Multi-Object Detector YOLOv4-Tiny Enables High-Throughput Combinatorial and Spatially-Resolved Sorting of Cells in Microdroplets

Lewis Howell, Vasileios Anagnostidis, and Fabrice Gielen*

The encapsulation of cells together with micro-objects in monodispersed water-in-oil microdroplets offers a powerful means to perform quantitative biological studies within large cell populations. In such applications, accurate object detection is crucial to ensure control over the content for every compartment. In particular, the ability to rapidly count and localize objects is key to future applications in single-cell -omics, cellular aggregation, and cell-to-cell interactions. In this paper, the authors combine the Deep Learning object detector YOLOv4-tiny with microfluidic Image-Activated Droplet Sorting (DL-IADS), to perform flexible, label-free classification, counting, and localization of multiple micro-objects simultaneously and at high-throughput. They trained YOLOv4-tiny to detect SH-SY5Y cells, polyacrylamide beads, and cellular aggregates in a single model, with a precision of 92% for cells, 98% for beads, and 81% for aggregates. They exploit this accuracy and counting ability to implement a closed-loop feedback that enables controlled loading of microbeads via the automated adjustment of flow rates. They subsequently demonstrate the combinatorial sorting of co-encapsulated single cells and single beads based on real-time classification at up to 111 Hz, with enrichment factors of up to 145. Finally, they demonstrate spatially-resolved sorts by evaluating cell-to-cell distances in real-time to isolate cell doublets with high purity.

1. Introduction

Droplet microfluidics is a powerful technology that provides a quantitative biological platform to study single-cell heterogeneity or cell–cell interactions.[1–3] Single-cell -omic assays allow access to genetic information which can be correlated with cellular phenotypes, providing insight into tissue physiology and disease pathways that are impossible to infer from bulk analysis alone.[4] On the other hand, aggregation of cells

into multicellular assemblies and interaction between single cells can also be studied at high-throughput and used for generating physiologically relevant 3D cell cultures and the isolation of rare cells, for example, for cancer research.[2,5] The core principle of droplet-based microfluidics is to harness the precision of microfabricated devices to manipulate small volumes of reagents and enable high-throughput, low-waste experiments. To generate microdroplets, a shear force is applied at the interface between two immiscible phases, dividing the continuous flow into discrete picoliter droplets, such that individual cells or other objects can be encapsulated within. Incorporating assay components provides the ability to effectively control individual cellular micro-environments, isolated by the oil–water interface which acts as a barrier to diffusion.

A key aspect in cell encapsulation is the development of protocols in which the experimenter exerts control over the exact make-up of droplets, including the number of cells per micro-reactor and the addition of other discrete objects necessary to the high-throughput assay. For instance, co-encapsulation of individual cells together with individual hydrogel beads carrying barcoding DNA primers is used for the capture of DNA/RNA molecules and extraction of quantitative information on whole transcriptomes or genomes, across thousands of cells.[6,7] Likewise, cellular co-cultures typically require two single cells of different types to be placed in the same droplet.

Although used in many such applications, achieving controlled multi-object microfluidic capture on a droplet-by-droplet basis is far from trivial. The loading of discrete objects into droplets is usually stochastic and therefore fundamentally limited by Poisson statistics.[8] In single-cell applications, if a maximum 10% of droplets are permitted to contain more than one cell, then in a Poisson limited system, only 15.6% of the drops can theoretically contain exactly one cell, with most droplets being empty. In the study of cell–cell interactions or cell doublets, the theoretical maximum number of droplets having 2 cells of the same type is only ≈27%.

Therefore, specialized approaches to microfluidic encapsulation are needed to confer a higher degree of control with respect to the desired droplet contents, transitioning from stochastic

L. Howell, V. Anagnostidis, F. Gielen
University of Exeter
Living Systems Institute and College of Engineering
Mathematics and Physical Sciences
Stocker Road, Exeter EX4 4QD, UK
E-mail: f.gielen@exeter.ac.uk

(Poisson) to fully deterministic loading. Previous deterministic approaches typically introduce constraints on experimental parameters or have small operational windows that limit their use (e.g., aspect ratio, geometries, flow rates, etc.). For instance, high-aspect ratio channels can be used to provide self-organization of cells at high flow rates.[9] Passive microfluidic methods such as inertial ordering have demonstrated single-cell encapsulation of up to 80% efficiency but are limited by high flow rate requirements and applicability only to single objects at specific concentrations.[10] Close-packed ordering of soft objects (e.g., hydrogel beads) can be used to provide deterministic loading of single beads, but requires channel dimensions matching accurately the size of the beads.[11] Therefore, techniques to selectively control the number of micro-objects in droplets, "beating" Poisson statistics, are crucial in overcoming the inherent limitations of stochastic encapsulation.

To generalize micro-loading operations in cases where deterministic loading cannot be implemented, active sorting of subpopulations provides a means to select droplets of interest with high flexibility and high purity. Techniques such as image-activated droplet sorting (IADS) have proven effective for the identification and classification of a variety of micro-objects in bright-field images. To date, approaches based on template-matching and thresholding have been implemented, demonstrating the feasibility to reconcile imaging with high-throughput.[12–14] However, these methods do not fully leverage the spatial resolution offered by bright-field microscopy, lack robustness to differing object appearance, and cannot easily deal with multi-class, multi-object counting, particularly with overlapping objects. Furthermore, although progress has been recently made in imaging flow cytometers, they still require fluorescent cell stains to trigger acquisition of bright-field images and are not designed to handle water-in-oil emulsions.[15,16] Combining bright-field microscopy and label-free microfluidics allows rapid collection of large and temporally resolved datasets of images. Such datasets are particularly suitable for advanced image analysis and particularly machine learning tasks including image classification and object detection. This allows access to higher level information including the identification and localization of objects in individual droplets, as well as global loading distributions and characterization of device operation such as throughput and monodispersity.

Previously, we have demonstrated the use of convolutional neural networks (CNNs) to perform real-time sorting of cells and polyacrylamide beads.[17] However, traditional image classification CNNs were not well suited to the simultaneous identification and counting of multiple objects. Such is the task of object detectors, which can both classify and localize multiple classes of objects within bounding boxes. Amongst these is YOLOv4; the fourth iteration of the 'You Only Look Once' one-stage, anchor-based detector.[18] The model uses a single neural network to process the entire image and identify objects within, making YOLOv4 significantly faster than other two-stage techniques such as Fast R-CNN.[19] Reduced implementations of YOLOv4 such as YOLOv4-tiny further improve timing performance, simplifying the network architecture, and reducing the number of free parameters to make detection feasible for high-throughput applications at rates of hundreds of frames per second (FPS).[18]

A secondary challenge for optimized object loading in real-time in microfluidic experiments is the mostly manual, arbitrary, and often time-consuming process of adjusting experimental variables such as flow rates of the carrier oil and aqueous phases. Furthermore, drift correction is often needed during experiments to adjust for pressure fluctuations in the device, meaning constant operator supervision is required. In contrast, methods relying on feedback loops allow for automated corrections, optimizing device operation and function, and bypassing constraints of flow instabilities and strict requirements on microfluidic geometries. For instance, an image-based feedback loop has been developed to keep droplets monodisperse long-term.[20,21]

In this work, the recently developed object detection algorithm YOLOv4-tiny was implemented for multi-class counting of cells, cell aggregates, and polyacrylamide (PA) beads with both high precision and processing speed. We subsequently show integration of deep learning detections with a feedback control loop to enable controlled loading of soft PA beads with a chosen target occupancy. Using feedback, we demonstrate an on-demand combinatorial selection of droplets co-encapsulating both a single-cell and a single-bead. Finally, we utilize the localization ability of our system to study cell association by isolation of cell doublets using real-time calculation of cell-to-cell distances.

2. Results

In this study, we focused on the development of a generic platform for the rapid and accurate identification of micro-objects used in high-throughput screening experiments. To this end, we have implemented the deep learning object detector YOLOv4-tiny to detect, classify, and localize cells and beads encapsulated in water-in-oil droplets from bright-field images. The microfluidic device for droplet creation and sorting, along with the optical setup, was used as previously reported.[17] This setup offers a flexible platform for on-demand droplet sorting and enrichment, with droplets formed in the same device (**Figure 1**A).

In brief, image acquisition is triggered using a photodetector that detects light fluctuation at the interface between the aqueous droplets and surrounding oil, enabling imaging at reproducible locations. An iris diaphragm (I1) was placed after the collimated white LED light source to increase the effective depth of field at the focal plane. The image is subsequently grabbed, pre-processed, and passed to the YOLOv4-tiny object detector whose identifications are used to trigger electrodes, permitting selection of droplets having the desired content. In addition to multiclass detection, we have implemented a feedback mechanism for deterministic PA bead loading in which real-time counting is used to feedback the flow rates of syringe pumps, as described in Section 2.3 below.

2.1. Object Detection for Droplet Sorting

A primary requirement of our system is to recognize, count, and localize multiple micro-objects of different types using
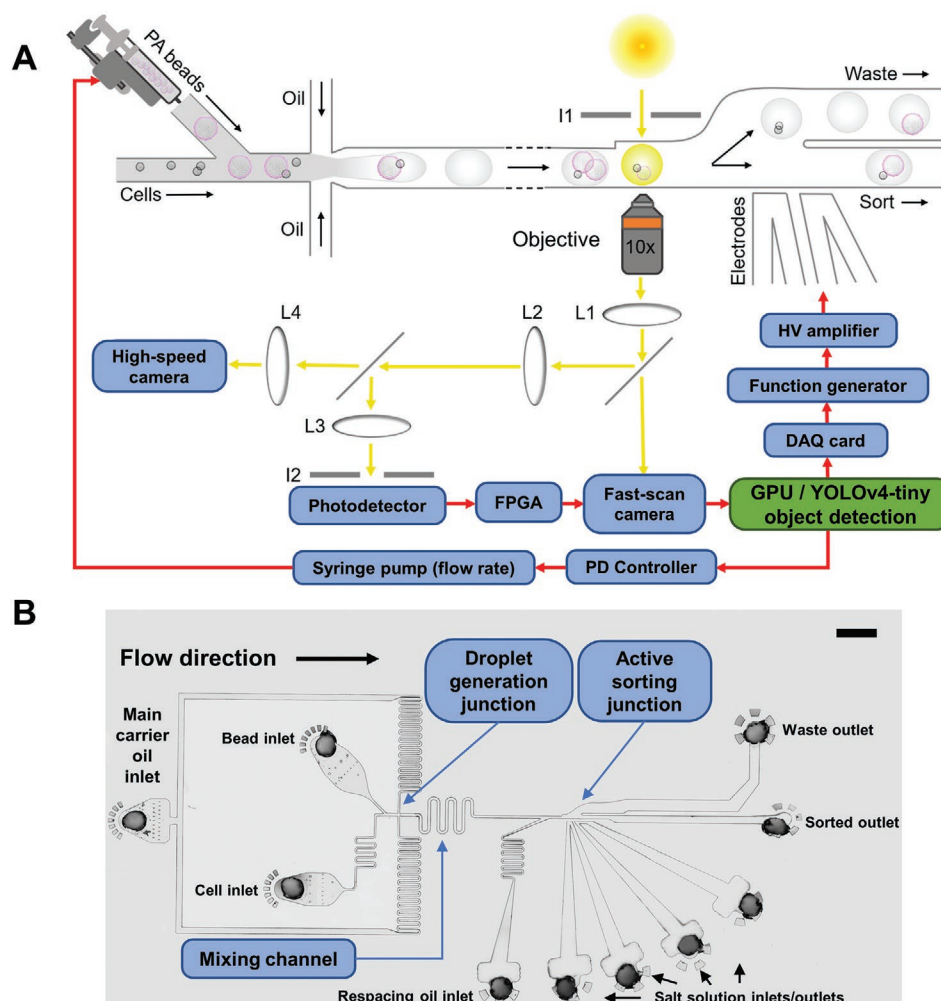
Figure 1. A) Schematic of the microfluidic platform for droplet sorting based on real-time object detection. Yellow arrows indicate the optical path and red arrows show the control sequence for feedback and sorting. The droplet generation and sorting junctions are illustrated, with black arrows representing the direction of flows. Droplets are imaged in bright-field with an inverted microscope, illuminated by white light passing through an iris (I1). After magnification by a 10× objective lens, the light is split towards a fast-scan camera, a photodetector, and a high-speed camera. Four lenses in the optical path allow the high-speed camera to capture an area covering the whole sorting junction (800 µm × 800 µm) while the fast-scan camera captures a close-cropped image of the droplet (160 µm × 160 µm). Light fluctuations due to a droplet passing by is detected with a photodetector placed directly behind an iris (I2), causing the FPGA to trigger acquisition of a single frame. The YOLOv4-tiny deep neural network detects micro-objects in the image, outputting a logical decision to the DAQ card and a correction to the proportional-derivative (PD) controller. To sort a droplet, an electric pulse from the function generator is relayed to the high voltage (HV) amplifier to deflect it by dielectrophoresis. Closed-loop feedback is completed with syringe pumps which vary the flow rate to optimize droplet loading. Plano-convex lenses L1, L2, L3, and L4 have focal lengths of 50, 50, 25.4, and 25.4 mm, respectively. B) Image of the microfluidic device, incorporating three main modules: a droplet generation junction with two inlets for cells and beads, respectively, a serpentine mixing channel and an active sorting junction. A respacing oil inlet increased inter droplet distances such that multiple droplets did not get selected simultaneously. Salt solutions were used to conduct the electric fields triggering dielectrophoretic motion. Scale bar: 2 mm.

a single deep learning model. A total of three classes were selected to demonstrate this: PA beads, SH-SY5Y cells, and cell aggregates. In this work, we define cell aggregates as any cell cluster of 3 or more contiguous cells which have grouped in suspension following dissociation from their monolayer culture and can be used as initial templates for growing multicellular spheroids.[22] The aggregate class is useful to detect single cells overlapping which are difficult to enumerate and can be used as a criterion for selection or exclusion.

To classify images of microdroplets, YOLOv4-tiny neural networks were trained on custom datasets obtained from microfluidic experiments. Ground truth datasets were constructed from a total of 2000 images, each of a single microdroplet containing n cells (SH-SY5Y) and m beads (65 µm in diameter, polyacrylamide) with n, m as integers greater than or equal to zero. When possible, individual cells within cell aggregates were labelled in addition to the aggregate itself which allows for the estimation of the number of cells making up each aggregate. Prior to labelling, a circular mask was applied to confine the region of interest to within the droplet diameter and images converted to greyscale. YOLOv4 format labels were created using the open-source tool LabelImg to annotate images in the ground truth dataset.[23] Objects were labelled by drawing bounding boxes fully closing objects and allocating the correct

**ADVANCED
SCIENCE NEWS**
www.advancedsciencenews.com

**ADVANCED
MATERIALS
TECHNOLOGIES**
www.advmattechnol.de

class, including partially obscured or out-of-frame objects with greater than 50% in view.

Following training set construction, training took place on a cloud virtual machine and took ≈1.5 h. As expected, the model loss fell exponentially as the iteration number increased (Figure S1, Supporting Information), dropping to below 0.3 in the latter half of the training which indicates strong feature learning from example images. Every 1000 iterations, the validation set was used to calculate the mean average precision (mAP), which plateaued at ≈90% after 4000 iterations. This mAP measures performance both in terms of classification ability (the types of objects in an image) and localization ability (the position of the objects in the image). mAP is calculated by the mean of the average precision over all classes, for a range of thresholds defining the minimum overlap between the predicted and ground truth bounding boxes (intersection over union, or IoU). The optimal model weights were selected at 4000 iterations, the threshold for maximal feature learning before overfitting—where further training causes a loss decrease without a corresponding increase of mAP. This is important as overfitted models have poor generalizability to unseen data, becoming overly complex to fit for noise or features specific to the training set.

To validate model performance, the network was evaluated against a hand-labelled, ground truth dataset of test images unseen by the computer during training. YOLOv4-tiny classifications show strong agreement with the ground truth (**Figure 2**A), effectively locating and classifying single and multiple objects, even in cases of overlapping or partially obscured objects in multi-object images. The model correctly identified cells with a variety of morphologies, in different brightness and contrasts modes, including many of those in proximity as doublets or part of larger aggregates. Experimental predictions were confirmed on over 10 000 images, acquired across 4 separate experiments, on two different dates, and manually labelled before comparison with YOLOv4-tiny predictions. The results can be summarized by confusion matrices (Figure 2B) showing the comparison between ground truth and YOLOv4-tiny object classification.

Figure 2B-(i) shows the overall number of true positive, false positive, and false negative predictions for each class, normalized by the total number of images. Cell counting was found to have a true positive rate of 92% with no false positives. PA beads had a true positive and false positive rate of 98.3% and 0.6%, respectively. This indicated that sorting based on the number of beads and/or cells would lead to a droplet population of high purity with a low proportion of unnecessarily wasted droplets. Selection of aggregates had a lower true positive rate of 80.7% and false positive rate of 3.8%, which may be partially attributed to ambiguities in the ground truth labelling of large aggregates.

The per class counting accuracy for beads, cells, and aggregates (Figure 2B-(ii), (iii), and (iv), respectively) shows high precision in counting beads (over 95.2% in all cases) and cells (from 96% for one cell down to 64.1% for 5 cells), which decreases with increasing cell number mainly due to the 3D orientation of cell clusters obscuring individual cells. Importantly, single cells and single beads were identified with high precision, a crucial criterion for selection in many single-cell applications.

2.1.1. Application to Detection of Cellular Aggregates

Even at fixed cell density, we observed variations in droplet-by-droplet cell loading numbers due to both Poisson statistics and pre-aggregation events. This is reflected in the histogram for the number of cells per droplet in a typical experiment (Figure S2A, Supporting Information) which does not follow expected Poisson statistics. We attribute this to the lack of stirring of the cell suspension during handling and injection.

The aggregate class detection helped us identify large groups of single cells overlapping and hard to enumerate and could be used as criterion for selection of multicellular clusters or exclusion during sorts. The number of cells detected per aggregate typically ranged from 3 to 20 with rapid decrease in occurrence beyond small cell clusters with 5 cells or fewer (Figure S2B, Supporting Information). The number of aggregates also enabled us to study the kinetics of cellular aggregation over the experimental course, following detachment from the plastic substrate of the culture dish by trypsinization. We generally observed a gradual decrease in the rate of single cells, coupled with an increase in the rate of aggregates (Figure S3, Supporting Information) despite cells being suspended in a density-matching solution. This apparent decrease in the incoming rate of single cells was up to 40% for experiments lasting over 15 min.

2.2. Timing Performance

The classification speed of neural networks is vital for integration within high-throughput microfluidic systems where a capacity to image tens to thousands of droplets per second is a requirement to enable population scale experiments. Here, the reduced architecture of YOLOv4-tiny is shown to be not only viable for real-time applications, but also for the very high working throughputs of microfluidics. The time for each processing step was measured using the python system-wide performance counter, at the highest available clock resolution. As with a previous study, microdroplet images were acquired individually on a photodetector trigger, rather than continuous processing of video data.[17] This had the advantage of allowing a higher effective throughput by preventing processing of redundant frames in between droplets. Once an image was acquired, it was grabbed from the camera buffer and pre-processed by resizing it to 416 × 416 pixels, the application of a circular mask, followed by normalization. The latency due to data transfer and pre-processing was found to be small, averaging less than 0.4 ms per image. On average, object detection took 9 ms per image, implying a maximum possible framerate of 111 FPS. At the cost of additional processing time, it is also useful to save the images and YOLO detection outputs when obtaining training data, or for verification of sorting accuracy. Of three different python computer vision libraries benchmarked (OpenCV, PIL, and Scikit-image), saving images in the JPEG format with OpenCV was most performant, allowing sorting and image acquisition at up to 101 FPS. Additional information including bounding box coordinates and individual confidence scores for all objects could
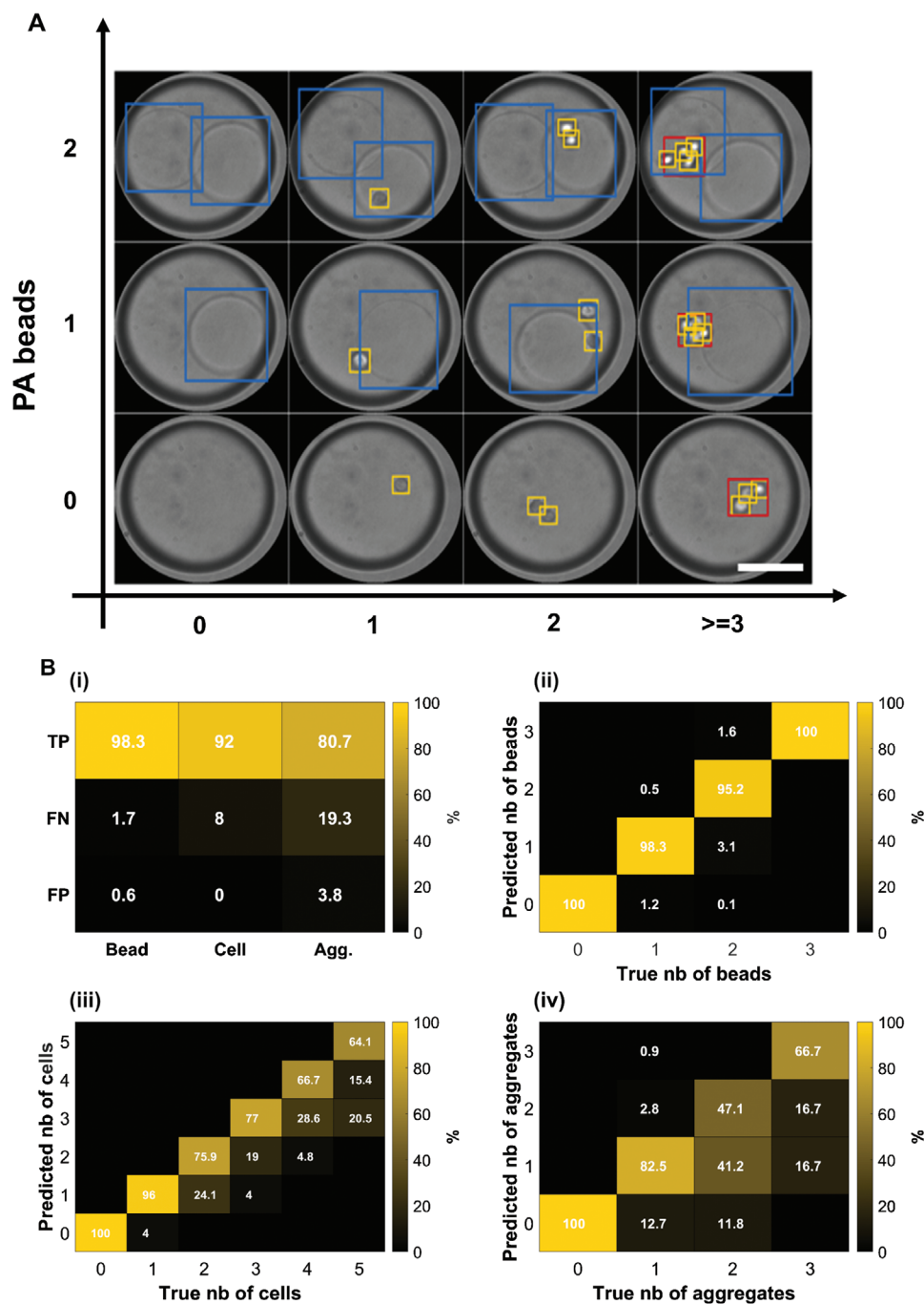
Figure 2. Object classification and detection with YOLOv4-tiny. A) Example images showing correct YOLOv4-tiny bounding boxes for combinations of beads and cells (blue: PA bead, yellow: cell, red: aggregate). Scale bar: 65 µm. B) (i) Normalized confusion matrix showing rates of true positive (TP), false positive (FP), and false negative (FN) predictions per class. (ii) to (iv) Normalized confusion matrix showing ground truth (true nb) versus number of detected objects (Predicted nb) for PA beads (ii), cells (iii), and aggregates (iv).

be recorded to allow full reconstruction of YOLO detections at up to 85 FPS. Rank-ordered timing for over 15 000 images showed that images were classified and saved with bounding boxes with a minimum timing of 11.5 ms, a median timing of 14 ms, and with only 3% of the images processed in over 20 ms (Figure S4, Supporting Information). However the experimental sorting rates achieved were lower than this theoretical maximum as the distance travelled by droplets between detection and triggering at high flow rates (above 35 µL min⁻¹ for total flow rates) exceeded the distance between imaging and sorting junction. A maximum sorting rate of ≈45 Hz was achieved at flow rates of 2 µL min⁻¹ for both cells and PA beads supernatant and 15 µL min⁻¹ for both the carrier and respacing oil phases.

**ADVANCED
SCIENCE NEWS**

www.advancedsciencenews.com

**ADVANCED
MATERIALS
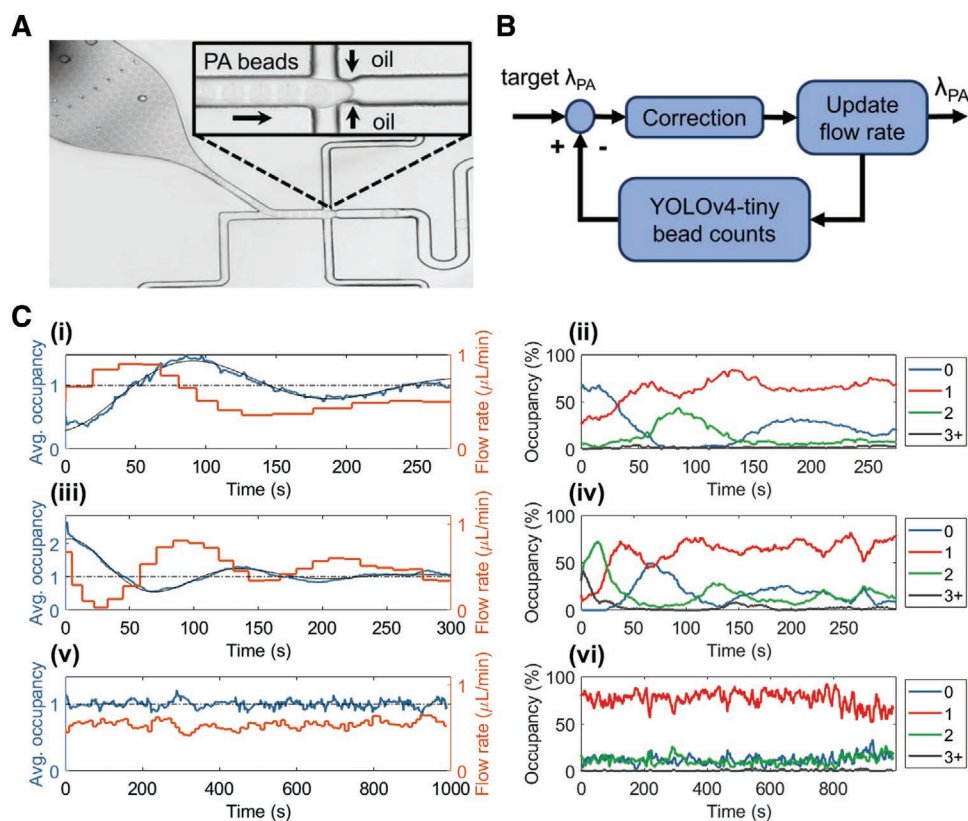TECHNOLOGIES**

www.advmattechnol.de

Figure 3. Closed-loop feedback for optimized encapsulation of PA beads. A) Image of bead packing at the entrance of the device. The inset highlights the multilayer packing just before encapsulation. Black arrows indicate flow directions. Scale bar: 200 μm. B) Feedback loop implementation flowchart: a target occupancy is set (target λ_{PA}), and the YOLOv4-tiny counting used to update the flow rate for the beads by calculating the difference between actual and target occupancies (correction). C) (i), (iii), (v): Average occupancy (blue line) and flow rate for the beads (orange line) as a function of time for proportional (i) and proportional-derivative ((iii) and v)) corrections. A horizontal dashed line represents the target occupancy of 1. The black curves in (i) and (iii) are the best fit for PA bead counts using a damped oscillation model. (ii), (iv), and (vi): Fraction of droplets containing a given number of beads as a function of time averaged over 200 images for proportional (ii) and proportional-derivative corrections ((iv) and (vi)).

2.3. Closed-Loop Feedback for Controlled PA Bead Encapsulation

Typical microfluidics encapsulation experiments require empirical determination of flow conditions that result in optimum object loading occupancies. However, tuning flows can be time-consuming, susceptible to inter-device variability, and may not be stable over the duration of a prolonged high-throughput experiment (e.g., over 10 min), requiring constant operator attention.

One such example is flow-rate dependent bead loading. Although fully deterministic loading has been demonstrated for the loading of soft microbeads in channel sizes closely matching the beads' diameters (in which they form a monolayer), this is generally not the case in other geometries. For example, in the deeper microchannels we have used, beads form a close-packed double layer upon loading (**Figure 3**A) so that several beads can be co-encapsulated in the same droplet. However, faster flow rates for PA delivery leads to a greater number of beads per droplet, making their distributions tuneable.[11] This opens a path for a fully integrated droplet loading system, where active feedback is

used to automatically control pump flow rates and optimize droplet loading for optimal operation and minimum waste (Figure 3).

High accuracy object detection and counting by YOLOv4-tiny can be used to adjust flows, creating a closed-loop feedback. We have implemented the scheme shown in Figure 3B where the correction factor represents the difference between a target and actual occupancy for PA beads (target λ_{PA} and λ_{PA}) whose average is measured over 200 consecutive images to prevent sudden changes in operation. We tested mainly a target occupancy of one bead per droplet to demonstrate deterministic loading but other occupancies, including non-integers, can be used. Figure S5, Supporting Information, shows an example feedback control for a target occupancy of 2. In a typical experiment, the flow rate for the oil phase and a PBS solution (not shown in Figure 3A) were kept constant at 12 and 1 μL min⁻¹ respectively, while the PA beads flow rate was determined by the state of the feedback loop. A proportional (*P*) or proportional-derivative (*PD*) feedback was used to update the flow rate of the PA beads:

$$u(t) = K_p\, e(t) + K_d\, \frac{de(t)}{dt} \qquad (1)$$

where u is the new flow rate for PA beads, t is time, e is the error, and K_p and K_d represent the proportional and derivative gains, respectively. The gain constants K_p and K_d were tested empirically to determine the optimal response of the loop. In absence of a derivative component, it was found that setting K_p above 0.5 resulted in highly undamped to slightly damped oscillations (Figure S6, Supporting Information). Generally, proportional correction alone did not fully stabilize the average loading occupancy over 30 min, leading to a periodic oscillation around the target value as seen in Figure 3C-(i),(ii). In contrast, adding a derivative gain term, which anticipates future error, resulted in slightly damped to critically damped oscillations (Figure 3C-(iii) to (vi) and Figure S7, Supporting Information), stabilizing occupancies around the target value much more efficiently.

The feedback was also robust to step changes in oil flow rates as shown in Figure S8, Supporting Information, as could happen in cases of channels being obstructed by foreign objects.

Using a 110 µm height channel and 65 µm beads, we were unable to reach perfect loading of one bead per droplet, even using pulseless syringe pumps, because of instabilities in flow delivery and irregular bead arrival at the flow-focusing junction. This suggests that the operational window to obtain such one-bead-one-droplet synchrony may be small. Indeed, the maximum proportion of single beads averaged over 200 droplets was ≈90%. Above 9% of the remaining 10% of droplets contained either zero or two beads, with few droplets containing 3 or more beads. Although capturing 4 or more beads was feasible, it was exceedingly rare at the low occupancies we targeted, for example, with 3 droplets containing 4 beads out of over 19 000 for Figure 3-(vi). The use of fast response pressure pumps may enable the required higher degree of control.[20] Furthermore, monodispersity was not found to change significantly with the adjustment of flow rates as the aqueous:oil ratio was kept above 1:5 at all times. Nonetheless, if significant, the adjustment of two or more flows could be implemented to keep the overall aqueous flow rate constant. Additionally, within the 110 µm device, multilayer packing allowed us to demonstrate encapsulation of beads around a target number higher than 1 (Figure S5, Supporting Information).

To test deterministic single-bead loading in geometries matching more closely the object size, we fabricated a device of height 70 µm, in which encapsulating more than 2 beads is rare because of the physical constraints (e.g., 13 occurrences of three beads in over 13 000 drops in Figure S9, Supporting Information).

In such geometries, single bead loading could reach 100% fidelity (Figure S9, Supporting Information), although flow instabilities prevented this loading to be stable over long time periods (>10 s). The higher throughput (≈40 Hz versus ≈15 Hz in 110 µm device) obtained using this device for the same flow rates resulted in more frequent flow rate adjustments such that the PD correction previously used did not damp the oscillations as heavily as in the 110 µm device. However, this demonstrates the possibility to autonomously reach optimum one-to-one loading of soft beads. Overall, with devices of different heights, we show that the method can be flexibly applied to devices operating in different regimes (with higher throughput and lower bead occupancy for smaller devices).

2.3.1. Analysis of Damped Oscillations

We could fit the average loading occupancy by modelling the time response with a damped oscillator model. We used the following equation to fit the curves shown in Figure 3C-(i),(iii):

$$y = Ce^{-bt}\sin(wt+\varphi)+D \tag{2}$$

where C, b, ω, ϕ, and D are constants and t is time. The exponential envelope of the damped oscillations is defined by the exponential term b which dictates the strength of the damping. In the examples given in Figure 3C, it was found to be $6.4\,10^{-3}$ s^{-1} versus $12.5\,10^{-3}$ s^{-1} for P and PD corrections, respectively). The fits were closely matching experimental data, with average occupancy (coefficient D) very close to the target occupancy of 1 bead per droplet (0.994 ± 0.002 for Figure 3C-(i), 1.034 ± 0.006 for Figure 3C-(iii)) for 5-min experiments. Natural oscillations were calculated from the coefficient ω to be ≈17 s, indicating a slow time response of the system as observed from the time delay between flow rate adjustments and average loading curves.

2.4. Combinatorial Object Sorting

Integration of powerful object detectors with microfluidic sorting allows for new types of droplet selection not previously possible, using multiclass counting and localization as selection criteria. To demonstrate this, we designed an experiment to showcase selection of multiple discrete objects for co-encapsulation events of exactly one cell and one bead as required in emerging screening techniques such as single-cell sequencing or drug discovery.[24] To this aim, cells and PA beads were co-flown and co-encapsulated in droplets, imaged, classified by YOLOv4-tiny neural networks and subsequently sorted in real-time.

A summary of four 'single-cell-single-bead' sorting experiments performed on three separate dates is displayed in **Table 1**.

Table 1. Summary of the results for four 'single-cell-single-bead' distinct sorting experiments.

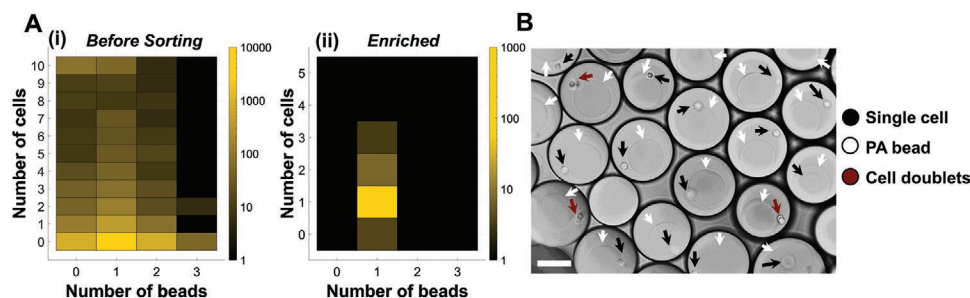| Experiment number	Cell density [10^6 mL^{-1}]	Number of droplets screened	Number of droplets sorted	Average throughput [Hz]	Number of false detections for beads	Number of false detections for cells	Overall percentage true detections
1	1.5	13 423	610	22	0 (0%)	15 (2.5%)	97.5
2	1.5	13 930	1006	14.5	32 (3.2%)	30 (3.0%)	93.8
3	0.15	19 351	115	20.2	2 (1.7%)	5 (4.3%)	93.9
4	0.15	15 382	99	7.3	9 (9%)	2 (2%)	88.9

2101053 (7 of 12)

Figure 4. A) Enrichment of droplets containing exactly one cell and one bead. i) Initial combinations of beads and cells for >13 000 droplets prior to sorting. ii) Enriched droplet population containing exactly a single cell and a single PA bead after sorting by YOLOv4-tiny. B) Example image of sorted droplets containing a single cell (black arrows) and a single bead (white arrows). Red arrows indicate false positive cell doublets. Scale bar: 65 µm.

We exemplify two different regimes for these sorts: at high cell density, we show that we can rapidly accumulate a large sorted fraction while keeping purity high (experiments 1 and 2 have an average 8 single cells per 100 droplets), while at lower cell densities (experiments 3 and 4 have an average 8 single cells per 1000 droplets), we demonstrate high enrichment rates. Each of these experiments included over 13 000 droplets, at throughput ranging from 7.3 to 22 Hz. In all experiments, the sorted droplet fraction showed significant droplet enrichment in the target occupancy.

The multi-object distribution for droplet loading before and after sorting can be displayed through heatmaps as shown in **Figure** 4A-(i),(ii) (which corresponds to experiment 1 in Table 1), visually displaying the change in combinations of cells and beads before/after sorting. All experiments shown implemented the closed-loop feedback mechanism described in the previous section.

The low cell density experiment (0.15 million mL^{-1}) allowed us to check for the ability to enrich for a target droplet population. For instance, in experiment 3 (Table 1), sorting increased the population of single-cell-single-bead droplets from ≈0.65% before sorting to ≈94% when active sorting by deep neural networks was incorporated, representing an overall enrichment factor (ratio of the fractions of single cell + single bead after and before sorting) of 145 in the target droplet population. Only 11 droplets were wrongly classified as waste (i.e., droplets that do meet the sorting criteria) out of a total of 126 amongst 19 351 droplets.

Correct channel sorting was confirmed using a high-speed camera and images of sorted droplets were hand-labelled to check detection accuracy, with strict negative labelling of unclear objects. Strict negative labelling was done by expert judgement to ensure that images classified as true positive are indeed true positives. The overall percentage of true positive events containing 1 cell and 1 bead ranged between 88.9% to 97.5%. The predominant cause for false detections was the presence of cell doublets, especially those where cells were overlapping to the point of near total eclipse at the time of imaging. This could be minimized by more stringent cell straining protocols and addition of additional reagents preventing cellular attachment. Further, oil satellite droplets were sometimes falsely recognized as cells, and cells contacting the droplet interface were occasionally missed. It was also observed that cells and PA beads were sometimes not detected in cases of low contrast. These issues represent a small fraction (a maximum ≈5%) of the total

number of detections in a typical experiment but could be further alleviated by ensuring cleaner microfluidic conditions to prevent oil contaminants in the device, by using fixed optical components to decrease variability in object appearance, and by improving the imaging contrast by use of higher power light sources. It may be possible to further improve the precision of detections by retraining on a data set with additional training examples, especially those with varying contrast and in the presence of contaminants to improve sensitivity.

In practice, extra erroneous droplets may arise from device priming and flow instabilities that lead to droplets transiently leaking into the sorting outlet, lowering the effective enrichment rate. Therefore, we have verified purity obtained by manual labelling of droplets for the sorted droplet fraction corresponding to experiment 2 in Table 1 and found that ≈85% were true positives with most incorrect droplets (determined by reviewing the saved image collection) corresponding to cell doublets (Figure 4B).

2.5. Sorting of Cell Doublets Using Real-Time Relative Object Locations

The long-held view in flow cytometry type experiments is to discard cell doublets as they contribute to blurring of cellular identity and individual cell responses.[25] However, cell doublets or clusters are known to have important functional roles in many biological processes relying on dynamic interactions. For instance, immune-cell pairs are formed in response to viral infection and autoimmune diseases.[26] In contrast to training a model to recognize cell pairs, we have used the localization ability of the object detector to distinguish cells in close proximity to those which are more distant. An example histogram of cell pair distances and representative images of cell pairs are shown in **Figures** 5A and 5B. A distinct peak corresponding to cell-to-cell distance of ≈15 µm can be seen. The additional calculations needed for working out intercell distances increases analysis time for sorted droplets to ≈41 ms in total, reducing the theoretical maximum throughput by a factor of 3 compared to classification and counting alone. We therefore implemented a two-step process of droplet generation followed by re-injection of the droplets using the same sorting device. This enabled flowing droplets at lower rates by dividing the total oil flows twofold in the device compared to in-line generation and sorting. With this strategy, we could sort cell doublets at a
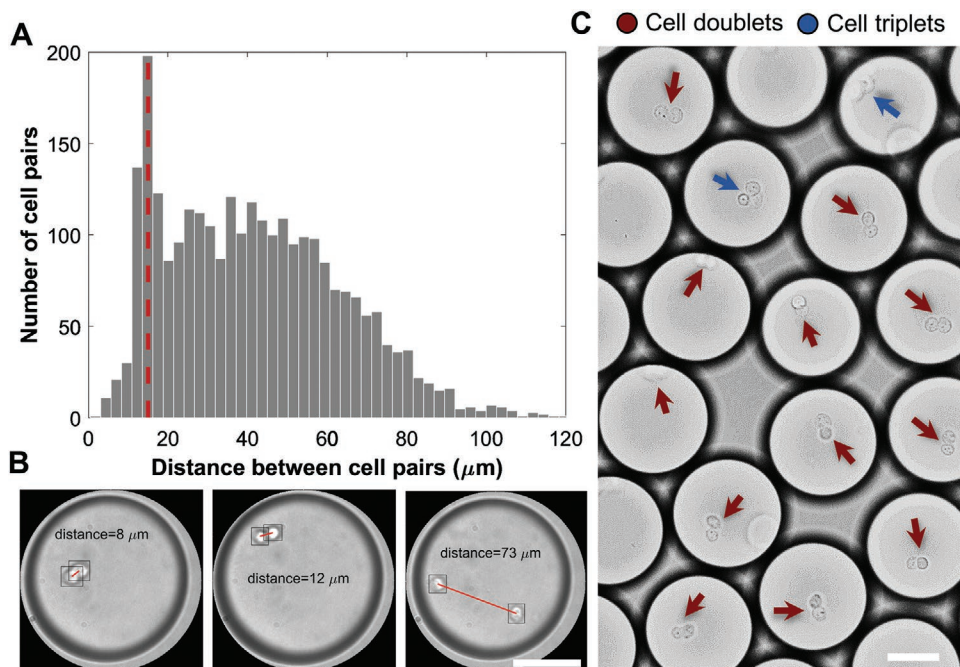
**ADVANCED
SCIENCE NEWS**

www.advancedsciencenews.com

**ADVANCED
MATERIALS
TECHNOLOGIES**

www.advmattechnol.de

Figure 5. Spatially-resolved selections: cell-to-cell distances are used as criterion for droplet sorting. A) Representative distribution of cell pair distances for a total of 2701 cell pairs. A distance of 15 μm (red dashed line), corresponding to the average cell diameter is used as a threshold for sorting. B) Example images of cell pairs with red line representing the calculated distances between the centers of the cell bounding boxes. C) Image of sorted droplets containing mostly cell doublets. Scale bar: 65 μm.

maximum rate ≈10 Hz. We set a cell-to-cell distance threshold equal to 15 μm (the average cell diameter), ensuring the selection of cells in close contact. An image of the sorted fraction for a typical sorting experiment (Figure 5C) shows high purity in cell doublets (≈80% of the sorted population) with most of the incorrect droplets corresponding to cell triplets.

The model selection contained ≈20% cell triplets that were counted as false positives, presumably because the 3D orientation of small cell clusters renders classification difficult, which was consistent with the expected number of misclassified events shown in Figure 2B-(iii). However, our selection shows that the probability for two cells to be close-by but not attached during imaging is very low as no cells were found to be separated after selection.

3. Discussion and Outlook

Modern deep learning object detectors such as YOLOv4-tiny stand to provide powerful capabilities for droplet microfluidics, especially for single-cell sorting applications. Such neural networks can be trained to localize and classify any number of micro-objects without explicit definition and with high adaptability, instead learning their general features from a set of example images. Here, three classes were trained: PA beads, SH-SY5Y cells, and cellular aggregates. Detections were comparable in their precision to hand-labelled images and are robust to experimental variation such as the presence of foreign objects (e.g., oil droplet satellites, dust particles). This enabled the same model to be used on different days without need for drift correction. Use in real-time microfluidic combinatorial experiments highlights the network's ability to label objects in complex environments with many overlapping objects. Object detection with inference of bounding boxes is also useful for confirming correct training of such networks, whereas decisions made by traditional CNN image classifiers cannot be so easily explicated (although interpretative techniques such as integrated gradients, XRAI are emerging).[27]

In **Table 2**, we summarize the most relevant image-activated sorters including the ones implementing neural networks approaches, summarizing the advantages of object detection methods over previously demonstrated workflows.

The speed of object detector CNNs are already shown to be appropriate for integration with high-throughput techniques, but could be pushed significantly further using improved GPU hardware, such as the new Ampere microarchitecture which represents a significant improvement in computational power, with deep learning performance three times higher than the GTX 1080Ti used in this study. Further performance increase could be achieved using TensorRT, NVIDIA's high-performance neural network inference optimizer which converts a TensorFlow graph to a more efficient structure for improved latency, throughput, and efficiency.[29] Recent works have implemented the TensorRT framework giving a ≈5× increase in inference time.[28] Practically, achieving higher sorting rates will provide other experimental challenges, in particular balancing the need for high-contrast, high-resolution images with decreased luminosities at high shutter speeds. This will require a combination of high-power light sources and scan cameras with low acquisition times to reduce motion blur. The ability to trigger images further away from the sorting junction whilst still monitoring sorting will also enable faster sorts up to the theoretical rates of over 100 Hz provided by object detectors such as YOLOv4-tiny.

Table 2. Comparison between YOLOv4-tiny and other image analyses methods from previous studies.

Reference	Type of image analysis	Droplets used in study	Multiclass counting and sorting	Spatially-resolved sorting	Sorting Throughput
Anagnostidis et al.[17]	CNN, 3 convolutional layers	Yes	No	No	40 Hz
Lee et al.[27]	Resnet-18 (CNN, 18 layers)	No	No	No	82.8 Hz
Template-matching Girault et al.[12]	Template-matching	Yes	Yes	No	10 Hz
Isozaki et al.[28]	CNN, 6 convolutional layers	No	No	No	1133 Hz
Sesen et al.[13]	Image thresholding	Yes	Yes	Not shown	4 Hz
This study	YOLOv4-tiny	Yes	Yes	Yes	45 Hz

We have demonstrated that object detectors allow for the precise selection of a droplet subpopulation based on both the number and type of micro-objects encapsulated. In this study, co-encapsulated single-cell and single-beads formed the criteria for model selections, demonstrating the applicability to double Poisson sorting, a configuration used in an increasing number of novel microfluidic-based methods.[6,30] The same detector could be used to sort cell doublets using location information in real-time. enabling cell–cell interactions to be studied label-free and combined with other downstream assays. Further, although only 3 classes were implemented in this study, YOLOv4-tiny can been trained for many independent classes (e.g., 80 classes in the MS COCO dataset), offering potential for further customization of microfluidic experiments.[18] For example, this method could be scaled to a single general model able to recognize a variety of biological objects of interest.[31]

In the future, coupling high-magnification microscopy with multimodal detection (e.g., with fluorescence read-outs), it may be possible to perform cellular phenotyping in blood samples, detect morphological changes during cell differentiation, or analyze intratumor heterogeneity indicative of cancer.[31,32] This offers a broad range of novel possibilities for future studies in analytical biology, with the training of new object detectors limited only by the availability and quality of training datasets and the labor-intensive labelling process.

4. Conclusions

In conclusion, we have demonstrated selection using real-time object detection, including single cells and the targeted isolation of cell doublets based on inter-cell distances. Results obtained with this platform will improve the biological outcome, predictability, and quality of single-cell -omics experiments. They will also pave the way for performing assays interrogating the spatial distributions of micro-objects. Example applications include the screening of antibody libraries against cellular targets using bead display systems, or the quantification of cellular aggregation events.[33,34] Coupled to hardware interfacing and closed-loop feedbacks, we have also demonstrated the potential for such platforms to develop into robust, operator-free machines able to conduct biological experimentation autonomously. This showcases the trend towards generic, adaptable platforms that can self-adjust towards optimal functionality across a range of experimental conditions (e.g., fluid viscosities, temperature changes, etc.).

In the future, low or dual magnification imaging could be used to monitor several areas simultaneously with different target process variables (e.g., cell loading, sorting, adjustment of throughput). Finally, we foresee other areas of research that are enabled by image analyses such as the development and study of biodegradable materials, biofilm formation, the enzymatic degradation of polymer microparticles or rheological studies of soft microgels in microchannels.

5. Experimental Section

Microfluidics Platform for Droplet Imaging and Real-Time Object Detection: The microfluidic device for droplet creation and in-line sorting was used as previously reported and the chip CAD design available online on DropBase (openwetware.org/wiki/DropBase).[17] The device consisted of one inlet for the continuous oil phase and two inlets for the aqueous phases to allow for simultaneous cell and PA bead loading. For the carrier oil, 1% w/w surfactant (008-Fluorosurfactant, Ran Biotechnologies) dissolved in HFE-7500 (3M) was used. Aqueous and carrier phases were loaded into 1 mL glass syringes (SGE) and 5 mL plastic syringes (BD Plastipak), respectively, and their flow rates controlled by pulseless syringe pumps (Nemesys, Cetoni). The PA beads were 65 μm in diameter (Droplet Genomics) and suspended in 10 mM Tris-HCl (pH 8.0), 137 mM NaCl, 2.7 mM KCl, 10 mM EDTA, and 0.1% v/v Triton X-100. For closed-loop feedback, the Nemesys pumps Qmix SDK was used to update the flow rates of syringe pumps using Python control sequences.

Cell Culture: The human neuroblastoma cell line SHSY-5Y (Sigma) was cultured in Dulbecco's Modified Eagle Medium (DMEM, Gibco), supplemented with 10% Fetal Bovine Serum (Gibco), 1% Glutamax (100×, Gibco), and 1% Non-Essential Amino Acids (Gibco) at 37 °C with 5% CO_2. The medium was filtered using a 0.2 μm Sartorius filter prior usage. Cells were passaged at ≈80% confluency using 0.05% trypsin-EDTA (1×, Gibco). For all the experiments, the cells were passaged at ≈80% confluency using 0.05% trypsin-EDTA (1×, Gibco), resuspended in fresh media containing 15% (v/v) OptiPrep density gradient medium (Sigma), counted (density 1.5×10^6 cells mL^{-1}), and passed through a 20 μm cell strainer (Falcon, Fisher Scientific) to minimize cellular aggregation.

Droplet Selection Experiments: For single-cell-single-bead sorting experiments, the flow rates used were 1 μL min^{-1} for aqueous solutions and 12 μL min^{-1} for the carrier oil phases. Droplets were subsequently respaced by an additional insertion of the same carrier oil to better frame individual droplets and facilitate in-line sorting.

For cell doublet selections, droplets were generated using the microfluidic sorting device (Figure 1) and collected in a reservoir 0.5 mL PCR tube (Eppendorf). For generation, the flow rates for the oil, cells, and second aqueous phase (PA bead supernatant) were 30, 3, and 3 μL min^{-1}, respectively. The droplets were then collected for 10 min and

**ADVANCED
SCIENCE NEWS**

www.advancedsciencenews.com

**ADVANCED
MATERIALS
TECHNOLOGIES**

www.advmattechnol.de

re-injected into the sorting device through the bead inlet (Figure 1B) at a flow rate of 2 µL min^{-1}. Oil was flown in the main carrier oil inlet at 12 µL min^{-1} for respacing the droplets.

A dedicated optical setup built around an inverted trinocular microscope (SP-98-I, Brunel) allowed for the bright-field imaging of microdroplets in real time, capturing close-cropped images of individual droplets using a triggered acquisition approach as previously reported (Figure 1A).[17] The input of a photodetector (PDA36A, Thorlabs) was processed by a low latency Field-Programmable Gate Array (FPGA) controller, which in turn outputs a 5 V trigger to the fast scan CCD camera (Pike F-032B, Allied Vision).

Images of individual droplets are immediately fetched, and a single NVIDIA (GTX 1080Ti) GPU was used to evaluate the YOLOv4-tiny model frame by frame in the Python 3.7.0 API of TensorFlow 2.3.0rc0, running on a Windows 7, 64-bit operating system with an Intel i5-6500 3.2 GHz processor and 32 GB DDR4 RAM. If desired, images and bounding box information was also saved to an SSD drive. The resulting classification was compared to a chosen sorting criterion, with a probability threshold to activate sorting set at 0.25. If the logical condition for sorting is met, a USB data acquisition card (DAQ, USB-6009, National Instruments) board relays into a function generator which creates a 10 kHz rectangular signal at 8 Vpp for 20 ms. This pulse is amplified 100-fold, and the electric field propagates through a saturated 5 M salt solution flowed into dedicated electrode channels close to the sorting junction. This non-linear field causes selected droplets to be deflected towards the sorting channel via dielectrophoresis, while unsorted droplets proceed to the waste.[35] A high-speed camera (Miro ex4, Vision Research) was used to confirm correct sorting.

Training YOLOv4-Tiny to Detect Micro-Objects: YOLOv4-tiny neural networks were trained on custom datasets obtained from microfluidic experiments. A stratified split was then used to create the training set (70%), validation set (15%), and test set (15%). This ensured subsets contained an equal number of examples from each class, shuffling the images, and selecting randomly from each stratum to avoid bias. A total of 1400 images were used in training, 300 for network validation, and 300 reserved for testing, which were not seen by the network during training. In the 1400 test images, a total of 1571 beads, 2134 cells, and 317 aggregates were present; as many images contained more than a single object. In addition, images with foreign objects such as dust particles or oil microdroplets, as well as different focus and contrast were deliberately included to create a model which is robust to varying experimental conditions and the presence of contaminants. YOLOv4-tiny training was conducted using Python 3.7.10 on a cloud virtual machine provided by Google Colaboratory. A single GPU (Tesla T4 GPU, NVIDIA) was configured with CUDA Toolkit 11.0, cuDNN 7.6.5, and the Darknet framework built from YOLOv4-tiny implementation for Windows.[18] To train 3 classes, the network was constructed with a batch size 64 and subdivisions of 16, in addition to an image height and width of 416 pixels with 1 color channel, max batches of 6000, and 24 filters for each convolutional layer sequentially before each of the 3 YOLO layers. To increase the rate of convergence, pre-trained weights were used as a starting point for transfer learning. The YOLO network was then trained over a total of 6000 iterations and the weights saved every 100 iterations. The curve for training loss and corresponding validation precision is shown in Figure S1, Supporting Information. Finally, the optimum network weights were converted to a TensorFlow protocol buffer file with both the graph definition and model weights needed to run detections on the local machine connected to the microscope rig.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Keywords

[1] S. Sohrabi, N. Kassir, K. M. Moraveji, *RSC Adv.* **2020**, *10*, 27560.
[2] S. Sarkar, W. Kang, S. Jiang, K. Li, S. Ray, E. Luther, A. R. Ivanov, Y. Fu, T. Konry, *Lab Chip* **2020**, *20*, 2317.
[3] P. Mair, F. Gielen, F. Hollfelder, *Curr. Opin. Chem. Biol.* **2017**, *37*, 137.
[4] A. M. Thompson, A. L. Paguirigan, J. E. Kreutz, J. P. Radich, D. T. Chiu, *Lab Chip* **2014**, *14*, 3135.
[5] H. Kleine-Brüggeney, L. D. Van Vliet, C. Mulas, F. Gielen, C. C. Agley, J. C. Silva, A. Smith, K. Chalut, F. Hollfelder, *Small* **2019**, *15*, 1804576.
[6] B. Demaree, C. L. Delley, H. N. Vasudevan, C. A. C. Peretz, D. Ruff, C. C. Smith, A. R. Abate, *Nat. Commun.* **2021**, *12*, 1583.
[7] A. M. Klein, L. Mazutis, I. Akartuna, N. Tallapragada, A. Veres, V. Li, L. Peshkin, D. A. Weitz, M. W. Kirschner, *Cell* **2015**, *161*, 1187.
[8] D. J. Collins, A. Neild, A. deMello, A. Q. Liu, Y. Ai, *Lab Chip* **2015**, *15*, 3439.
[9] J. F. Edd, D. Di Carlo, K. J. Humphry, S. Koster, D. Irimia, D. A. Weitz, M. Toner, *Lab Chip* **2008**, *8*, 1262.
[10] H. S. Moon, K. Je, J. W. Min, D. Park, K. Y. Han, S. H. Shin, W. Y. Park, C. E. Yoo, S. H. Kim, *Lab Chip* **2018**, *18*, 775.
[11] A. R. Abate, C. H. Chen, J. J. Agresti, D. A. Weitz, *Lab Chip* **2009**, *9*, 2628.
[12] M. Girault, H. Kim, H. Arakawa, K. Matsuura, M. Odaka, A. Hattori, H. Terazono, K. Yasuda, *Sci. Rep.* **2017**, *7*, 40072.
[13] M. Sesen, G. Whyte, *Sci. Rep.* **2020**, *10*, 8736.
[14] E. Zang, S. Brandes, M. Tovar, K. Martin, F. Mech, P. Horbert, T. Henkel, M. T. Figge, M. Roth, *Lab Chip* **2013**, *13*, 3707.
[15] D. A. Basiji, *Principles of Amnis Imaging Flow Cytometry, Imaging Flow Cytometry*, Humana Press, New York, NY **2016**, p. 13.
[16] J. G. Burel, M. Pomaznoy, C. S. L. Arlehamn, G. Seumois, P. Vijayanand, A. Sette, B. Peters, *Cytometry, Part A* **2020**, *97*, 1127.
[17] V. Anagnostidis, B. Sherlock, J. Metz, P. Mair, F. Hollfelder, F. Gielen, *Lab Chip* **2020**, *20*, 889.
[18] A. Bochkovskiy, C.-Y. W. , H.-Y. M. Liao, *arXiv* **2020**, abs/2004.10934.
[19] R. Girshick, *IEEE Int. Conf. Comput. Vision (ICCV)* **2015**, 1440.
[20] D. F. Crawford, C. A. Smith, G. Whyte, *Sci. Rep.* **2017**, *7*, 10545.
[21] N. Wang, R. Liu, N. Asmare, C. H. Chu, O. Civelekoglu, A. F. Sarioglu, *Lab Chip* **2021**, *21*, 1916.

[22] R. Z. Lin, H. Y. Chang, *Biotechnol. J.* **2008**, *3*, 1172.

[23] D. Tzutalin, *LabelImg. Git Code.*, **2015**, https://github.com/tzutalin/labelImg.

[24] F. Lan, B. Demaree, N. Ahmed, A. R. Abate, *Nat. Biotechnol.* **2017**, *35*, 640.

[25] A. Cosma, *Cytometry A* **2020**, *97*, 768.

[26] S. C. Bendall, *Nat. Biotechnol.* **2020**, *38*, 559.

[27] K. Lee, S. E. Kim, J. Doh, K. Kim, W. K. Chung, *Lab Chip* **2021**, https://doi.org/10.1039/D0LC00747A.

[28] A. Isozaki, H. Mikami, H. Tezuka, H. Matsumura, K. Huang, M. Akamine, K. Hiramatsu, T. Iino, T. Ito, H. Karakawa, Y. Kasai, Y. Li, Y. Nakagawa, S. Ohnuki, T. Ota, Y. Qian, S. Sakuma, T. Sekiya, Y. Shirasaki, N. Suzuki, E. Tayyabi, T. Wakamiya, M. Xu, M. Yamagishi, H. Yan, Q. Yu, S. Yan, D. Yuan, W. Zhang, Y. Zhao, F. Arai, R. E. Campbell, C. Danelon, D. Di Carlo, K. Hiraki, Y. Hoshino, Y. Hosokawa, M. Inaba, A. Nakagawa, Y. Ohya, M. Oikawa, S. Uemura, Y. Ozeki, T. Sugimura, N. Nitta, K. Goda, *Lab Chip* **2020**, *20*, 2263.

[29] F. Abdurahman, K. A. Fante, M. Aliy, *BMC Bioinformatics* **2021**, *22*, 112.

[30] X. Zhang, T. Li, F. Liu, Y. Chen, J. Yao, Z. Li, Y. Huang, J. Wang, *Mol. Cell* **2019**, *73*, 130.

[31] F. Michor, K. Polyak, *Cancer Prev. Res.* **2010**, *3*, 1361.

[32] J. Zhang, X. Wei, R. Zeng, F. Xu, X. Li, *Future Sci. OA* **2017**, *3*, FSO187.

[33] L. Diamante, P. Gatti-Lafranconi, Y. Schaerli, F. Hollfelder, *Protein Eng., Des. Sel.* **2013**, *26*, 713.

[34] A. M. White, Y. Zhang, J. G. Shamul, J. Xu, E. A. Kwizera, B. Jiang, X. He, *Small* **2021**, *17*, 2100491.

[35] K. Ahn, C. Kerbage, T. P. Hunt, R. M. Westervelt, D. R. Link, D. A. Weitz, *Appl. Phys. Lett.* **2006**, *88*, 024104.

2101053 (12 of 12)