different type of remote sensing data, demonstrating how approaches can be designed to leverage a wide range of data structures and scales. The first case study used a lightweight feature based fully connected network to accommodate the noise and density of 3D point clouds. In contrast, the second and third case studies used a grid based convolutional neural network architecture as these are better suited to multispectral 2D and 2.5D datasets. In particular, the second case study demonstrated that differing LiDAR data representations can be used to enhance the interpretability of these types of datasets. By adapting the model design to the data type, task complexity and data availability, this research demonstrates that neural network based approaches can detect objects at a level of accuracy comparable to or surpassing other published approaches, with the caveat that direct accuracy comparisons are not possible due to each study being evaluated against its own bespoke datasets and tasks.

In the three case studies of this research, methodological improvements were developed that increased the accuracy on their target datasets. For the bolt detection algorithm in the first case study, accuracy was improved by combining feature generation techniques from both the mobile robotics and remote sensing research communities, demonstrating how cross disciplinary knowledge transfer can aid in real world technology application. For the LiDAR applications in the second case study, transfer learning from a similar domain was proven to be a powerful method for applying deep learning techniques to smaller datasets, depending upon the availability of a suitably similar transfer learning model. In the third case study, the accuracy improvements came from a novel confidence based ensembling, which additionally allowed for greater

coverage in cloudy regions. This case study also demonstrated that building and training deep learning models without transfer learning is a viable strategy for specific tasks and data types, provided the training data can be generated.

Each case study also included a practical post detection example of how the results from these algorithms can be used to add knowledge and value to remote sensing datasets. The bolt detection algorithm in the first case study was used to feed into a bolt location database, allowing a mine to record the as-built bolting pattern and to generate bolting quality assurance reports, supporting intelligent machines and mining information modelling. The legacy mining pit detection algorithm in the second case study was extended to infer mineralisation and lode direction, allowing historians and stakeholders to gain further knowledge of a landscape's history even in cases where records have been lost. The third case study, which examined small scale mining in Ghana provided an accurate national scale dataset that will allow governments and NGOs to gain valuable insights into the patterns of mining, migration and urbanisation.

Taken together, the three case studies of this thesis establish a flexible and effective approach to utilising these technologies in the mining sector. The final chapter of this thesis brings together the implementation knowledge gained into a framework that will aid in the application of these techniques to other tasks. Through comparing and contrasting the approaches selected in the three case studies, the framework sets out effective methods of identifying, developing and applying these technologies within the mining sector across a range of tasks, data structures and end users.

Appendices

Appendix A – Chapter 3

A-1 Datasets

3D point cloud data from two mines, collected by Matthew Eyre and Jane Gallwey – not publicly available.

A-2 Software

Leica Cyclone V.9.3.1

Python libraries: Scikit-Learn, OpenCV, Pandas, Numpy

CloudCompare v2.9.1

Appendix B – Chapter 4

Datasets, software and workflow from paper 'Bringing Lunar LiDAR Back Down to Earth: Mapping Our Industrial Heritage Through Deep Transfer Learning'

B-1 Datasets

Lidar Composite Digital Surface Model England 50cm resolution [ASC geospatial data], Scale 1:2000, Tiles: sx6780, sx6781, sx6782, sx6880, sx6680, sx6681, sx6070, sx6071, sx6170, sx6171, sx6270, sx6570. Updated: 5 January 2016, Open Government Licence, Using: EDINA LiDAR Digimap Service, <https://digimap.edina.ac.uk>, Downloaded: 2019-03-07 16:43:54.775

Lidar Composite Digital Surface Model England 25cm resolution [ASC geospatial data], Scale 1:1000, Tiles: se0166 Updated: 5 January 2016, Open Government Licence, Using: EDINA LiDAR Digimap Service, <https://digimap.edina.ac.uk>, Downloaded: 2019-06-10 13:37:54.891

High Resolution (25cm) Vertical Aerial Imagery (2011, 2015) Scale 1:500, Tiles: sx6780, sx6781, sx6782, sx6680, sx6681, sx6070, sx6071, sx6170, sx6171, sx6270, sx6570, se0166, Updated: 25 October 2015, Getmapping, Using: EDINA Aerial Digimap Service, <https://digimap.edina.ac.uk>, Downloaded: 2019-03-09 11:26:46.049

1:2500 County Series 1st Edition [TIFF geospatial data], Scale 1:2500, Tiles: devo-sx6780-1, devo-sx6781-1, devo-sx6782-1, devo-sx680-1, devo-sx6681-1, devo-sx6070-1, devo-sx6071-1, devo-sx6170-1, devo-sx6171-1, devo-sx6270-1, devo-sx6570-1. Updated: 30 November 2010, Historic, Using: EDINA Historic Digimap Service, <https://digimap.edina.ac.uk>, Downloaded: 2019-03-07 15:03:20.971

B-2 Software

ArcGIS Pro 2.3.1

Relief Visualisation Toolbox 1.3

Python libraries: Scikit Image, OpenCV, Pandas, Numpy, H5py, Matplotlib, Keras, Tensorflow

B-3 Workflow

1) Import layers to ArcGIS and create training data from desktop survey.

2) Quality control the DEM using an elevation void fill operation

3) Export training data for deep learning using appropriate object radius and tile overlap. For generating training data do not tick 'export no feature tiles', for test data this box should be checked. Use Pascal VOC as the type of label to be exported. This step creates .tiff images, .tfw geotiff world files (georeferencing information) and .xml label files for every image tile.

4) Remove the default output folder structure and put all files in one folder.

5) If required, the image tiles can be batch fed into RVT to create other visualisations which will match the existing georeferencing and label files.

6) A modified version of the 'input_data_gen.py' script found at https://github.com/silburt/DeepMoon is then run to generate hdf5 database files from the labelled images. The modifications are minor and will depend on the machine and data considerations for a particular user. This script outputs two database files for each of the training, validation and test datasets; one for the images and one for the pits.

7) The 'run_model_train.py' script also from https://github.com/silburt/DeepMoon is then run, with the parameters set appropriately for the dataset and computer used. This script calls the 'model_train.py' script which also needs to be modified to load the model provided from https://zenodo.org/record/1133969#.XVUZi-j0k2w in either Keras 1 or Keras 2 format depending on machine setup. This allows the model to be fine tuned from the endpoint of the Lunar study rather than trained from scratch.

8) To test the model a modified version of 'Using Zenodo Data.ipynb' found in https://github.com/silburt/DeepMoon/tree/master/docs is run. This jupyter notebook can be modified to remove the sections on crater drawing and loop through whichever directory the test image database is found in. Python image libraries can be used to save the generated image masks.

9) These image masks have the same extents as the original image tiles output from ArcGIS, therefore if the .tfw files from the original export are copied into the folder with the masks the masks will import into ArcGIS already georeferenced correctly.

10) Merge the tiles into a single raster. This is the end of the workflow if using the pixel based masks. To aid interpretation it is recommended to use a graduated colour scheme where 0 probability is mapped to fully transparent pixels.

11) To generate vector data first remove all pixels below a set threshold, for example 0.4.

12) Use raster calculator to convert this layer to integer form =(int(raster * 100)).

13) Use tool 'convert raster to polygon'.

14) Tidy up generated polygons by merging edges that touch and splitting polygons that don't touch into separate objects.

Appendix C – Chapter 5

C-1 Dataset

Tile Code[17]	Date	Tile Code	Date	Tile Code	Date	Tile Code	Date
Year 1		**Year 2**		**Year 3**		**Year 4**	
NYM	24-Nov-15	NWM_W	11-Dec-16	NYM	13-Dec-17	NWN_E	23-Dec-18
NWN_E	04-Dec-15	NVN	11-Dec-16	NYM	02-Jan-18	NWM_W	26-Dec-18
NWP_E	04-Dec-15	NWN_W	11-Dec-16	NXL	02-Jan-18	NVM	26-Dec-18
NWN_W	07-Dec-15	NWL	28-Dec-16	NWN_E	02-Jan-18	NVN	26-Dec-18
NWP_W	07-Dec-15	NYN	28-Dec-16	NWM_W	05-Jan-18	NWM_E	28-Dec-18
NVN	17-Dec-15	NXN	28-Dec-16	NWN_W	05-Jan-18	NXM	28-Dec-18
NWN_W	17-Dec-15	NWN_E	28-Dec-16	NWM_E	12-Jan-18	NYM	28-Dec-18
NWP_W	17-Dec-15	NWP_E	28-Dec-16	NYM	12-Jan-18	NWL	28-Dec-18
NWM_E	24-Dec-15	NWM_W	31-Dec-16	NYN	12-Jan-18	NXL	28-Dec-18
NXM	24-Dec-15	NVM	31-Dec-16	NWM_W	15-Jan-18	NWN_W	31-Dec-18
NYM	24-Dec-15	NWP_W	31-Dec-16	NVM	15-Jan-18	NWP_W	31-Dec-18
NZM	24-Dec-15	NXP	07-Jan-17	NWN_W	15-Jan-18	NYN	02-Jan-19
NYL	24-Dec-15	NWP_E	07-Jan-17	NWL	17-Jan-18	NXN	02-Jan-19
NWL	24-Dec-15	NWM_E	27-Jan-17	NXL	17-Jan-18	NWN_E	02-Jan-19
NXL	24-Dec-15	NXM	27-Jan-17	NWM_E	22-Jan-18	NWP_E	02-Jan-19
NYN	24-Dec-15	NYM	27-Jan-17	NXM	22-Jan-18	NWM_W	05-Jan-19
NXN	24-Dec-15	NWL	27-Jan-17	NWL	22-Jan-18	NVN	05-Jan-19
NWN_E	24-Dec-15	NXL	27-Jan-17	NXL	22-Jan-18	NWN_W	05-Jan-19
NWP_E	24-Dec-15	NYN	27-Jan-17	NYN	22-Jan-18	NWP_W	05-Jan-19
NWM_W	27-Dec-15	NXN	27-Jan-17	NXN	22-Jan-18	NWM_E	07-Jan-19
NVM	27-Dec-15	NWN_E	27-Jan-17	NWN_E	22-Jan-18	NWP_E	07-Jan-19
NWN_W	27-Dec-15	NVN	30-Jan-17	NWM_E	27-Jan-18	NWN_W	20-Jan-19
NWM_W	06-Jan-16	NWN_W	30-Jan-17	NXM	27-Jan-18	NXM	22-Jan-19
NVM	06-Jan-16	NWP_W	30-Jan-17	NYM	27-Jan-18	NYM	22-Jan-19
NVN	16-Jan-16	NWN_E	06-Feb-17	NYN	27-Jan-18	NYN	22-Jan-19
NWP_W	16-Jan-16	NWP_E	06-Feb-17	NXN	27-Jan-18	NXN	22-Jan-19
NWM_E	23-Jan-16	NWP_W	09-Feb-17	NWN_E	27-Jan-18	NWM_E	11-Feb-19
NXM	23-Jan-16	NWM_W	19-Feb-17	NXM	01-Feb-18	NWN_E	11-Feb-19
NWL	23-Jan-16	NWN_W	19-Feb-17	NXL	01-Feb-18	NWM_W	14-Feb-19
NXL	23-Jan-16	NXN	08-Mar-17	NYN	01-Feb-18	NVN	14-Feb-19
NYN	23-Jan-16	NWM_E	28-Mar-17	NWM_W	19-Feb-18	NWN_W	14-Feb-19
NXN	23-Jan-16	NXM	28-Mar-17	NVM	19-Feb-18	NWL	16-Feb-19
NWN_E	23-Jan-16	NYM	28-Mar-17	NVN	19-Feb-18	NXL	16-Feb-19
NWP_E	23-Jan-16	NWL	28-Mar-17	NWN_W	19-Feb-18	NWM_E	21-Feb-19
NWM_W	26-Jan-16	NXL	28-Mar-17	NWP_W	19-Feb-18	NXM	21-Feb-19
NVN	26-Jan-16	NYN	28-Mar-17	NXP	03-Mar-18	NYM	21-Feb-19
NWN_W	26-Jan-16	NXM	17-Apr-17	NWP_E	03-Mar-18	NWN_E	21-Feb-19
NYM	12-Feb-16	NYM	17-Apr-17	NWM_W	06-Mar-18	NWM_E	26-Feb-19
NXP	03-Mar-16	NZM	17-Apr-17	NVN	06-Mar-18	NXM	26-Feb-19
NWM_E	22-Apr-16	NYL	17-Apr-17	NWN_W	06-Mar-18	NYM	26-Feb-19

---

[17] XXX_E and XXX_W refer to tiles on the Sentinel swath boundary which are split diagonally into East and West tiles.

NXM	22-Apr-16	NWL	17-Apr-17	NWP_W	06-Mar-18	NWL	26-Feb-19
NWL	22-Apr-16	NXL	17-Apr-17	NWP_E	23-Mar-18	NXL	26-Feb-19
NXL	22-Apr-16	NYN	17-Apr-17	NWP_W	20-Apr-18	NXN	26-Feb-19
NYN	22-Apr-16	NWM_E	27-Apr-17	NXN	22-Apr-18	NWN_E	26-Feb-19
NXN	22-Apr-16	NXN	27-Apr-17	NWP_E	22-Apr-18	NXN	03-Mar-19
				NXM	27-Apr-18	NWN_E	03-Mar-19
				NYM	27-Apr-18	NWM_W	11-Mar-19
				NZM	27-Apr-18	NVM	11-Mar-19
				NYL	27-Apr-18	NVN	11-Mar-19
				NWL	27-Apr-18	NWN_W	11-Mar-19
				NXL	27-Apr-18	NWL	13-Mar-19
				NYN	27-Apr-18	NXL	13-Mar-19
						NWP_E	13-Mar-19
						NVN	26-Mar-19
						NWN_W	26-Mar-19
						NWP_W	26-Mar-19
						NWM_E	02-Apr-19
						NXM	02-Apr-19
						NXP	02-Apr-19
						NYN	02-Apr-19
						NXN	02-Apr-19
						NWN_E	02-Apr-19
						NWP_E	02-Apr-19
						NWP_W	10-Apr-19
						NXM	22-Apr-19
						NYM	22-Apr-19
						NZM	22-Apr-19
						NYL	22-Apr-19
						NYN	22-Apr-19

C-2 Software

ArcGIS Pro 2.3.1

MATLAB 2019b + Deep Learning Toolbox

Sen2Cor_v2.5.5

C-3 Change mapping schema

Year x		Year x+1		Change sub-class		Change macro-class
Mine	2	Mine	2	No Change	0	No Change
Built	7	Mine	2	Built to Mine	5	New Mined
Vegetation	11	Mine	2	Vegetation to Mine	9	New Mined
Cloud	0	Mine	2	Cloud to mine	-2	New Mined
Mine	2	Built	7	Mine to Built	-5	New Built
Built	7	Built	7	No Change	0	No Change
Vegetation	11	Built	7	Vegetation to Built	4	New Built
Cloud	0	Built	7	Cloud to Built	-7	New Built
Mine	2	Vegetation	11	Mine to Vegetation	-9	New Vegetation
Built	7	Vegetation	11	Built to Vegetation	-4	New Vegetation
Vegetation	11	Vegetation	11	No Change	0	No Change
Cloud	0	Vegetation	11	Cloud to Vegetation	-11	New Vegetation
Mine	2	Cloud	0	Mine to Cloud	2	No Change*
Built	7	Cloud	0	Built to Cloud	7	No Change*
Vegetation	11	Cloud	0	Vegetation to Cloud	11	No Change*
Cloud	0	Cloud	0	No Change	0	No Change*

* Where the change class was from any ground class to the cloud class, the pixel was assumed unchanged and its prior year's value was used.

References

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., Kudlur, M., Levenberg, J., Monga, R., Moore, S., Murray, D.G., Steiner, B., Tucker, P., Vasudevan, V., Warden, P., Wicke, M., Yu, Y., Zheng, X., 2016. TensorFlow: A System for Large-Scale Machine Learning, in: 12th USENIX Conference on Operating Systems Design and Implementation, OSDI'16. USENIX Association, USA, pp. 265–283.

Abdoulatif, A.S., Baratoux, D., Baratoux, L., Ahoussi, E.K., Yao, A., Colin, F., Jean Kan, K., Kamagaté, B., Zié, O., Gnamba Emmanuel Franck, G., 2019. Remote-sensing mapping (Landsat) of abandoned mining wastes over time by artisanal gold mining in Koma Bangou (Liptako, Niger)., in: AGU Fall Meeting Abstracts.

Abellán, A., Calvet, J., Vilaplana, J.M., Blanchard, J., 2010. Detection and spatial prediction of rockfalls by means of terrestrial laser scanner monitoring. Geomorphology 119, 162–171. https://doi.org/https://doi.org/10.1016/j.geomorph.2010.03.016

Aber, J.S., Marzolff, I., Ries, J., 2010. Small-Format Aerial Photography: Principles, Techniques and Geoscience Applications. Elsevier B.V., Amsterdam.

Agrawal, A., Nakazawa, A., Takemura, H., 2009. MMM-classification of 3D Range Data, in: IEEE International Conference on Robotics and Automation. Kobe, pp. 2003–2008. https://doi.org/10.1109/ROBOT.2009.5152539

Alamiyan-Harandi, F., Derhami, V., Jamshidi, F., 2020. Combination of Recurrent Neural Network and Deep Learning for Robot Navigation Task in Off-Road Environment. Robotica 38, 1450–1462. https://doi.org/10.1017/S0263574719001565

Albee, A.L., Arvidson, R.E., Palluconi, F., Thorpe, T., 2001. Overview of the Mars Global Surveyor mission. J. Geophys. Res. Planets 106, 23291–23316. https://doi.org/10.1029/2000JE001306

Ali, D., Frimpong, S., 2020. Artificial intelligence, machine learning and process automation: existing knowledge frontier and way forward for mining sector. Artif. Intell. Rev. 53, 6025–6042. https://doi.org/10.1007/s10462-020-09841-6

Alvi, M., Zisserman, A., Nellåker, C., 2019. Turning a Blind Eye: Explicit Removal of Biases and Variation from Deep Neural Network Embeddings, in: Leal-Taixé, L., Roth, S. (Eds.), Computer Vision – ECCV 2018 Workshops. Springer International Publishing, Cham, Switzerland, pp. 556–572. https://doi.org/10.1007/978-3-030-11009-3_34

Anguelov, D., Taskar, B., Chatalbashev, V., Koller, D., Gupta, D., Heitz, G., Ng, A., 2005. Discriminative Learning of Markov Random Fields for Segmentation of 3D Scan Data, in: 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05). IEEE, pp. 169–176. https://doi.org/10.1109/CVPR.2005.133

Arandjelovic, R., Gronat, P., Torii, A., Pajdla, T., Sivic, J., 2018. NetVLAD: CNN Architecture for Weakly Supervised Place Recognition. IEEE Trans. Pattern Anal. Mach. Intell. 40, 1437–1451. https://doi.org/10.1109/TPAMI.2017.2711011

Arazo, E., Ortego, D., Albert, P., O'Connor, N.E., McGuinness, K., 2020. Pseudo-Labeling and Confirmation Bias in Deep Semi-Supervised Learning, in: 2020 International Joint Conference on Neural Networks (IJCNN). pp. 1–8. https://doi.org/10.1109/IJCNN48605.2020.9207304

Armeni, I., Sax, S., Zamir, A.R., Savarese, S., 2017. Joint 2D-3D-Semantic Data for Indoor Scene Understanding. arXiv Prepr.

Ashby, A., van Etten, E., Lund, M., 2016. Pitfalls of gold mine sites in care and maintenance, in: Fourie, A., Tibbett, M. (Eds.), 11th International Conference on Mine Closure. Australian Centre for Geomechanics, Perth, pp. 313–324. https://doi.org/10.36487/ACG_rep/1608_22_Ashby

Asner, G.P., Llactayo, W., Tupayachi, R., Luna, E.R., 2013. Elevated rates of gold mining in the Amazon revealed through high-resolution monitoring. Natl. Acad. Sci. 110, 18454–18459. https://doi.org/10.1073/pnas.1318271110

Asner, G.P., Tupayachi, R., 2016. Accelerated losses of protected forests from gold mining in the Peruvian Amazon. Environ. Res. Lett. 12, 094004. https://doi.org/10.1088/1748-9326/aa7dab

Attard, L., Debono, C.J., Valentino, G., Di Castro, M., 2018. Tunnel inspection using photogrammetric techniques and image processing: A review. ISPRS J. Photogramm. Remote Sens. 144, 180–188. https://doi.org/10.1016/j.isprsjprs.2018.07.010

Azmin, B.A., Bell, I., Sameti, B., Ko, D., Tafazoli, S., 2016. Porta Metrics TM safety and consistancy of results in image-based fragmentation analysis, in: International Society of Explosives Engineers.

Badrinarayanan, V., Kendall, A., Cipolla, R., 2017. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. IEEE Trans. Pattern Anal. Mach. Intell. 39, 2481–2495. https://doi.org/10.1109/TPAMI.2016.2644615

Balabin, R.M., Safieva, R.Z., Lomakina, E.I., 2010. Gasoline classification using near infrared (NIR) spectroscopy data: Comparison of multivariate techniques. Anal. Chim. Acta 671, 27–35. https://doi.org/10.1016/j.aca.2010.05.013

Balado, J., Díaz-Vilariño, L., Arias, P., González-Jorge, H., 2018. Automatic classification of urban ground elements from mobile laser scanning data. Autom. Constr. 86, 226–239. https://doi.org/10.1016/j.autcon.2017.09.004

Balaniuk, R., Isupova, O., Reece, S., 2020. Mining and Tailings Dam Detection in Satellite Imagery Using Deep Learning. Sensors 20, 6936. https://doi.org/10.3390/s20236936

Ball, J.E., Anderson, D.T., Chan, C.S., 2017. Comprehensive survey of deep learning in remote sensing: theories, tools, and challenges for the community. J. Appl. Remote Sens. 11, 042609. https://doi.org/10.1117/1.JRS.11.042609

Bamford, T., Esmaeili, K., Schoellig, A.P., 2017. A real-time analysis of post-blast rock fragmentation using UAV technology. Int. J. Mining, Reclam. Environ. 31, 439–456. https://doi.org/10.1080/17480930.2017.1339170

Bansah, K.J., Dumakor-Dupey, N.K., Kansake, B.A., Assan, E., Bekui, P., 2018. Socioeconomic and environmental assessment of informal artisanal and small-scale mining in Ghana. J. Clean. Prod. 202, 465–475. https://doi.org/10.1016/j.jclepro.2018.08.150

Barenblitt, A., Payton, A., Lagomasino, D., Fatoyinbo, T., Asare, K., Aidoo, K., Pigott, H., Som, C.K., Seidu, O., Smeets, L., others, 2020. Spatial distribution of artisanal goldmining in Ghana: Using machine learning and Google Earth Engine to quantify conversion of vegetation to gold mines, in: AGU Fall Meeting 2020.

Barich, A., 2019. Earth AI to start autonomous rig in April before further exploration expansion. S&P Glob. Mark. Intell.

Basommi, P.L., Guan, Q., Cheng, D., 2015. Exploring Land use and Land cover change in the mining areas of Wa East District, Ghana using Satellite Imagery. Open Geosci. 7, 618–626. https://doi.org/10.1515/geo-2015-0058

Bassier, M., Van Genechten, B., Vergauwen, M., 2019. Classification of sensor independent point cloud data of building objects using random forests. J. Build. Eng. 21, 468–477. https://doi.org/10.1016/j.jobe.2018.04.027

Bayer, B.E., 1975. Color Imaging Array. United States Patent No. 3971065.

Behley, J., Garbade, M., Milioto, A., Quenzel, J., Behnke, S., Stachniss, C., Gall, J., 2019. SemanticKITTI: A Dataset for Semantic Scene Understanding of LiDAR Sequences, in: 2019 IEEE/CVF International Conference on Computer Vision (ICCV). IEEE, pp. 9296–9306. https://doi.org/10.1109/ICCV.2019.00939

Behley, J., Kersting, K., Schulz, D., Steinhage, V., Cremers, A.B., 2010. Learning to hash logistic regression for fast 3D scan point classification, in: 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems. IEEE, pp. 5960–5965. https://doi.org/10.1109/IROS.2010.5650093

Bello, S.A., Yu, S., Wang, C., Adam, J.M., Li, J., 2020. Review: Deep Learning on 3D Point Clouds. Remote Sens. 12, 1729. https://doi.org/10.3390/rs12111729

Bengio, Y., 2012. Practical recommendations for gradient-based training of deep architectures, in: Montavon, G., Orr, G.B., Müller, K.-R. (Eds.), Neural Networks: Tricks of the Trade. Springer, Berlin, Germany.

Bengio, Y., LeCun, Y., 2007. Scaling learning algorithms towards AI, in: Bottou, L., Chapelle, O., DeCoste, D., Weston, J. (Eds.), Large-Scale Kernel Machines. MIT Press, pp. 321–359.

Beraldin, J.-A., Blais, F., Lohr, U., 2010. Laser Scanning Technology, in: Vosselman, G., Maas, H.-G. (Eds.), Airborne and Terrestrial Laser Scanning. Whittles Publishing, Caithness, UK, pp. 1–44.

Berger, M., Moreno, J., Johannessen, J.A., Levelt, P.F., Hanssen, R.F., 2012. ESA's sentinel missions in support of Earth system science. Remote Sens. Environ. 120, 84–90. https://doi.org/10.1016/j.rse.2011.07.023

BGS, 2016. Tile: EW338 Devon.

Bissiri, Y., Baiden, G., Filion, S., Saari, A., 2008. Automated surveying device for underground navigation. Min. Technol. 117, 71–82. https://doi.org/10.1179/174328608X362640

Blakely, R.J., Simpson, R.W., 1986. Approximating edges of source bodies from magnetic or gravity anomalies. Geophysics 51, 1494–1498. https://doi.org/10.1190/1.1442197

Blaschke, T., Hay, G.J., Kelly, M., Lang, S., Hofmann, P., Addink, E., Queiroz Feitosa, R., van der Meer, F., van der Werff, H., van Coillie, F., Tiede, D., 2014. Geographic Object-Based Image Analysis – Towards a new paradigm. ISPRS J. Photogramm. Remote Sens. 87, 180–191. https://doi.org/10.1016/j.isprsjprs.2013.09.014

Blomley, R., Jutzi, B., Weinmann, M., 2016. Classification of airborne laser scanning data using geometric multi-scale features and different neighbourhood types. ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci. III–3, 169–176. https://doi.org/10.5194/isprsannals-III-3-169-2016

Boadi, S., Nsor, C.A., Antobre, O.O., Acquah, E., 2016. An analysis of illegal mining on the Offin shelterbelt forest reserve, Ghana: Implications on community livelihood. J. Sustain. Min. 15, 115–119. https://doi.org/10.1016/j.jsm.2016.12.001

Boakye, E., Anyemedu, F.O.K., Quaye-Ballard, J.A., Donkor, E.A., 2020. Spatio-temporal analysis of land use/cover changes in the Pra River Basin, Ghana. Appl. Geomatics 12, 83–93. https://doi.org/10.1007/s12518-019-00278-3

Body, D., 2014. The Changing Face of Surveying within the Mining Industry.

Bonetto, S., Facello, A., Ferrero, A.M., Umili, G., 2015. A tool for semi-automatic linear feature detection based on DTM. Comput. Geosci. 75, 1–12. https://doi.org/10.1016/j.cageo.2014.10.005

Borodinov, N., Neumayer, S., Kalinin, S. V, Ovchinnikova, O.S., Vasudevan, R.K., Jesse, S., 2019. Deep neural networks for understanding noisy data applied to physical property extraction in scanning probe microscopy. npj Comput. Mater. 5, 25. https://doi.org/10.1038/s41524-019-0148-5

Borthwick, J.R., 2009. Mining haul truck pose estimation and load profiling using stereo vision. University of British Columbia.

Botchwey, G., Crawford, G., Loubere, N., Lu, J., 2019. South-South Irregular Migration: The Impacts of China's Informal Gold Rush in Ghana. Int. Migr. 57, 310–328. https://doi.org/10.1111/imig.12518

Boulch, A., 2019. Generalizing discrete convolutions for unstructured point clouds, in: Biasotti, S., Lavoué, G., Falcidieno, B., Pratikakis, I. (Eds.), Eurographics Workshop on 3D Object Retrieval. pp. 1–11. https://doi.org/10.2312/3dor.20191064

Boulch, A., Guerry, J., Le Saux, B., Audebert, N., 2018. SnapNet: 3D point cloud semantic labeling with 2D deep segmentation networks. Comput. Graph. 71, 189–198. https://doi.org/10.1016/j.cag.2017.11.010

Bradski, G., 2000. The OpenCV Library. Dr. Dobb's J. Softw. Tools.

Brandt, M., Tucker, C.J., Kariryaa, A., Rasmussen, K., Abel, C., Small, J., Chave, J., Rasmussen, L.V., Hiernaux, P., Diouf, A.A., Kergoat, L., Mertz, O., Igel, C., Gieseke, F., Schöning, J., Li, S., Melocik, K., Meyer, J., Sinno, S., Romero, E., Glennie, E., Montagu, A., Dendoncker, M., Fensholt, R., 2020. An unexpectedly large count of

trees in the West African Sahara and Sahel. Nature 587, 78–82.
https://doi.org/10.1038/s41586-020-2824-5

Breiman, L., 2001. Random Forests. Mach. Learn. 45, 5–32.
https://doi.org/https://doi.org/10.1023/A:1010933404324

Bressan, T.S., Kehl de Souza, M., Girelli, T.J., Junior, F.C., 2020. Evaluation of
machine learning methods for lithology classification using geophysical data.
Comput. Geosci. 139, 104475. https://doi.org/10.1016/j.cageo.2020.104475

Briese, C., 2010. Extraction of Digital Terrain Models, in: Vosselman, G., Maas, H.-
G. (Eds.), Airborne and Terrestrial Laser Scanning. Whittles Publishing, Dunbeath.

Burkov, A., 2019. The Hundred-Page Machine Learning Book. Andriy Burkov,
Quebec City, Canada.

Caballero Espejo, J., Messinger, M., Román-Dañobeytia, F., Ascorra, C., Fernandez,
L., Silman, M., 2018. Deforestation and Forest Degradation Due to Gold Mining in
the Peruvian Amazon: A 34-Year Perspective. Remote Sens. 10, 1903.
https://doi.org/10.3390/rs10121903

Campbell, A.D., Thurley, M.J., 2017. Application of laser scanning to measure
fragmentation in underground mines. Min. Technol. 126, 1–8.
https://doi.org/10.1080/14749009.2017.1296668

Carrivick, J.L., Smith, M.W., Quincey, D.J., 2016. Structure from Motion in the
Geosciences. John Wiley & Sons, Oxford, UK.

Caruana, R., Niculescu-Mizil, A., 2006. An empirical comparison of supervised
learning algorithms, in: 23rd International Conference on Machine Learning - ICML
'06. ACM Press, New York, New York, USA, pp. 161–168.
https://doi.org/10.1145/1143844.1143865

Chapelle, O., Zien, A., 2005. Semi-Supervised Classification by Low Density
Separation. AISTATS 57–64.

Chehata, N., Guo, L., Mallet, C., 2009. Airborne LiDAR feature selection for urban classification using random forests. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. 38, 207–212.

Chen, S., Walske, M.L., Davies, I.J., 2018. Rapid mapping and analysing rock mass discontinuities with 3D terrestrial laser scanning in the underground excavation. Int. J. Rock Mech. Min. Sci. 110, 28–35. https://doi.org/10.1016/j.ijrmms.2018.07.012

Chen, W., Li, X., He, H., Wang, L., 2017. A Review of Fine-Scale Land Use and Land Cover Classification in Open-Pit Mining Areas by Remote Sensing Techniques. Remote Sens. 10, 15. https://doi.org/10.3390/rs10010015

Chen, Z., Liu, X., Yang, J., Little, E., Zhou, Y., 2020. Deep learning-based method for SEM image segmentation in mineral characterization, an example from Duvernay Shale samples in Western Canada Sedimentary Basin. Comput. Geosci. 138, 104450. https://doi.org/10.1016/j.cageo.2020.104450

Cheng, G., Zhou, P., Han, J., 2016. Learning Rotation-Invariant Convolutional Neural Networks for Object Detection in VHR Optical Remote Sensing Images. IEEE Trans. Geosci. Remote Sens. 54, 7405–7415. https://doi.org/10.1109/TGRS.2016.2601622

Cherabier, I., Hane, C., Oswald, M.R., Pollefeys, M., 2016. Multi-Label Semantic 3D Reconstruction Using Voxel Blocks, in: 2016 Fourth International Conference on 3D Vision (3DV). IEEE, pp. 601–610. https://doi.org/10.1109/3DV.2016.68

Chevrel, M., Courtois, M., Weill, G., 1981. The SPOT satellite remote sensing mission. Photogramm. Eng. Remote Sensing 47, 1163–1171.

Chiang, A., Yeh, Y., 2015. Anomaly Detection Ensembles: In Defense of the Average, in: 2015 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT). pp. 207–210. https://doi.org/10.1109/WI-IAT.2015.260

Chollet, F., 2015. Keras.

Chollet, F., 2017. Deep Learning with Python. Manning Publications, Shelter Island, NY, USA.

Chowdhary, A., Jadhav, A., Jaiswal, A.K., Mukherjee, B., Shah, U., Gawade, P.S., Jagyasi, B., Vegesna, R.G.V., Mohammad, N., Contractor, G., 2019. Detection of coal stockpiles using geospatial satellite images, in: Schulz, K., Nikolakopoulos, K.G., Michel, U. (Eds.), Earth Resources and Environmental Remote Sensing/GIS Applications X. SPIE, p. 27. https://doi.org/10.1117/12.2532523

Cireşan, D.C., Meier, U., Gambardella, L.M., Schmidhuber, J., 2010. Deep, Big, Simple Neural Nets for Handwritten Digit Recognition. Neural Comput. 22, 3207–3220. https://doi.org/10.1162/NECO_a_00052

Cochran, W.G., 1977. Sampling Techniques, 3rd ed. John Wiley & Sons, New York, NY, USA.

Coggan, J., Wetherelt, A., Gwynn, X.P., Flynn, Z.N., 2007. Comparison of Hand-mapping With Remote Data Capture Systems For Effective Rock Mass Characterisation, in: 11th ISRM Congress. International Society for Rock Mechanics and Rock Engineering, Lisbon, Portugal.

Congedo, L., 2016. Semi-Automatic Classification Plugin Documentation. Release 6.0.1.1. https://doi.org/10.13140/RG.2.2.29474.02242/1

Corbett, T., O'Faircheallaigh, C., Regan, A., 2017. 'Designated areas' and the regulation of artisanal and small-scale mining. Land use policy 68, 393–401. https://doi.org/10.1016/j.landusepol.2017.08.004

Coulter, L.L., Stow, D.A., Tsai, Y.-H., Ibanez, N., Shih, H., Kerr, A., Benza, M., Weeks, J.R., Mensah, F., 2016. Classification and assessment of land cover and land use change in southern Ghana using dense stacks of Landsat 7 ETM+ imagery. Remote Sens. Environ. 184, 396–409. https://doi.org/10.1016/j.rse.2016.07.016

Cracknell, M.J., Reading, A.M., 2014. Geological mapping using remote sensing data: A comparison of five machine learning algorithms, their response to variations in the spatial distribution of training data and the use of explicit spatial information. Comput. Geosci. 63, 22–33. https://doi.org/10.1016/j.cageo.2013.10.008

Cryderman, C., Mah, S.B., Shufletoski, A., 2014. Evaluation of UAV Photogrammetric Accuracy for Mapping and Earthworks Computations. Geomatica 68, 309–317. https://doi.org/10.5623/cig2014-405

Data Study Group Team, 2020. Data Study Group Final Report: SenSat. https://doi.org/10.5281/zenodo.3878499

Dekker, F., 2019. Cocoa not main cause of deforestation in Ghana [WWW Document]. Satelligence. URL https://satelligence.com/news/2019/5/17/cocoa-not-main-cause-of-deforestation-in-ghana (accessed 12.4.19).

Demirel, N., Emil, M.K., Duzgun, H.S., 2011. Surface coal mine area monitoring using multi-temporal high-resolution satellite imagery. Int. J. Coal Geol. 86, 3–11. https://doi.org/10.1016/j.coal.2010.11.010

Deng, J., Dong, W., Socher, R., Li, L., Kai Li, Li Fei-Fei, 2009. ImageNet: A large-scale hierarchical image database, in: 2009 IEEE Conference on Computer Vision and Pattern Recognition. pp. 248–255. https://doi.org/10.1109/CVPR.2009.5206848

Dietterich, T.G., 2000. Ensemble Methods in Machine Learning, in: Multiple Classifier Systems: First International Workshop, MCS 2000, Lecture Notes in Computer Science. Springer, Berlin, Germany, pp. 1–15. https://doi.org/10.1007/3-540-45014-9_1

DigitalGlobe, 2015. World Imagery Basemap.

Dines, H.G., 1988. The metalliferous mining region of south west England, 3rd ed. British Geological Survey, Nottingham, UK.

Doneus, M., 2013. Openness as Visualization Technique for Interpretative Mapping of Airborne Lidar Derived Digital Terrain Models. Remote Sens. 5, 6427–6442. https://doi.org/10.3390/rs5126427

Dong, P., Chen, Q., 2017. Lidar remote sensing and applications. CRC Press, Boca Raton, FL, USA.

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese,

F., Bargellini, P., 2012. Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. Remote Sens. Environ. 120, 25–36. https://doi.org/10.1016/j.rse.2011.11.026

Durrant-Whyte, H., Bailey, T., 2006. Simultaneous localization and mapping: part I. IEEE Robot. Autom. Mag. 13, 99–110. https://doi.org/10.1109/MRA.2006.1638022

Environment Agency, 2020. LiDAR Capture Programme [WWW Document]. URL https://environment.maps.arcgis.com/apps/opsdashboard/index.html#/27d3192741c e48dfabdeb20f3d740bc7 (accessed 1.19.21).

ESRI, 2019. ArcGIS Pro.

Ester, M., Kriegel, H.-P., Sander, J., Xu, X., 1996. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise, in: Second International Conference on Knowledge Discovery and Data Mining. AAAI Press, Portland, Oregon, pp. 226–231.

European Space Agency, 2019. Level-2A [WWW Document]. Prod. Types User Guid. URL https://sentinel.esa.int/web/sentinel/user-guides/sentinel-2-msi/product-types/level-2a (accessed 12.5.19).

European Space Agency, 2015. Sentinel-2 User Handbook.

Eyre, M., Wetherelt, A., Coggan, J., 2016. Evaluation of automated underground mapping solutions for mining and civil engineering applications. J. Appl. Remote Sens. 10, 046011. https://doi.org/10.1117/1.JRS.10.046011

Eyre, M.L., Foster, P.J., Jobling-Purser, J., Coggan, J., 2015. The benefits of laser scanning and 3D modelling in accident investigation: in a mining context. Min. Technol. 124, 73–77. https://doi.org/10.1179/1743286315Y.0000000004

Fekete, S., Diederichs, M., 2013. Integration of three-dimensional laser scanning with discontinuum modelling for stability analysis of tunnels in blocky rockmasses. Int. J. Rock Mech. Min. Sci. 57, 11–23. https://doi.org/10.1016/j.ijrmms.2012.08.003

Fekete, S., Diederichs, M., Lato, M., 2010. Geotechnical and operational applications for 3-dimensional laser scanning in drill and blast tunnels. Tunn. Undergr. Sp. Technol. 25, 614–628. https://doi.org/10.1016/j.tust.2010.04.008

Fernández-Lozano, J., Gutiérrez-Alonso, G., Fernández-Morán, M.Á., 2015. Using airborne LiDAR sensing technology and aerial orthoimages to unravel roman water supply systems and gold works in NW Spain (Eria valley, León). J. Archaeol. Sci. 53, 356–373. https://doi.org/10.1016/j.jas.2014.11.003

Ferreira, E., Brito, M., Balaniuk, R., Alvim, M.S., Santos, J.A. dos, 2020. Brazildam: A Benchmark Dataset For Tailings Dam Detection, in: 2020 IEEE Latin American GRSS & ISPRS Remote Sensing Conference (LAGIRS). IEEE, pp. 339–344. https://doi.org/10.1109/LAGIRS48042.2020.9165620

Foley, J.D., van Dam, A., Feiner, S.K., Hughes, J.F., 1990. Computer Graphics: Principles and Practice (2nd Ed.). Addison-Wesley Longman Publishing Co., Inc., USA.

Folleco, A., Khoshgoftaar, T.M., Van Hulse, J., Bullard, L., 2008. Identifying learners robust to low quality data, in: 2008 IEEE International Conference on Information Reuse and Integration. IEEE, pp. 190–195. https://doi.org/10.1109/IRI.2008.4583028

Fonte, J., Pires, H., Gonçalves-Seco, L., Matias Rodriguez, R., Lima, A., 2014. Archaeological research of ancient mining landscapes in Galicia (Spain) using Airborne Laser Scanning data, in: Ancient Mining Landscapes in Western Europe. Boticas, Portugal. https://doi.org/10.13140/2.1.3122.5608

Forestry Commission, 2010. Shapefiles of Forest and Game Reserves in Ghana 2010.

Forkuor, G., Benewinde Zoungrana, J.-B., Dimobe, K., Ouattara, B., Vadrevu, K.P., Tondoh, J.E., 2020. Above-ground biomass mapping in West African dryland forest using Sentinel-1 and 2 datasets - A case study. Remote Sens. Environ. 236, 111496. https://doi.org/10.1016/j.rse.2019.111496

Freeland, T., Heung, B., Burley, D. V., Clark, G., Knudby, A., 2016. Automated feature extraction for prospection and analysis of monumental earthworks from aerial

LiDAR in the Kingdom of Tonga. J. Archaeol. Sci. 69, 64–74. https://doi.org/10.1016/j.jas.2016.04.011

Frei, E., Kung, J., Bukowski, R., 2004. High-Definition Surveying (Hds): a New Era in Reality Capture, in: Thies, M., Koch, B., Spiecker, H., Weinacker, H. (Eds.), International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences. ISPRS, Freiburg, Germany, pp. 262–271.

Frimpong, S., Li, Y., Agarwal, S., 2007. Frontier research in dynamic control, vision and collision avoidance for truck haulage, in: 2007 SME Annual Meeting. Society of Manufacturing Engineers.

Froese, C.R., Mei, S., 2008. Mapping and monitoring coal mine subsidence using LiDAR and InSAR, in: GeoEdmonton2008. Edmunton, Canada, pp. 1127–1133.

Fu, W., Ma, J., Chen, P., Chen, F., 2020. Remote Sensing Satellites for Digital Earth, in: Guo, H., Goodchild, M.F., Annoni, A. (Eds.), Manual of Digital Earth. Springer Singapore, Singapore, pp. 55–123. https://doi.org/10.1007/978-981-32-9915-3_3

Fukushima, K., 1988. Neocognitron: A hierarchical neural network capable of visual pattern recognition. Neural Networks 1, 119–130. https://doi.org/10.1016/0893-6080(88)90014-7

Ganić, A., Milutinović, A., Tokalić, R., Ognjanović, S., 2011. Measuring methods for cross sections of underground mine chambers. Podzemn. Rad. 0.

Gao, H., Mao, J., Zhou, J., Huang, Z., Wang, L., Xu, W., 2015. Are You Talking to a Machine? Dataset and Methods for Multilingual Image Question Answering, in: 28th International Conference on Neural Information Processing Systems - Volume 2, NIPS'15. MIT Press, Cambridge, MA, USA, pp. 2296–2304.

Garcia-Garcia, A., Orts-Escolano, S., Oprea, S., Villena-Martinez, V., Martinez-Gonzalez, P., Garcia-Rodriguez, J., 2018. A survey on deep learning techniques for image and video semantic segmentation. Appl. Soft Comput. 70, 41–65. https://doi.org/10.1016/j.asoc.2018.05.018

Gascon, F., Bouzinac, C., Thépaut, O., Jung, M., Francesconi, B., Louis, J., Lonjou, V., Lafrance, B., Massera, S., Gaudel-Vacaresse, A., Languille, F., Alhammoud, B., Viallefont, F., Pflug, B., Bieniarz, J., Clerc, S., Pessiot, L., Trémas, T., Cadau, E., De Bonis, R., Isola, C., Martimort, P., Fernandez, V., 2017. Copernicus Sentinel-2A Calibration and Products Validation Status. Remote Sens. 9, 584. https://doi.org/10.3390/rs9060584

Gebhardt, S., Payzer, E., Salemann, L., Fettinger, A., Rotenberg, E., Seher, C., 2009. Polygons, point-clouds, and voxels, A comparison of high-fidelity terrain representations. Fall Simul. Interoperability Work. 2009, 2009 Fall SIW 357–365.

Geiger, A., Lenz, P., Stiller, C., Urtasun, R., 2013. Vision meets robotics: The KITTI dataset. Int. J. Rob. Res. 32, 1231–1237. https://doi.org/10.1177/0278364913491297

Geman, S., Geman, D., 1984. Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. IEEE Trans. Pattern Anal. Mach. Intell. PAMI-6, 721–741. https://doi.org/10.1109/TPAMI.1984.4767596

Géron, A., 2017. Hands-On Machine Learning with Scikit-Learn & Tensorflow, 1st ed. O'Reilly, Sebastopol, CA, USA.

Ghosh, S., Das, N., Das, I., Maulik, U., 2019. Understanding Deep Learning Techniques for Image Segmentation. ACM Comput. Surv. 52, 1–35. https://doi.org/10.1145/3329784

Giannoni, L., Lange, F., Tachtsidis, I., 2018. Hyperspectral imaging solutions for brain tissue metabolic and hemodynamic monitoring: past, current and future developments. J. Opt. 20, 044009. https://doi.org/10.1088/2040-8986/aab3a6

Gigli, G., Casagli, N., 2011. Semi-automatic extraction of rock mass structural data from high resolution LiDAR point clouds. Int. J. Rock Mech. Min. Sci. 48, 187–198. https://doi.org/10.1016/j.ijrmms.2010.11.009

Gikas, V., 2012. Three-Dimensional Laser Scanning for Geometry Documentation and Construction Management of Highway Tunnels during Excavation. Sensors 12, 11249–11270. https://doi.org/10.3390/s120811249

Girardeau-Montaut, D., 2016. Cloud compare documentation [WWW Document]. URL http://www.cloudcompare.org/doc/wiki/index.php?title=Main_Page (accessed 10.23.18).

Girshick, R., 2015. Fast R-CNN, in: 2015 IEEE International Conference on Computer Vision (ICCV). IEEE, pp. 1440–1448. https://doi.org/10.1109/ICCV.2015.169

Girshick, R., Donahue, J., Darrell, T., Malik, J., 2014. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation, in: 2014 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, pp. 580–587. https://doi.org/10.1109/CVPR.2014.81

González-Jorge, H., Rodríguez-Gonzálvez, P., Shen, Y., Lagüela, S., Díaz-Vilariño, L., Lindenbergh, R., González-Aguilera, D., Arias, P., 2018. Metrological intercomparison of six terrestrial laser scanning systems. IET Sci. Meas. Technol. 12, 218–222. https://doi.org/10.1049/iet-smt.2017.0209

Goodfellow, I., Bengio, Y., Courville, A., 2016. Deep Learning. MIT Press, Cambridge, MA, USA.

Grebby, S., Cunningham, D., Naden, J., Tansey, K., 2012. Application of airborne LiDAR data and airborne multispectral imagery to structural mapping of the upper section of the Troodos ophiolite, Cyprus. Int. J. Earth Sci. 101, 1645–1660. https://doi.org/10.1007/s00531-011-0742-3

Grigorescu, S., Trasnea, B., Cocias, T., Macesanu, G., 2020. A survey of deep learning techniques for autonomous driving. J. F. Robot. 37, 362–386. https://doi.org/10.1002/rob.21918

Grotzinger, J.P., Crisp, J., Vasavada, A.R., Anderson, R.C., Baker, C.J., Barry, R., Blake, D.F., Conrad, P., Edgett, K.S., Ferdowski, B., Gellert, R., Gilbert, J.B., Golombek, M., Gómez-Elvira, J., Hassler, D.M., Jandura, L., Litvak, M., Mahaffy, P., Maki, J., Meyer, M., Malin, M.C., Mitrofanov, I., Simmonds, J.J., Vaniman, D., Welch, R. V, Wiens, R.C., 2012. Mars Science Laboratory Mission and Science Investigation. Space Sci. Rev. 170, 5–56. https://doi.org/10.1007/s11214-012-9892-2

Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., Liu, T., Wang, X., Wang, G., Cai, J., Chen, T., 2018. Recent advances in convolutional neural networks. Pattern Recognit. 77, 354–377. https://doi.org/10.1016/j.patcog.2017.10.013

Guyot, A., Hubert-Moy, L., Lorho, T., 2018. Detecting Neolithic Burial Mounds from LiDAR-Derived Elevation Data Using a Multi-Scale Approach and Machine Learning Techniques. Remote Sens. 10, 225. https://doi.org/10.3390/rs10020225

Hackel, T., Savinov, N., Ladicky, L., Wegner, J.D., Schindler, K., Pollefeys, M., 2017a. Semantic3d.net: A new large-scale point cloud classification benchmark. ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci. IV-1/W1, 91–98. https://doi.org/10.5194/isprs-annals-IV-1-W1-91-2017

Hackel, T., Wegner, J.D., Schindler, K., 2017b. Joint classification and contour extraction of large 3D point clouds. ISPRS J. Photogramm. Remote Sens. 130, 231–245. https://doi.org/10.1016/j.isprsjprs.2017.05.012

Hadsell, R., Sermanet, P., Ben, J., Erkan, A., Scoffier, M., Kavukcuoglu, K., Muller, U., LeCun, Y., 2009. Learning long-range vision for autonomous off-road driving. J. F. Robot. 26, 120–144. https://doi.org/10.1002/rob.20276

Hamilton Jenkin, A.K., 1974. Mines of Devon Volume 1: The Southern Area. David and Charles (Holdings) Limited, Newton Abbot, UK.

Hanan, N.P., Anchang, J.Y., 2020. Satellites could soon map every tree on Earth. Nature 587, 42–43. https://doi.org/10.1038/d41586-020-02830-3

Hanke, K., Hiebel, G., Kovacs, K., Moser, M., 2009. Surveying and Geoinformation - Contributions To an Interdisciplinary Special Research Program on the History of Mining Activities, in: 22nd CIPA Symposium. Kyoto, Japan.

Haslam, R., Howard, A.S., 2017. TELLUS How — an exemplar approach for end-user innovation and follow up of a major NERC environmental survey project.

He, K., Gkioxari, G., Dollar, P., Girshick, R., 2017. Mask R-CNN, in: 2017 IEEE International Conference on Computer Vision (ICCV). IEEE, pp. 2980–2988. https://doi.org/10.1109/ICCV.2017.322

Hecht-Nielsen, R., 1992. Theory of the Backpropagation Neural Network, in: Wechsler, H. (Ed.), Neural Networks for Perception. Academic Press, pp. 65–93.

Hesse, R., 2010. LiDAR-derived Local Relief Models - a new tool for archaeological prospection. Archaeol. Prospect. 17, 67–72. https://doi.org/10.1002/arp.374

Hilson, G., 2002. The future of small-scale mining: environmental and socioeconomic perspectives. Futures 34, 863–872. https://doi.org/10.1016/S0016-3287(02)00044-7

Hilson, G., 2001. A contextual review of the Ghanaian small-scale mining industry. Mining, Miner. Sustain. Dev. 29.

Hilson, G., Gatsinzi, A., 2014. A rocky road ahead? Critical reflections on the futures of small-scale mining in sub-Saharan Africa. Futures 62, 1–9. https://doi.org/10.1016/j.futures.2014.05.006

Hilson, G., Hilson, A., Adu-Darko, E., 2014. Chinese participation in Ghana's informal gold mining economy: Drivers, implications and clarifications. J. Rural Stud. 34, 292–303. https://doi.org/10.1016/j.jrurstud.2014.03.001

Hilson, G., Maconachie, R., 2020. Artisanal and small-scale mining and the Sustainable Development Goals: Opportunities and new directions for sub-Saharan Africa. Geoforum 111, 125–141. https://doi.org/10.1016/j.geoforum.2019.09.006

Hinton, G., Deng, L., Yu, D., Dahl, G., Mohamed, A., Jaitly, N., Senior, A., Vanhoucke, V., Nguyen, P., Sainath, T., Kingsbury, B., 2012. Deep Neural Networks for Acoustic Modeling in Speech Recognition: The Shared Views of Four Research Groups. IEEE Signal Process. Mag. 29, 82–97. https://doi.org/10.1109/MSP.2012.2205597

Hinton, G.E., 2012. A Practical Guide to Training Restricted Boltzmann Machines, in: Montavon, G., Orr, G.B., Müller, K.-R. (Eds.), Neural Networks: Tricks of the Trade. Springer, Berlin, Germany.

Hinton, G.E., Osindero, S., Teh, Y.-W., 2006. A Fast Learning Algorithm for Deep Belief Nets. Neural Comput. 18, 1527–1554. https://doi.org/10.1162/neco.2006.18.7.1527

Historic England, 2018. Using Airborne Lidar in Archaeological Survey: The Light Fantastic. Swindon, UK.

Hoek, E., Brown, E., 1982. Underground Excavations in Rock. Institution of Mining and Metallurgy, Boca Raton, FL, USA.

Hoeser, T., Kuenzer, C., 2020. Object Detection and Image Segmentation with Deep Learning on Earth Observation Data: A Review-Part I: Evolution and Recent Trends. Remote Sens. 12, 1667. https://doi.org/10.3390/rs12101667

Holmes, F., 2019. This AI Company Is The Future Of Gold Exploration [WWW Document]. Forbes. URL https://www.forbes.com/sites/greatspeculations/2019/02/11/this-ai-company-is-the-future-of-gold-exploration/#57693c8b2ae3

Hood, S.B., Cracknell, M.J., Gazley, M.F., Reading, A.M., 2019. Improved supervised classification of bedrock in areas of transported overburden: Applying domain expertise at Kerkasha, Eritrea. Appl. Comput. Geosci. 3–4, 100001. https://doi.org/10.1016/j.acags.2019.100001

Hornik, K., 1991. Approximation capabilities of multilayer feedforward networks. Neural Networks 4, 251–257. https://doi.org/10.1016/0893-6080(91)90009-T

Hough, P.V.C., 1962. Method and means for recognizing complex patterns. United States Patent No. 3069654 A.

Howarth, K., Van Komen, D.F., Neilsen, T.B., Knobles, D.P., Dahl, P.H., Dall'Osto, D.R., 2019. Effect of signal to noise ratio on a convolutional neural network for source ranging and environmental classification. J. Acoust. Soc. Am. 146, 2961–2962. https://doi.org/10.1121/1.5137292

Hu, Q., Yang, B., Xie, L., Rosa, S., Guo, Y., Wang, Z., Trigoni, N., Markham, A., 2020. RandLA-Net: Efficient Semantic Segmentation of Large-Scale Point Clouds,

in: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, pp. 11105–11114. https://doi.org/10.1109/CVPR42600.2020.01112

Hu, W., Wu, L., 2016. Ground deformation extraction using visible images and LiDAR data in mining area. ISPRS - Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. XLI-B7, 505–512. https://doi.org/10.5194/isprs-archives-XLI-B7-505-2016

Huang, C., Li, Y., Loy, C.C., Tang, X., 2016. Learning Deep Representation for Imbalanced Classification, in: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, pp. 5375–5384. https://doi.org/10.1109/CVPR.2016.580

Huang, J., Rathod, V., Sun, C., Zhu, M., Korattikara, A., Fathi, A., Fischer, I., Wojna, Z., Song, Y., Guadarrama, S., Murphy, K., 2017. Speed/Accuracy Trade-Offs for Modern Convolutional Object Detectors, in: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, pp. 3296–3297. https://doi.org/10.1109/CVPR.2017.351

Huang, Z., Fu, H., Chen, W., Zhang, J., Huang, H., 2018. Damage detection and quantitative analysis of shield tunnel structure. Autom. Constr. 94, 303–316. https://doi.org/10.1016/j.autcon.2018.07.006

Hubel, D.H., Wiesel, T.N., 1968. Receptive fields and functional architecture of monkey striate cortex. J. Physiol. 195, 215–243. https://doi.org/10.1113/jphysiol.1968.sp008455

Hubel, D.H., Wiesel, T.N., 1959. Receptive fields of single neurones in the cat's striate cortex. J. Physiol. 148, 574–591. https://doi.org/10.1113/jphysiol.1959.sp006308

Humboldt State University, 2019. Introduction to Remote Sensing [WWW Document]. GSP 216. URL http://gsp.humboldt.edu/OLM/Courses/GSP_216_Online/lesson3-1/bands.html (accessed 6.30.20).

Iglovikov, V., Mushinskiy, S., Osin, V., 2017. Satellite Imagery Feature Detection using Deep Convolutional Neural Network: A Kaggle Competition. arXiv Prepr.

Intergovernmental Forum on Mining Minerals Metals and Sustainable Development, 2017. Global Trends in Artisanal and Small-Scale Mining (ASM): A review of key numbers and issues. Winnipeg, Canada.

Jackisch, R., Lorenz, S., Zimmermann, R., Möckel, R., Gloaguen, R., 2018. Drone-Borne Hyperspectral Monitoring of Acid Mine Drainage: An Example from the Sokolov Lignite District. Remote Sens. 10, 385. https://doi.org/10.3390/rs10030385

Jacobson, N.P., Gupta, M.R., 2005. Design goals and solutions for display of hyperspectral images. IEEE Trans. Geosci. Remote Sens. 43, 2684–2692. https://doi.org/10.1109/TGRS.2005.857623

Jain, V., Seung, H.S., 2008. Natural Image Denoising with Convolutional Networks, in: Proceedings of the 21st International Conference on Neural Information Processing Systems, NIPS'08. Curran Associates Inc., Red Hook, NY, USA, pp. 769–776.

Jensen, J.R., 2007. Remote Sensing of the Environment: An Earth Resource Perspective, 2nd ed. Pearson, London, UK.

Jinfei Wang, Howarth, P.J., 1990. Use of the Hough transform in automated lineament. IEEE Trans. Geosci. Remote Sens. 28, 561–567. https://doi.org/10.1109/TGRS.1990.572949

Justice, C., Townshend, J., 2002. Special issue on the moderate resolution imaging spectroradiometer (MODIS): a new generation of land surface monitoring. Remote Sens. Environ. 83, 1–2. https://doi.org/10.1016/S0034-4257(02)00083-4

Jutzi, B., Gross, H., 2009. Nearest neighbour classification on laser point clouds to gain object structures from buildings, in: ISPRS Hannover Workshop 2009 High-Resolution Earth Imaging for Geospatial Information. ISPRS, Hannover, Germany, p. 6.

Kaelbling, L.P., Littman, M.L., Moore, A.W., 1996. Reinforcement Learning: A Survey. J. Artif. Intell. Res. 4, 237–285. https://doi.org/10.1613/jair.301

Kalman, R.E., 1960. A New Approach to Linear Filtering and Prediction Problems. J. Basic Eng. 82, 35–45. https://doi.org/10.1115/1.3662552

Karan, S.K., Samadder, S.R., Maiti, S.K., 2016. Assessment of the capability of remote sensing and GIS techniques for monitoring reclamation success in coal mine degraded lands. J. Environ. Manage. 182, 272–283. https://doi.org/10.1016/j.jenvman.2016.07.070

Kelly, M., Blanchard, S.D., Kersten, E., Koy, K., 2011. Terrestrial Remotely Sensed Imagery in Support of Public Health: New Avenues of Research Using Object-Based Image Analysis. Remote Sens. 3, 2321–2345. https://doi.org/10.3390/rs3112321

Kemker, R., Salvaggio, C., Kanan, C., 2018. Algorithms for semantic segmentation of multispectral remote sensing imagery using deep learning. ISPRS J. Photogramm. Remote Sens. 145, 60–77. https://doi.org/10.1016/j.isprsjprs.2018.04.014

Kirsch, M., Lorenz, S., Zimmermann, R., Tusa, L., Möckel, R., Hödl, P., Booysen, R., Khodadadzadeh, M., Gloaguen, R., 2018. Integration of Terrestrial and Drone-Borne Hyperspectral and Photogrammetric Sensing Methods for Exploration Mapping and Mining Monitoring. Remote Sens. 10, 1366. https://doi.org/10.3390/rs10091366

Kogut, T., Weistock, M., 2019. Classifying airborne bathymetry data using the Random Forest algorithm. Remote Sens. Lett. 10, 874–882. https://doi.org/10.1080/2150704X.2019.1629710

Kokalj, Ž., Hesse, R., 2017. Airborne laser scanning raster data visualization, Prostor, Kraj, Čas, Prostor, kraj, čas. ZRC SAZU, Založba ZRC, Ljubljana, Slovenia. https://doi.org/10.3986/9789612549848

Kokalj, Ž., Somrak, M., 2019. Why Not a Single Image? Combining Visualizations to Facilitate Fieldwork and On-Screen Mapping. Remote Sens. 11, 747. https://doi.org/10.3390/rs11070747

Konoplich, G. V, Putin, E.O., Filchenkov, A.A., 2016. Application of deep learning to the problem of vehicle detection in UAV images, in: 2016 XIX IEEE International Conference on Soft Computing and Measurements (SCM). IEEE, pp. 4–6. https://doi.org/10.1109/SCM.2016.7519666

Korosec, K., 2019. Uber unveils new Volvo self-driving vehicle in a step toward robotaxi service [WWW Document]. TechCrunch. URL https://techcrunch.com/2019/06/12/uber-unveils-new-volvo-self-driving-vehicle-in-a-step-toward-robotaxi-service/?guccounter=1&guce_referrer=aHR0cHM6Ly93d3cuZ29vZ2xlLmNvbS8&guce_referrer_sig=AQAAANMFgP7arLZO_Y6W5KVkJEe87FH7ESDXvbQoed-FELH3T2TlSSLaQt5BzcCHI

Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. ImageNet classification with deep convolutional neural networks, in: NIPS'12: Proceedings of the 25th International Conference on Neural Information Processing Systems. pp. 1097–1105.

Kroupi, E., Kesa, M., Navarro-Sánchez, V.D., Saeed, S., Pelloquin, C., Alhaddad, B., Moreno, L., Soria-Frisch, A., Ruffini, G., 2019. Deep convolutional neural networks for land-cover classification with Sentinel-2 images. J. Appl. Remote Sens. 13, 1. https://doi.org/10.1117/1.JRS.13.024525

Kumar, A., Sattigeri, P., Fletcher, P.T., 2017. Semi-Supervised Learning with GANs: Manifold Invariance with Improved Inference, in: Proceedings of the 31st International Conference on Neural Information Processing Systems, NIPS'17. Curran Associates Inc., Red Hook, NY, USA, pp. 5540–5550.

Kusimi, J.M., 2008. Assessing land use and land cover change in the Wassa West District of Ghana using remote sensing. GeoJournal 71, 249–259. https://doi.org/10.1007/s10708-008-9172-6

Labou, I., Benoit, M., Baratoux, L., Grégoire, M., Ndiaye, P.M., Thebaud, N., Béziat, D., Debat, P., 2020. Petrological and geochemical study of Birimian ultramafic rocks within the West African Craton: Insights from Mako (Senegal) and Loraboué (Burkina Faso) lherzolite/harzburgite/wehrlite associations. J. African Earth Sci. 162, 103677. https://doi.org/10.1016/j.jafrearsci.2019.103677

Lafferty, J., Zhu, X., Liu, Y., 2004. Kernel conditional random fields, in: Twenty-First International Conference on Machine Learning - ICML '04, ICML '04. ACM Press, New York, New York, USA, p. 64. https://doi.org/10.1145/1015330.1015337

Landrieu, L., Mallet, C., Weinmann, M., 2017a. Comparison of belief propagation and graph-cut approaches for contextual classification of 3D lidar point cloud data, in: 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). IEEE, pp. 2768–2771. https://doi.org/10.1109/IGARSS.2017.8127571

Landrieu, L., Raguet, H., Vallet, B., Mallet, C., Weinmann, M., 2017b. A structured regularization framework for spatially smoothing semantic labelings of 3D point clouds. ISPRS J. Photogramm. Remote Sens. 132, 102–118. https://doi.org/10.1016/j.isprsjprs.2017.08.010

Landrieu, L., Simonovsky, M., 2018. Large-Scale Point Cloud Semantic Segmentation with Superpoint Graphs, in: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, pp. 4558–4567. https://doi.org/10.1109/CVPR.2018.00479

Latifovic, R., Pouliot, D., Campbell, J., 2018. Assessment of Convolution Neural Networks for Surficial Geology Mapping in the South Rae Geological Region, Northwest Territories, Canada. Remote Sens. 10, 307. https://doi.org/10.3390/rs10020307

Lato, M., Diederichs, M.S., Hutchinson, D.J., Harrap, R., 2009. Optimization of LiDAR scanning and processing for automated structural evaluation of discontinuities in rockmasses. Int. J. Rock Mech. Min. Sci. 46, 194–199. https://doi.org/10.1016/j.ijrmms.2008.04.007

LeCun, Y., 2019. Deep Learning Hardware: Past, Present, and Future, in: 2019 IEEE International Solid- State Circuits Conference - (ISSCC). IEEE, pp. 12–19. https://doi.org/10.1109/ISSCC.2019.8662396

LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521, 436–444. https://doi.org/10.1038/nature14539

LeCun, Y., Boser, B., Denker, J.S., Henderson, D., Howard, R.E., Hubbard, W., Jackel, L.D., 1989a. Backpropagation Applied to Handwritten Zip Code Recognition. Neural Comput. 1, 541–551. https://doi.org/10.1162/neco.1989.1.4.541

LeCun, Y., Boser, B., Denker, J.S., Henderson, D., Howard, R.E., Hubbard, W., Jackel, L.D., 1989b. Handwritten digit recognition with a back-propagation network, in: NIPS'89: Proceedings of the 2nd International Conference on Neural Information Processing Systems. pp. 396–404.

Lecun, Y., Bottou, L., Bengio, Y., Haffner, P., 1998. Gradient-based learning applied to document recognition. Proc. IEEE 86, 2278–2324. https://doi.org/10.1109/5.726791

Lee, S., Choi, Y., 2016. Reviews of unmanned aerial vehicle (drone) technology trends and its applications in the mining industry. Geosystem Eng. 19, 197–204. https://doi.org/10.1080/12269328.2016.1162115

Lehtomäki, M., Jaakkola, A., Hyyppä, J., Kukko, A., Kaartinen, H., 2010. Detection of Vertical Pole-Like Objects in a Road Environment Using Vehicle-Based Laser Scanning Data. Remote Sens. 2, 641–664. https://doi.org/10.3390/rs2030641

Lehtomaki, M., Jaakkola, A., Hyyppa, J., Lampinen, J., Kaartinen, H., Kukko, A., Puttonen, E., Hyyppa, H., 2016. Object Classification and Recognition From Mobile Laser Scanning Point Clouds in a Road Environment. IEEE Trans. Geosci. Remote Sens. 54, 1226–1239. https://doi.org/10.1109/TGRS.2015.2476502

Leroux, L., Congedo, L., Bellón, B., Gaetano, R., Bégué, A., 2018. Land Cover Mapping Using Sentinel-2 Images and the Semi-Automatic Classification Plugin: A Northern Burkina Faso Case Study, in: QGIS and Applications in Agriculture and Forest. John Wiley & Sons, Inc., Hoboken, NJ, USA, pp. 119–151. https://doi.org/10.1002/9781119457107.ch4

Li, C.C., 2017. Principles of rockbolting design. J. Rock Mech. Geotech. Eng. 9, 396–414. https://doi.org/10.1016/j.jrmge.2017.04.002

Li, F.-F., Krishna, R., Xu, D., 2020. CS231n: Convolutional Neural Networks for Visual Recognition [WWW Document]. URL http://cs231n.stanford.edu/ (accessed 1.11.21).

Lillesand, T.M., Kiefer, R.W., Chipman, J.W., 2015. Remote Sensing and Image Interpretation, 7th ed. John Wiley & Sons, Hoboken, NJ, USA.

Lin, L.-H., Lawrence, P.D., Hall, R., 2013. Robust outdoor stereo vision SLAM for heavy machine rotation sensing. Mach. Vis. Appl. 24, 205–226. https://doi.org/10.1007/s00138-011-0380-6

Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., Zitnick, C.L., 2014. Microsoft COCO: Common Objects in Context, in: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (Eds.), Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Springer International Publishing, Cham, Switzerland, pp. 740–755. https://doi.org/10.1007/978-3-319-10602-1_48

Linder, W., 2014. Digital Photogrammetry: Theory and Applications. Springer, Berlin, Germany.

Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., Berg, A.C., 2016. SSD: Single Shot MultiBox Detector, in: Leibe, B., Matas, J., Sebe, N., Welling, M. (Eds.), Computer Vision – ECCV 2016. ECCV 2016. Lecture Notes in Computer Science. Springer International Publishing, Cham, Switzerland, pp. 21–37. https://doi.org/10.1007/978-3-319-46448-0_2

Liu, W., Wang, Z., Liu, X., Zeng, N., Liu, Y., Alsaadi, F.E., 2017. A survey of deep neural network architectures and their applications. Neurocomputing 234, 11–26. https://doi.org/10.1016/j.neucom.2016.12.038

Liu, X., Wang, H., Jing, H., Shao, A., Wang, L., 2020. Research on Intelligent Identification of Rock Types Based on Faster R-CNN Method. IEEE Access 8, 21804–21812. https://doi.org/10.1109/ACCESS.2020.2968515

Lobo, F. de L., Souza-Filho, P.W.M., Novo, E.M.L. de M., Carlos, F.M., Barbosa, C.C.F., 2018. Mapping Mining Areas in the Brazilian Amazon Using MSI/Sentinel-2 Imagery (2017). Remote Sens. 10, 1178. https://doi.org/10.3390/rs10081178

Long, N.Q., Buczek, M.M., Hien, L.P., Szlapińska, S.A., Nam, B.X., Nghia, N.V., Cuong, C.X., 2018. Accuracy assessment of mine walls' surface models derived from terrestrial laser scanning. Int. J. Coal Sci. Technol. 5, 328–338. https://doi.org/10.1007/s40789-018-0218-1

Luhmann, T., Robson, S., Kyle, S., Boehm, J., 2014. Close-Range Photogrammetry and 3D Imaging, 2nd ed. Walter de Gruyter, Berlin, Germany.

Maennling, N., Toledano, P., 2019. Seven trends shaping the future of the mining and metals industry [WWW Document]. World Econ. Forum. URL https://www.weforum.org/agenda/2019/03/seven-trends-shaping-the-future-of-the-mining-and-metals-sector/ (accessed 1.11.21).

Majurski, M., Manescu, P., Padi, S., Schaub, N., Hotaling, N., Simon, C., Bajcsy, P., 2019. Cell Image Segmentation Using Generative Adversarial Networks, Transfer Learning, and Augmentations, in: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). IEEE, pp. 1114–1122. https://doi.org/10.1109/CVPRW.2019.00145

Mallast, U., Gloaguen, R., Geyer, S., Rödiger, T., Siebert, C., 2011. Semi-automatic extraction of lineaments from remote sensing data and the derivation of groundwater flow-paths. Hydrol. Earth Syst. Sci. Discuss. 8, 1399–1431. https://doi.org/10.5194/hessd-8-1399-2011

Marshall, J.A., Bonchis, A., Nebot, E., Scheding, S., 2016. Robotics in Mining, in: Siciliano, B., Khatib, O. (Eds.), Springer Handbook of Robotics. Springer International Publishing, Cham, Switzerland, pp. 1549–1572.

Martin, P.G., Payton, O.D., Fardoulis, J.S., Richards, D.A., Scott, T.B., 2015. The use of unmanned aerial systems for the mapping of legacy uranium mines. J. Environ. Radioact. 143, 135–140. https://doi.org/10.1016/j.jenvrad.2015.02.004

Martínez-Sánchez, J., Puente, I., GonzálezJorge, H., Riveiro, B., Arias, P., 2016. Automatic thickness and volume estimation of sprayed concrete on anchored retaining walls from terrestrial LiDAR data. ISPRS - Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. XLI-B5, 521–526. https://doi.org/10.5194/isprsarchives-XLI-B5-521-2016

Masoud, A., Koike, K., 2017. Applicability of computer-aided comprehensive tool (LINDA: LINeament Detection and Analysis) and shaded digital elevation model for

characterizing and interpreting morphotectonic features from lineaments. Comput. Geosci. 106, 89–100. https://doi.org/10.1016/j.cageo.2017.06.006

MathWorks, 2019. Semantic Segmentation of Multispectral Images Using Deep Learning [WWW Document]. R2019b Doc. Examples. URL https://uk.mathworks.com/help/images/multispectral-semantic-segmentation-using-deep-learning.html (accessed 1.11.21).

Maturana, D., Scherer, S., 2015. VoxNet: A 3D Convolutional Neural Network for real-time object recognition, in: 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, pp. 922–928. https://doi.org/10.1109/IROS.2015.7353481

Maxwell, A.E., Pourmohammadi, P., Poyner, J.D., 2020. Mapping the Topographic Features of Mining-Related Valley Fills Using Mask R-CNN Deep Learning and Digital Elevation Data. Remote Sens. 12, 547. https://doi.org/10.3390/rs12030547

Maxwell, A.E., Warner, T.A., Strager, M.P., Pal, M., 2014. Combining RapidEye Satellite Imagery and Lidar for Mapping of Mining and Mine Reclamation. Photogramm. Eng. Remote Sens. 80, 179–189. https://doi.org/10.14358/PERS.80.2.179-189

McCulloch, W.S., Pitts, W., 1943. A logical calculus of the ideas immanent in nervous activity. Bull. Math. Biophys. 5, 115–133. https://doi.org/10.1007/BF02478259

Mcdonald, J., 2013. New Data and Techniques for Evaluating Subsidence from Abandoned Underground Mines in Ohio, in: American Association of Petroleum Geologists Annual Conference and Exhibition. Pittsburgh, PA, USA.

Mcdonald, J., 2011. Mapping Abandoned Mines Using Imagery and LiDAR from the Ohio Statewide Imagery Program, in: Digital Mapping Techniques '11Affiliation: Association of American State Geologists and the U.S. Geological Survey.

McGlone, J.C., Mikhail, E.M., Bethel, J.S., Mullen, R., 2004. Manual of photogrammetry.

McLeod, T., Samson, C., Labrie, M., Shehata, K., Mah, J., Lai, P., Wang, L., Elder, J.H., 2013. Using Video Acquired from an Unmanned Aerial Vehicle (UAV) to Measure Fracture Orientation in an Open-Pit Mine. Geomatica 67, 173–180. https://doi.org/10.5623/cig2013-036

McQuilken, J., Garvin, H., 2016. Artisanal and small-scale gold mining in Ghana. Evidence to inform an 'action dialogue,' International Institute for Environment and Development. London, UK.

Mcquillan, A., 2013. Comparison of photogrammetry and survey laser scanning output data for use in mapping joints in open cut highwalls, in: 13th Coal Operators' Conference. The Australasian Institute of Mining and Metallurgy & Mine Managers Association of Australia, Wollongong, Australia, pp. 347–354.

Middleton, M., Schnur, T., Sorjonen-Ward, P., Hyvönen, E., 2015. Geological lineament interpretation using the object-based image analysis approach: results of semi-automated analyses versus visual interpretation, in: Novel Technologies for Greenfield Exploration. Geological Survey of Finland, pp. 135–154.

Ministry of Lands and Natural Resources, 2019. Government committed to the fight against illegal mining [WWW Document]. Online Artic. URL http://mlnr.gov.gh/index.php/government-committed-to-the-fight-against-illegal-mining/ (accessed 2.9.20).

Minsky, M., Papert, S., 1969. Perceptrons., Perceptrons. M.I.T. Press, Oxford, England.

Monsalve, J.J., Baggett, J., Bishop, R., Ripepi, N., 2019. Application of laser scanning for rock mass characterization and discrete fracture network generation in an underground limestone mine. Int. J. Min. Sci. Technol. 29, 131–137. https://doi.org/10.1016/j.ijmst.2018.11.009

Moore, K., 2019. Goldcorp and IBM pioneer orebody discovery and predictability solution [WWW Document]. MiningWorld Mag. URL http://miningworld.com/goldcorp-and-ibm-pioneer-orebody-discovery-and-predictability-solution/ (accessed 1.12.21).

Moraes, R., Valiati, J.F., Gavião Neto, W.P., 2013. Document-level sentiment classification: An empirical comparison between SVM and ANN. Expert Syst. Appl. 40, 621–633. https://doi.org/10.1016/j.eswa.2012.07.059

Moudrý, V., Gdulová, K., Fogl, M., Klápště, P., Urban, R., Komárek, J., Moudrá, L., Štroner, M., Barták, V., Solský, M., 2019. Comparison of leaf-off and leaf-on combined UAV imagery and airborne LiDAR for assessment of a post-mining site terrain and vegetation structure: Prospects for monitoring hazards and restoration success. Appl. Geogr. 104, 32–41. https://doi.org/10.1016/j.apgeog.2019.02.002

MXNet, 2020. Data Augmentation with Masks [WWW Document]. MXNet Tutorials. URL https://mxnet.apache.org/versions/1.2.1/tutorials/python/data_augmentation_with_masks.html (accessed 9.1.20).

Munappy, A., Bosch, J., Olsson, H.H., Arpteg, A., Brinne, B., 2019. Data Management Challenges for Deep Learning, in: 2019 45th Euromicro Conference on Software Engineering and Advanced Applications (SEAA). pp. 140–147. https://doi.org/10.1109/SEAA.2019.00030

Naprstek, T., 2020. New Methods for the Interpolation and Interpretation of Lineaments in Aeromagnetic Data. Laurentian Université.

Newman, P., 2010. Environment, Antecedent and Adventure: Tin and Copper Mining on Dartmoor, Devon, c.1700-1914. University of Leicester.

Ng, A., 2017. Improving Deep Neural Networks: Hyperparameter tuning, Regularization and Optimization [WWW Document]. URL https://www.coursera.org/learn/deep-neural-network/home/welcome (accessed 1.12.21).

Nielsen, M.A., 2015. Neural Networks and Deep Learning. Determination Press.

Niemeyer, J., Rottensteiner, F., Soergel, U., 2014. Contextual classification of lidar data and building object detection in urban areas. ISPRS J. Photogramm. Remote Sens. 87, 152–165. https://doi.org/10.1016/j.isprsjprs.2013.11.001

Nogueira, K., Penatti, O.A.B., dos Santos, J.A., 2017. Towards better exploiting convolutional neural networks for remote sensing scene classification. Pattern Recognit. 61, 539–556. https://doi.org/10.1016/j.patcog.2016.07.001

Northern Mine Research Society, 1980. British Mining Vol 13: The Mines of Grassington Moor. Northern Mine Research Society, Sheffield, UK.

Nuijten, R.J.G., Kooistra, L., De Deyn, G.B., 2019. Using Unmanned Aerial Systems (UAS) and Object-Based Image Analysis (OBIA) for Measuring Plant-Soil Feedback Effects on Crop Productivity. Drones 3, 54. https://doi.org/10.3390/drones3030054

Nygren, P., Jasinski, M., 2016. A Comparative Study of Segmentation and Classification Methods for 3D Point Clouds. University of Gothenburg.

O'Leary, D.W., Friedman, J.D., Pohn, H.A., 1976. Lineament, linear, lineation: Some proposed new standards for old terms. Geol. Soc. Am. Bull. 87, 1463.

Öberg, F., 2013. Method and system for monitoring and documenting installation of rock reinforcement bolt. United States Patent No. 8606542 B2.

Obodai, J., Adjei, K.A., Odai, S.N., Lumor, M., 2019. Land use/land cover dynamics using landsat data in a gold mining basin-the Ankobra, Ghana. Remote Sens. Appl. Soc. Environ. 13, 247–256. https://doi.org/10.1016/j.rsase.2018.10.007

Olofsson, P., Foody, G.M., Herold, M., Stehman, S. V., Woodcock, C.E., Wulder, M.A., 2014. Good practices for estimating area and assessing accuracy of land change. Remote Sens. Environ. 148, 42–57. https://doi.org/10.1016/j.rse.2014.02.015

Onederra, I., Thurley, M.J., Catalan, A., 2015a. Measuring blast fragmentation at Esperanza mine using high-resolution 3D laser scanning. Min. Technol. 124, 34–36. https://doi.org/10.1179/1743286314Y.0000000076

Onederra, I., Thurley, M.J., Catalan, A., 2015b. Measuring blast fragmentation at Esperanza mine using high-resolution 3D laser scanning. Min. Technol. 124, 34–36. https://doi.org/10.1179/1743286314Y.0000000076

Owusu-Nimo, F., Mantey, J., Nyarko, K.B., Appiah-Effah, E., Aubynn, A., 2018. Spatial distribution patterns of illegal artisanal small scale gold mining (Galamsey) operations in Ghana: A focus on the Western Region. Heliyon 4, e00534. https://doi.org/10.1016/j.heliyon.2018.e00534

Owusu, K., Emmanuel, A.K., Musah-Surugu, I.J., Yankson, P.W.K., 2019. The effects of 2015 El Nino on smallholder maize production in the transitional ecological zone of Ghana. Int. J. Clim. Chang. Strateg. Manag. 11, 609–621. https://doi.org/10.1108/IJCCSM-02-2018-0014

Padró, J.-C., Carabassa, V., Balagué, J., Brotons, L., Alcañiz, J.M., Pons, X., 2019. Monitoring opencast mine restorations using Unmanned Aerial System (UAS) imagery. Sci. Total Environ. 657, 1602–1614. https://doi.org/10.1016/j.scitotenv.2018.12.156

Pal, M., 2005. Random forest classifier for remote sensing classification. Int. J. Remote Sens. 26, 217–222. https://doi.org/10.1080/01431160412331269698

Palafox, L.F., Hamilton, C.W., Scheidt, S.P., Alvarez, A.M., 2017. Automated detection of geological landforms on Mars using Convolutional Neural Networks. Comput. Geosci. 101, 48–56. https://doi.org/10.1016/j.cageo.2016.12.015

Palamara, D.R., Nicholson, M., Flentje, P., Baafi, E., Brassington, G.M., 2007. An evaluation of airborne laser scan data for coalmine subsidence mapping. Int. J. Remote Sens. 28, 3181–3203. https://doi.org/10.1080/01431160600993439

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., 2019. PyTorch: An Imperative Style, High-Performance Deep Learning Library, in: Advances in Neural Information Processing Systems 32. pp. 8024–8035.

Pauly, M., Keiser, R., Gross, M., 2003. Multi-scale Feature Extraction on Point-Sampled Surfaces. Comput. Graph. Forum 22, 281–289. https://doi.org/10.1111/1467-8659.00675

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A.,

Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E., 2011. Scikit-learn: Machine Learning in Python. J. Mach. Learn. Res. 12, 2825–2830.

Pelletier, C., Valero, S., Inglada, J., Champion, N., Marais Sicre, C., Dedieu, G., 2017. Effect of Training Class Label Noise on Classification Performances for Land Cover Mapping with Satellite Image Time Series. Remote Sens. 9, 173. https://doi.org/10.3390/rs9020173

Peng, D., Zhang, Y., Guan, H., 2019. End-to-End Change Detection for High Resolution Satellite Images Using Improved UNet++. Remote Sens. 11, 1382. https://doi.org/10.3390/rs11111382

Pires de Lima, R., Duarte, D., Nicholson, C., Slatt, R., Marfurt, K.J., 2020. Petrographic microfacies classification with deep convolutional neural networks. Comput. Geosci. 142, 104481. https://doi.org/10.1016/j.cageo.2020.104481

Price, R., Cornelius, M., Burnside, L., Miller, B., 2020. Mine Planning and Selection of Autonomous Trucks, in: Topal, E. (Ed.), 28th International Symposium on Mine Planning and Equipment Selection - MPES 2019. Springer International Publishing, Cham, Switzerland, pp. 203–212. https://doi.org/10.1007/978-3-030-33954-8_26

Prno, J., Scott Slocombe, D., 2012. Exploring the origins of 'social license to operate' in the mining sector: Perspectives from governance and sustainability theories. Resour. Policy 37, 346–357. https://doi.org/10.1016/j.resourpol.2012.04.002

Procopio, M.J., Mulligan, J., Grudic, G., 2009. Learning terrain segmentation with classifier ensembles for autonomous robot navigation in unstructured environments. J. F. Robot. 26, 145–175. https://doi.org/10.1002/rob.20279

Puente, I., González-Jorge, H., Martínez-Sánchez, J., Arias, P., 2013. Review of mobile mapping and surveying technologies. Measurement 46, 2127–2145. https://doi.org/10.1016/j.measurement.2013.03.006

Python 3.6.7, 2019.

Qi, C.R., Yi, L., Su, H., Guibas, L.J., 2017. PointNet++: Deep hierarchical feature learning on point sets in a metric space, in: Luxburg, U. Von, Guyon, I.M., Bengio,

S., Wallach, H.M., Fergus, R. (Eds.), NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems. Curran Associates, Red Hook, NY, USA, pp. 5105–5114.

Rahnama, M., Gloaguen, R., 2014. TecLines: A MATLAB-Based Toolbox for Tectonic Lineament Analysis from Satellite Images and DEMs, Part 2: Line Segments Linking and Merging. Remote Sens. 6, 11468–11493. https://doi.org/10.3390/rs61111468

Raina, R., Madhavan, A., Ng, A.Y., 2009. Large-scale deep unsupervised learning using graphics processors, in: 26th Annual International Conference on Machine Learning - ICML '09. ACM Press, New York, New York, USA, pp. 1–8. https://doi.org/10.1145/1553374.1553486

Rajaee, M., Obiri, S., Green, A., Long, R., Cobbina, S., Nartey, V., Buck, D., Antwi, E., Basu, N., 2015. Integrated Assessment of Artisanal and Small-Scale Gold Mining in Ghana—Part 2: Natural Sciences Review. Int. J. Environ. Res. Public Health 12, 8971–9011. https://doi.org/10.3390/ijerph120808971

Ramezani, M., Nouranian, S., Bell, I., Sameti, B., Tafazoli, S., 2017. Fast Rock Segmentation Using Artificial Intelligence to Approach Human-Level Accuracy. J. Explos. Eng.

Rau, J.Y., Jhan, J.P., Hsu, Y.C., 2015. Analysis of oblique aerial images for land cover and point cloud classification in an Urban environment. IEEE Trans. Geosci. Remote Sens. 53, 1304–1319. https://doi.org/10.1109/TGRS.2014.2337658

Rauhala, A., Tuomela, A., Davids, C., Rossi, P., 2017. UAV Remote Sensing Surveillance of a Mine Tailings Impoundment in Sub-Arctic Conditions. Remote Sens. 9, 1318. https://doi.org/10.3390/rs9121318

Razavian, A.S., Azizpour, H., Sullivan, J., Carlsson, S., 2014. CNN Features Off-the-Shelf: An Astounding Baseline for Recognition, in: 2014 IEEE Conference on Computer Vision and Pattern Recognition Workshops. IEEE, pp. 512–519. https://doi.org/10.1109/CVPRW.2014.131

Redmon, J., Divvala, S., Girshick, R., Farhadi, A., 2016. You Only Look Once: Unified, Real-Time Object Detection, in: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, pp. 779–788. https://doi.org/10.1109/CVPR.2016.91

Redmon, J., Farhadi, A., 2017. YOLO9000: Better, Faster, Stronger, in: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, pp. 6517–6525. https://doi.org/10.1109/CVPR.2017.690

Relief Visualisation Toolbox (RVT), n.d.

Ren, S., He, K., Girshick, R., Sun, J., 2017. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. IEEE Trans. Pattern Anal. Mach. Intell. 39, 1137–1149. https://doi.org/10.1109/TPAMI.2016.2577031

Ren, Y., Zhu, C., Xiao, S., 2018. Small Object Detection in Optical Remote Sensing Images via Modified Faster R-CNN. Appl. Sci. 8, 813. https://doi.org/10.3390/app8050813

Richardson, P.H.G., 1992. British Mining Vol. 44: Mines of Dartmoor and the Tamar Valley. Sheffield, UK.

Riegler, G., Ulusoy, A.O., Geiger, A., 2017. OctNet: Learning Deep 3D Representations at High Resolutions, in: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, pp. 6620–6629. https://doi.org/10.1109/CVPR.2017.701

Rodner, E., Simon, M., Fisher, R., Denzler, J., 2016. Fine-grained Recognition in the Noisy Wild: Sensitivity Analysis of Convolutional Neural Networks Approaches, in: Procedings of the British Machine Vision Conference 2016. British Machine Vision Association, pp. 60.1-60.13. https://doi.org/10.5244/C.30.60

Rodriguez-Galiano, V., Sanchez-Castillo, M., Chica-Olmo, M., Chica-Rivas, M., 2015. Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines. Ore Geol. Rev. 71, 804–818. https://doi.org/10.1016/j.oregeorev.2015.01.001

Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation, in: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (Eds.), Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. Springer International Publishing, Cham, Switzerland, pp. 234–241. https://doi.org/10.1007/978-3-319-24574-4_28

Rosenblatt, F., 1958. The perceptron: A probabilistic model for information storage and organization in the brain. Psychol. Rev. 65, 386–408. https://doi.org/10.1037/h0042519

Rosser, N., Lim, M., Petley, D., Dunning, S., Allison, R., 2007. Patterns of precursory rockfall prior to slope failure. J. Geophys. Res. 112, F04014. https://doi.org/10.1029/2006JF000642

Rowe, T.P., Foster, C.L.N., 1887. Observations on Balleswidden Mine. Trans. R. Geol. Soc. Cornwall x.

Ruff, T.M., 2004. Advances in proximity detection technologies for surface mining equipment, in: 34th Annual Institute on Mining Health, Health, Safety and Research. Salt Lake City, UT, USA.

Rumelhart, D.E., Hinton, G.E., Williams, R.J., 1986. Learning representations by back-propagating errors. Nature 323, 533–536. https://doi.org/10.1038/323533a0

Rumelhart, D.E., McClelland, J.L., 1986. Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Vol. 1: Foundations. MIT Press, Cambridge, MA, USA.

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A.C., Fei-Fei, L., 2015. ImageNet Large Scale Visual Recognition Challenge. Int. J. Comput. Vis. 115, 211–252. https://doi.org/10.1007/s11263-015-0816-y

Rusu, R.B., 2010. Semantic 3D Object Maps for Everyday Manipulation in Human Living Environments. KI - Künstliche Intelligenz 24, 345–348. https://doi.org/10.1007/s13218-010-0059-6

Rusu, R.B., Blodow, N., Beetz, M., 2009. Fast Point Feature Histograms (FPFH) for 3D registration, in: 2009 IEEE International Conference on Robotics and Automation. IEEE, pp. 3212–3217. https://doi.org/10.1109/ROBOT.2009.5152473

Rusu, R.B., Cousins, S., 2011. 3D is here: Point Cloud Library (PCL), in: 2011 IEEE International Conference on Robotics and Automation. IEEE, pp. 1–4. https://doi.org/10.1109/ICRA.2011.5980567

Samuel, A.L., 1959. Some Studies in Machine Learning Using the Game of Checkers. IBM J. Res. Dev. 3, 210–229. https://doi.org/10.1147/rd.33.0210

Sapling Learning, 2015. Electromagnetic Spectrum [WWW Document]. Chempendix. URL https://sites.google.com/site/chempendix/em-spectrum (accessed 8.5.20).

Sayab, M., Aerden, D., Paananen, M., Saarela, P., 2018. Virtual Structural Analysis of Jokisivu Open Pit Using 'Structure-from-Motion' Unmanned Aerial Vehicles (UAV) Photogrammetry: Implications for Structurally-Controlled Gold Deposits in Southwest Finland. Remote Sens. 10, 1296. https://doi.org/10.3390/rs10081296

Schach, R., Garshol, K., Heltzen, A.M., 1979. Rock Bolting - A Practical Handbook. Pergamon Press, Oxford.

Schubert, E., Sander, J., Ester, M., Kriegel, H.P., Xu, X., 2017. DBSCAN Revisited, Revisited. ACM Trans. Database Syst. 42, 1–21. https://doi.org/10.1145/3068335

Schuegraf, P., Bittner, K., 2019. Automatic Building Footprint Extraction from Multi-Resolution Remote Sensing Images Using a Hybrid FCN. ISPRS Int. J. Geo-Information 8, 191. https://doi.org/10.3390/ijgi8040191

Scrivener, R.C., 2006. Cornubian granites and mineralization of SW England, in: The Geology of England and Wales. The Geological Society of London, pp. 257–267. https://doi.org/10.1144/GOEWP.11

Sevara, C., Pregesbauer, M., Doneus, M., Verhoeven, G., Trinks, I., 2016. Pixel versus object — A comparison of strategies for the semi-automated mapping of archaeological features using airborne laser scanning data. J. Archaeol. Sci. Reports 5, 485–498. https://doi.org/10.1016/j.jasrep.2015.12.023

Shao, X., Zhu, H., Guo, D., Zheng, R., Wei, J., 2020. Research on Detection of Large Coal Blockage at the Transfer Point of Belt Conveyor Based on Improved Mask R-CNN. IOP Conf. Ser. Earth Environ. Sci. 440, 052028. https://doi.org/10.1088/1755-1315/440/5/052028

Shariati, H., Yeraliyev, A., Terai, B., Tafazoli, S., Ramezani, M., 2019. Towards autonomous mining via intelligent excavators. 2019 IEEE Conf. Comput. Vis. Pattern Recognit. 26–32.

Shelhamer, E., Long, J., Darrell, T., 2017. Fully Convolutional Networks for Semantic Segmentation. IEEE Trans. Pattern Anal. Mach. Intell. 39, 640–651. https://doi.org/10.1109/TPAMI.2016.2572683

Shendryk, Y., Rist, Y., Ticehurst, C., Thorburn, P., 2019. Deep learning for multi-modal classification of cloud, shadow and land cover scenes in PlanetScope and Sentinel-2 imagery. ISPRS J. Photogramm. Remote Sens. 157, 124–136. https://doi.org/10.1016/j.isprsjprs.2019.08.018

Shorten, C., Khoshgoftaar, T.M., 2019. A survey on Image Data Augmentation for Deep Learning. J. Big Data 6, 60. https://doi.org/10.1186/s40537-019-0197-0

Signoroni, A., Savardi, M., Baronio, A., Benini, S., 2019. Deep Learning Meets Hyperspectral Image Analysis: A Multidisciplinary Review. J. Imaging 5, 52. https://doi.org/10.3390/jimaging5050052

Silburt, A., Ali-Dib, M., Zhu, C., Jackson, A., Valencia, D., Kissin, Y., Tamayo, D., Menou, K., 2019. Lunar crater identification via deep learning. Icarus 317, 27–38. https://doi.org/10.1016/j.icarus.2018.06.022

Šilhavý, J., Minár, J., Mentlík, P., Sládek, J., 2016. A new artefacts resistant method for automatic lineament extraction using Multi-Hillshade Hierarchic Clustering (MHHC). Comput. Geosci. 92, 9–20. https://doi.org/10.1016/j.cageo.2016.03.015

Simic Milas, A., Cracknell, A.P., Warner, T.A., 2018. Drones – the third generation source of remote sensing data. Int. J. Remote Sens. 39, 7125–7137. https://doi.org/10.1080/01431161.2018.1523832

Simons, B., Shail, R.K., Andersen, J.C.Ø., 2016. The petrogenesis of the Early Permian Variscan granites of the Cornubian Batholith: Lower plate post-collisional peraluminous magmatism in the Rhenohercynian Zone of SW England. Lithos 260, 76–94. https://doi.org/10.1016/j.lithos.2016.05.010

Snapir, B., Simms, D.M., Waine, T.W., 2017. Mapping the expansion of galamsey gold mines in the cocoa growing area of Ghana using optical remote sensing. Int. J. Appl. Earth Obs. Geoinf. 58, 225–233. https://doi.org/10.1016/j.jag.2017.02.009

Soilán, M., Riveiro, B., Martínez-Sánchez, J., Arias, P., 2017. Segmentation and classification of road markings using MLS data. ISPRS J. Photogramm. Remote Sens. 123, 94–103. https://doi.org/10.1016/j.isprsjprs.2016.11.011

Soilán, Sánchez-Rodríguez, Río-Barral, Perez-Collazo, Arias, Riveiro, 2019. Review of Laser Scanning Technologies and Their Applications for Road and Railway Infrastructure Monitoring. Infrastructures 4, 58. https://doi.org/10.3390/infrastructures4040058

Somua-Gyimah, G., Frimpong, S., Nyaaba, W., Gbadam, E., 2019. A computer vision system for terrain recognition and object detection tasks in mining and construction environments, in: 2019 SME Annual Conference and Expo and CMA 121st National Western Mining Conference. Society for Mining, Metallurgy and Exploration (SME), Denver, CO, USA.

Soto-Pinto, C., Arellano-Baeza, A., Sánchez, G., 2013. A new code for automatic detection and analysis of the lineament patterns for geophysical and geological purposes (ADALGEO). Comput. Geosci. 57, 93–103. https://doi.org/10.1016/j.cageo.2013.03.019

Souza-Filho, P., Nascimento, W., Santos, D., Weber, E., Silva, R., Siqueira, J., 2018. A GEOBIA Approach for Multitemporal Land-Cover and Land-Use Change Analysis in a Tropical Watershed in the Southeastern Amazon. Remote Sens. 10, 1683. https://doi.org/10.3390/rs10111683

Srivastava, N., Hinton, G.E., Krizhevsky, A., Sutskever, I., Salakhutdinov, R., 2014. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. J. Mach. Learn. Res. 15, 1929–1958.

Stewart, P., 2018. Automated, Real Time Processing, Analysis, Mapping and Reporting of Data for the Detection of Geotechnical Features. Australia Patent No. 201180 A1.

Strobl, C., Boulesteix, A.-L., Kneib, T., Augustin, T., Zeileis, A., 2008. Conditional variable importance for random forests. BMC Bioinformatics 9, 307. https://doi.org/10.1186/1471-2105-9-307

Sturzenegger, M., Stead, D., 2009. Close-range terrestrial digital photogrammetry and terrestrial laser scanning for discontinuity characterization on rock cuts. Eng. Geol. 106, 163–182. https://doi.org/10.1016/j.enggeo.2009.03.004

Su, H., Maji, S., Kalogerakis, E., Learned-Miller, E., 2015. Multi-view Convolutional Neural Networks for 3D Shape Recognition. https://doi.org/10.1109/ICCV.2015.114

Sun, T., Li, H., Wu, K., Chen, F., Zhu, Z., Hu, Z., 2020. Data-Driven Predictive Modelling of Mineral Prospectivity Using Machine Learning and Deep Learning Methods: A Case Study from Southern Jiangxi Province, China. Minerals 10, 102. https://doi.org/10.3390/min10020102

Sun, Y., Zhang, X., Xin, Q., Huang, J., 2018. Developing a multi-filter convolutional neural network for semantic segmentation using high-resolution aerial imagery and LiDAR data. ISPRS J. Photogramm. Remote Sens. 143, 3–14. https://doi.org/10.1016/j.isprsjprs.2018.06.005

Sunderhauf, N., Shirazi, S., Dayoub, F., Upcroft, B., Milford, M., 2015. On the performance of ConvNet features for place recognition, in: 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, pp. 4297–4304. https://doi.org/10.1109/IROS.2015.7353986

Szegedy, C., Wei Liu, Yangqing Jia, Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A., 2015. Going deeper with convolutions, in: 2015

IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 1–9. https://doi.org/10.1109/CVPR.2015.7298594

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z., 2016. Rethinking the Inception Architecture for Computer Vision, in: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, pp. 2818–2826. https://doi.org/10.1109/CVPR.2016.308

Tajbakhsh, N., Jeyaseelan, L., Li, Q., Chiang, J.N., Wu, Z., Ding, X., 2020. Embracing imperfect datasets: A review of deep learning solutions for medical image segmentation. Med. Image Anal. 63, 101693. https://doi.org/10.1016/j.media.2020.101693

Tan, K., Cheng, X., Ju, Q., Wu, S., 2016. Correction of Mobile TLS Intensity Data for Water Leakage Spots Detection in Metro Tunnels. IEEE Geosci. Remote Sens. Lett. 13, 1711–1715. https://doi.org/10.1109/LGRS.2016.2605158

Tang, J., Deng, C., Huang, G.-B., Zhao, B., 2015. Compressed-Domain Ship Detection on Spaceborne Optical Image Using Deep Neural Network and Extreme Learning Machine. IEEE Trans. Geosci. Remote Sens. 53, 1174–1185. https://doi.org/10.1109/TGRS.2014.2335751

Taravat, A., Del Frate, F., Cornaro, C., Vergari, S., 2015. Neural Networks and Support Vector Machine Algorithms for Automatic Cloud Classification of Whole-Sky Ground-Based Images. IEEE Geosci. Remote Sens. Lett. 12, 666–670. https://doi.org/10.1109/LGRS.2014.2356616

Tchapmi, L., Choy, C., Armeni, I., Gwak, J., Savarese, S., 2017. SEGCloud: Semantic Segmentation of 3D Point Clouds, in: 2017 International Conference on 3D Vision (3DV). IEEE, pp. 537–547. https://doi.org/10.1109/3DV.2017.00067

Teledyne Optech, 2020. Lynx Camera Solutions [WWW Document]. URL https://www.teledyneoptech.com/en/products/mobile-survey/lynx-cameras/ (accessed 8.5.20).

Telmer, K., Stapper, D., 2007. Evaluating and Monitoring Small Scale Gold Mining and Mercury Use: Building a Knowledge-base with Satelite Imagery and Field Work. Victoria, Canada.

The MathWorks, 2020. MATLAB Deep Learning Toolkit.

Thomas, H., Qi, C.R., Deschaud, J.-E., Marcotegui, B., Goulette, F., Guibas, L., 2019. KPConv: Flexible and Deformable Convolution for Point Clouds, in: 2019 IEEE/CVF International Conference on Computer Vision (ICCV). IEEE, pp. 6410–6419. https://doi.org/10.1109/ICCV.2019.00651

Thurley, M.J., 2013. Automated Image Segmentation and Analysis of Rock Piles in an Open-Pit Mine, in: 2013 International Conference on Digital Image Computing: Techniques and Applications (DICTA). IEEE, pp. 1–8. https://doi.org/10.1109/DICTA.2013.6691484

Thurley, M.J., Ng, K.C., 2008. Identification and sizing of the entirely visible rocks from a 3D surface data segmentation of laboratory rock piles. Comput. Vis. Image Underst. 111, 170–178. https://doi.org/10.1016/j.cviu.2007.09.009

Tien Bui, D., Long, N.Q., Bui, X.-N., Nguyen, V.-N., Van Pham, C., Van Le, C., Ngo, P.-T.T., Bui, D.T., Kristoffersen, B., 2018. Lightweight Unmanned Aerial Vehicle and Structure-from-Motion Photogrammetry for Generating Digital Surface Model for Open-Pit Coal Mine Area and Its Accuracy Assessment, in: Tien Bui, D., Ngoc Do, A., Bui, H.-B., Hoang, N.-D. (Eds.), Advances and Applications in Geospatial Technology and Earth Resources. Springer International Publishing, Cham, Switzerland, pp. 17–33. https://doi.org/10.1007/978-3-319-68240-2_2

Tien Bui, D., Tuan, T.A., Klempe, H., Pradhan, B., Revhaug, I., 2016. Spatial prediction models for shallow landslide hazards: a comparative assessment of the efficacy of support vector machines, artificial neural networks, kernel logistic regression, and logistic model tree. Landslides 13, 361–378. https://doi.org/10.1007/s10346-015-0557-6

Titos, M., Bueno, A., Garcia, L., Benitez, C., 2018. A Deep Neural Networks Approach to Automatic Recognition Systems for Volcano-Seismic Events. IEEE J.

Sel. Top. Appl. Earth Obs. Remote Sens. 11, 1533–1544. https://doi.org/10.1109/JSTARS.2018.2803198

Tomè, D., Monti, F., Baroffio, L., Bondi, L., Tagliasacchi, M., Tubaro, S., 2016. Deep Convolutional Neural Networks for pedestrian detection. Signal Process. Image Commun. 47, 482–489. https://doi.org/10.1016/j.image.2016.05.007

Tong, X., Liu, Xiangfeng, Chen, P., Liu, Shijie, Luan, K., Li, L., Liu, Shuang, Liu, Xianglei, Xie, H., Jin, Y., Hong, Z., 2015. Integration of UAV-Based Photogrammetry and Terrestrial Laser Scanning for the Three-Dimensional Mapping and Monitoring of Open-Pit Mine Areas. Remote Sens. 7, 6635–6662. https://doi.org/10.3390/rs70606635

Toutin, T., Cheng, P., 2002. QuickBird – A Milestone for High Resolution Mapping. Earth Obs. Mag. 11, 14–18.

Triebel, R., Kersting, K., Burgard, W., 2006. Robust 3D scan point classification using associative Markov networks, in: 2006 IEEE International Conference on Robotics and Automation, 2006. ICRA 2006. IEEE, pp. 2603–2608. https://doi.org/10.1109/ROBOT.2006.1642094

Trier, Ø.D., Cowley, D.C., Waldeland, A.U., 2019. Using deep neural networks on airborne laser scanning data: Results from a case study of semi-automatic mapping of archaeological topography on Arran, Scotland. Archaeol. Prospect. 26, 165–175. https://doi.org/10.1002/arp.1731

Trier, Ø.D., Larsen, S.Ø., Solberg, R., 2009. Automatic detection of circular structures in high-resolution satellite images of agricultural land. Archaeol. Prospect. 16, 1–15. https://doi.org/10.1002/arp.339

Trier, Ø.D., Salberg, A.-B., Pilø, L.H., 2016. Semi-automatic mapping of charcoal kilns from airborne laser scanning data using deep learning, in: Matsumoto, M., Uleberg, E. (Eds.), 44th Conference on Computer Applications and Quantitative Methods in Archaeology. Archaeopress, Oxford, UK, pp. 219–231.

Trier, Ø.D., Zortea, M., Tonning, C., 2015. Automatic detection of mound structures in airborne laser scanning data. J. Archaeol. Sci. Reports 2, 69–79. https://doi.org/10.1016/j.jasrep.2015.01.005

Tuokuu, F.X.D., Idemudia, U., Gruber, J.S., Kayira, J., 2019. Identifying and clarifying environmental policy best practices for the mining industry–A systematic review. J. Clean. Prod. 222, 922–933. https://doi.org/10.1016/j.jclepro.2019.03.111

Uijlings, J.R.R., van de Sande, K.E.A., Gevers, T., Smeulders, A.W.M., 2013. Selective Search for Object Recognition. Int. J. Comput. Vis. 104, 154–171. https://doi.org/10.1007/s11263-013-0620-5

Vaiopoulos, A.D., Karantzalos, K., 2016. Pansharpening on the narrow VNIR and SWIR spectral bands of Sentinel-2. ISPRS - Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. XLI-B7, 723–730. https://doi.org/10.5194/isprsarchives-XLI-B7-723-2016

Valenta, R., Clark, A., O'Sullivan, R., Thomas, J., 2018. Estimating Geometallurgical Risk in Undeveloped Complex Orebodies, in: Procemin-Geomet 2018: 14th International Mineral Processing Conference and the 5th International Seminar on Geometallurgy. Gecamin, Santiago, Chile.

Valler, A., Andersson, A.-S., 2020. Epiroc and Orica unveil prototype system for first stages of underground automation [WWW Document]. URL https://www.epiroc.com/en-uk/newsroom/2020/epiroc-and-orica-unveil-prototype-system-for-first-stages-of-underground-automation (accessed 1.15.21).

van der Meer, F.D., van der Werff, H.M.A., van Ruitenbeek, F.J.A., 2014. Potential of ESA's Sentinel-2 for geological applications. Remote Sens. Environ. 148, 124–133. https://doi.org/10.1016/j.rse.2014.03.022

van der Meer, F.D., van der Werff, H.M.A., van Ruitenbeek, F.J.A., Hecker, C.A., Bakker, W.H., Noomen, M.F., van der Meijde, M., Carranza, E.J.M., Smeth, J.B. de, Woldai, T., 2012. Multi- and hyperspectral geologic remote sensing: A review. Int. J. Appl. Earth Obs. Geoinf. 14, 112–128. https://doi.org/10.1016/j.jag.2011.08.002

van der Merwe, J.W., Andersen, D.C., 2013. Applications and benefits of 3D laser scanning for the mining industry. J. South. African Inst. Min. Metall. 113, 213–219.

Vandapel, N., Huber, D.F., Kapuria, A., Hebert, M., 2004. Natural terrain classification using 3-d ladar data, in: IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA '04. 2004. IEEE, pp. 5117–5122. https://doi.org/10.1109/ROBOT.2004.1302529

Verschoof-van der Vaart, W.B., Lambers, K., 2019. Learning to Look at LiDAR: The Use of R-CNN in the Automated Detection of Archaeological Objects in LiDAR Data from the Netherlands. J. Comput. Appl. Archaeol. 2, 31–40. https://doi.org/10.5334/jcaa.32

Vo, A.-V., Truong-Hong, L., Laefer, D.F., Bertolotto, M., 2015. Octree-based region growing for point cloud segmentation. ISPRS J. Photogramm. Remote Sens. 104, 88–100. https://doi.org/10.1016/j.isprsjprs.2015.01.011

Vosselman, G., Coenen, M., Rottensteiner, F., 2017. Contextual segment-based classification of airborne laser scanner data. ISPRS J. Photogramm. Remote Sens. 128, 354–371. https://doi.org/10.1016/j.isprsjprs.2017.03.010

Vosselman, G., Klein, R., 2010. Visualisation and Structuring of Point Clouds, in: Vosselman, G., Maas, H.-G. (Eds.), Airborne and Terrestrial Laser Scanning. Whittles Publishing, Dunbeath, UK.

Vosselman, G., Maas, H.-G., 2010. Airborne and terrestrial laser scanning. Whittles Publishing, Dunbeath, UK.

Walsh, G., 2015. Leica ScanStation P-Series – Details that matter. Leica ScanStation - White Paper. Leica Geosystems AG.

Walsh, L., 2016. A review of terrestrial laser scanner calibration and the establishment of the Bentley Calibration Field. Curtin University.

Wang, C., Zhao, Z., Ren, Q., Xu, Y., Yu, Y., 2019. Dense U-net Based on Patch-Based Learning for Retinal Vessel Segmentation. Entropy 21, 168. https://doi.org/10.3390/e21020168

Wang, H., Jiang, J., Zhang, G., 2018. CraterIDNet: An End-to-End Fully Convolutional Neural Network for Crater Detection and Identification in Remotely Sensed Planetary Images. Remote Sens. 10, 1067. https://doi.org/10.3390/rs10071067

Wang, Q., Wu, L., Chen, S., Shu, D., Xu, Z., Li, F., Wang, R., 2014. Accuracy Evaluation of 3D Geometry from Low-Attitude UAV collections A case at Zijin Mine. ISPRS - Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. XL–4, 297–300. https://doi.org/10.5194/isprsarchives-XL-4-297-2014

Wang, W., Zhao, W., Huang, L., Vimarlund, V., Wang, Z., 2014. Applications of terrestrial laser scanning for tunnels: a review. J. Traffic Transp. Eng. (English Ed. 1, 325–337. https://doi.org/10.1016/S2095-7564(15)30279-8

Wang, Y., Tian, F., Huang, Y., Wang, J., Wei, C., 2015. Monitoring coal fires in Datong coalfield using multi-source remote sensing data. Trans. Nonferrous Met. Soc. China 25, 3421–3428. https://doi.org/10.1016/S1003-6326(15)63977-2

Wanli Ouyang, Xiaogang Wang, 2012. A discriminative deep model for pedestrian detection with occlusion handling, in: 2012 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, pp. 3258–3265. https://doi.org/10.1109/CVPR.2012.6248062

Weinmann, M., 2016. Reconstruction and Analysis of 3D Scenes, Reconstruction and Analysis of 3D Scenes: From Irregularly Distributed 3D Points to Object Classes. Springer International Publishing, Cham, Switzerland. https://doi.org/10.1007/978-3-319-29246-5

Weinmann, M., Jutzi, B., Hinz, S., Mallet, C., 2015a. Semantic point cloud interpretation based on optimal neighborhoods, relevant features and efficient classifiers. ISPRS J. Photogramm. Remote Sens. 105, 286–304. https://doi.org/10.1016/j.isprsjprs.2015.01.016

Weinmann, M., Urban, S., Hinz, S., Jutzi, B., Mallet, C., 2015b. Distinctive 2D and 3D features for automated large-scale scene analysis in urban areas. Comput. Graph. 49, 47–57. https://doi.org/10.1016/j.cag.2015.01.006

Weinmann, Martin, Weinmann, Michael, Mallet, C., Brédif, M., 2017. A Classification-Segmentation Framework for the Detection of Individual Trees in Dense MMS Point Cloud Data Acquired in Urban Areas. Remote Sens. 9, 277. https://doi.org/10.3390/rs9030277

Weisse, M., Goldman, E.D., 2019. The World Lost a Belgium-sized Area of Primary Rainforests Last Year [WWW Document]. World Resour. Inst. URL https://www.wri.org/blog/2019/04/world-lost-belgium-sized-area-primary-rainforests-last-year?utm_campaign=GFW&source=socialmediakit&utm_medium=gfwsocial&utm_term=2018tcl_4_2019 (accessed 12.4.19).

Weng, Q., 2012. An Introduction to Contemporary Remote Sensing. McGraw-Hill, New York, NY, USA.

Werner, T.T., Bebbington, A., Gregory, G., 2019. Assessing impacts of mining: Recent contributions from GIS and remote sensing. Extr. Ind. Soc. 6, 993–1012. https://doi.org/10.1016/j.exis.2019.06.011

Williams, D.L., Goward, S., Arvidson, T., 2006. Landsat: Yesterday, Today, and Tomorrow. Photogramm. Eng. Remote Sens. 72, 1171–1178. https://doi.org/10.14358/PERS.72.10.1171

Wilson, A.C., Roelofs, R., Stern, M., Srebro, N., Recht, B., 2017. The Marginal Value of Adaptive Gradient Methods in Machine Learning, in: Guyon, I., Luxburg, U. V, Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., Garnett, R. (Eds.), Advances in Neural Information Processing Systems 30. Curran Associates, Inc., Red Hook, NY, USA, pp. 4148–4158.

Winter, S., 2017. Uncovering England's landscape by 2020 [WWW Document]. URL https://environmentagency.blog.gov.uk/2017/12/30/uncovering-englands-landscape-by-2020/ (accessed 6.9.19).

Wired Magazine, 2020. Sandvik's automated loader is so precise it can make it through a glass maze [WWW Document]. URL https://www.wired.co.uk/article/sandvik-autonomous-loader (accessed 1.15.21).

Wold, S., Esbensen, K., Geladi, P., 1987. Principal component analysis. Chemom. Intell. Lab. Syst. 2, 37–52. https://doi.org/10.1016/0169-7439(87)80084-9

Wright, P., Stow, R., 1999. Detecting mining subsidence from space. Int. J. Remote Sens. 20, 1183–1188. https://doi.org/10.1080/014311699212939

Wu, D., Meng, Y., Zhan, K., Ma, F., 2018. A LiDAR SLAM Based on Point-Line Features for Underground Mining Vehicle, in: 2018 Chinese Automation Congress (CAC). IEEE, pp. 2879–2883. https://doi.org/10.1109/CAC.2018.8623075

Wurm, M., Stark, T., Zhu, X.X., Weigand, M., Taubenböck, H., 2019. Semantic segmentation of slums in satellite images using transfer learning on fully convolutional neural networks. ISPRS J. Photogramm. Remote Sens. 150, 59–69. https://doi.org/10.1016/j.isprsjprs.2019.02.006

Xiang, J., Chen, J., Sofia, G., Tian, Y., Tarolli, P., 2018. Open-pit mine geomorphic changes analysis using multi-temporal UAV survey. Environ. Earth Sci. 77, 220. https://doi.org/10.1007/s12665-018-7383-9

Xie, Y., Tian, J., Zhu, X.X., 2020. Linking Points With Labels in 3D: A Review of Point Cloud Semantic Segmentation. IEEE Geosci. Remote Sens. Mag. 8, 38–59. https://doi.org/10.1109/MGRS.2019.2937630

Xiong, Y., Zuo, R., 2020. Recognizing multivariate geochemical anomalies for mineral exploration by combining deep learning and one-class support vector machine. Comput. Geosci. 140, 104484. https://doi.org/10.1016/j.cageo.2020.104484

Xiong, Y., Zuo, R., Carranza, E.J.M., 2018. Mapping mineral prospectivity through big data analytics and a deep learning algorithm. Ore Geol. Rev. 102, 811–817. https://doi.org/10.1016/j.oregeorev.2018.10.006

Xu, T., Xu, L., Li, X., Yao, J., 2018. Detection of Water Leakage in Underground Tunnels Using Corrected Intensity Data and 3D Point Cloud of Terrestrial Laser Scanning. IEEE Access 6, 32471–32480. https://doi.org/10.1109/ACCESS.2018.2842797

Xueyun Chen, Shiming Xiang, Cheng-Lin Liu, Chun-Hong Pan, 2014. Vehicle Detection in Satellite Images by Hybrid Deep Convolutional Neural Networks. IEEE Geosci. Remote Sens. Lett. 11, 1797–1801. https://doi.org/10.1109/LGRS.2014.2309695

Yang, B., Fang, L., Li, J., 2013. Semi-automated extraction and delineation of 3D roads of street scene from mobile laser scanning point clouds. ISPRS J. Photogramm. Remote Sens. 79, 80–93. https://doi.org/10.1016/j.isprsjprs.2013.01.016

Yeomans, C.M., 2018. Enhancing the Geological Understanding of Southwest England Using Machine Learning Algorithms. University of Exeter.

Yeomans, C.M., 2017. Tellus South West data usage: a review (2014–2016), in: Jones, D.M. (Ed.), Open University Geological Society. Open University Geological Society, Exeter, pp. 51–61.

Yeomans, C.M., Middleton, M., Shail, R.K., Grebby, S., Lusty, P.A.J., 2019. Integrated Object-Based Image Analysis for semi-automated geological lineament detection in southwest England. Comput. Geosci. 123, 137–148. https://doi.org/10.1016/j.cageo.2018.11.005

Yu, H., Lu, X., Cheng, G., Ge, X., 2011. Detection and volume estimation of mining subsidence based on multi-temporal LiDAR data, in: 2011 19th International Conference on Geoinformatics. IEEE, pp. 1–6. https://doi.org/10.1109/GeoInformatics.2011.5980892

Yu, L., Porwal, A., Holden, E.-J., Dentith, M.C., 2012. Towards automatic lithological classification from remote sensing data using support vector machines. Comput. Geosci. 45, 229–239. https://doi.org/10.1016/j.cageo.2011.11.019

Yumus, D., Ozkazanc, Y., 2019. Land Cover Classification for Synthetic Aperture Radar Imagery by Using Unsupervised Methods, in: 2019 9th International Conference on Recent Advances in Space Technologies (RAST). IEEE, pp. 435–440. https://doi.org/10.1109/RAST.2019.8767877

Zeiler, M.D., Fergus, R., 2014. Visualizing and Understanding Convolutional Networks, in: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (Eds.), Computer Vision -- ECCV 2014. Springer International Publishing, Cham, Switzerland, pp. 818–833. https://doi.org/10.1007/978-3-319-10590-1_53

Zeng, F., Jacobson, A., Smith, D., Boswell, N., Peynot, T., Milford, M., 2019. LookUP: Vision-Only Real-Time Precise Underground Localisation for Autonomous Mining Vehicles, in: 2019 International Conference on Robotics and Automation (ICRA). IEEE, pp. 1444–1450. https://doi.org/10.1109/ICRA.2019.8794453

Zhang, J., Lin, X., Ning, X., 2013. SVM-Based Classification of Segmented Airborne LiDAR Point Clouds in Urban Areas. Remote Sens. 5, 3749–3775. https://doi.org/10.3390/rs5083749

Zhao, S., Zhang, D.M., Huang, H.W., 2020. Deep learning–based image instance segmentation for moisture marks of shield tunnel lining. Tunn. Undergr. Sp. Technol. 95, 103156. https://doi.org/10.1016/j.tust.2019.103156

Zhao, Yuan, Song, Ding, Lin, Liang, Zhang, 2019. Use of Unmanned Aerial Vehicle Imagery and Deep Learning UNet to Extract Rice Lodging. Sensors 19, 3859. https://doi.org/10.3390/s19183859

Zhou, Y., Tuzel, O., 2018. VoxelNet: End-to-End Learning for Point Cloud Based 3D Object Detection, in: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, pp. 4490–4499. https://doi.org/10.1109/CVPR.2018.00472

Zhu, X.X., Tuia, D., Mou, L., Xia, G.-S., Zhang, L., Xu, F., Fraundorfer, F., 2017. Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources. IEEE Geosci. Remote Sens. Mag. 5, 8–36. https://doi.org/10.1109/MGRS.2017.2762307

Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., Xiong, H., He, Q., 2021. A Comprehensive Survey on Transfer Learning. Proc. IEEE 109, 43–76. https://doi.org/10.1109/JPROC.2020.3004555

Zuber, M.T., Smith, D.E., Zellar, R.S., Neumann, G.A., Sun, X., Katz, R.B., Kleyner, I., Matuszeski, A., McGarry, J.F., Ott, M.N., Ramos-Izquierdo, L.A., Rowlands, D.D.,

Torrence, M.H., Zagwodzki, T.W., 2010. The Lunar Reconnaissance Orbiter Laser Ranging Investigation. Space Sci. Rev. 150, 63–80. https://doi.org/10.1007/s11214-009-9511-z

Zuo, R., Xiong, Y., Wang, J., Carranza, E.J.M., 2019. Deep learning and its application in geochemical mapping. Earth-Science Rev. 192, 1–14. https://doi.org/10.1016/j.earscirev.2019.02.023