

Oil Spill Segmentation via Adversarial f -Divergence Learning

Xingrui Yu, He Zhang, Chunbo Luo, *Member, IEEE*, Hairong Qi, *Fellow, IEEE*,
and Peng Ren, *Senior Member, IEEE*

Abstract—We develop an automatic oil spill segmentation method in terms of f -divergence minimization. We exploit f -divergence for measuring the disagreement between the distributions of ground-truth and generated oil spill segmentations. To render tractable optimization, we minimize the tight lower bound of the f -divergence by adversarial training a regressor and a generator, which are structured in different forms of deep neural networks separately. The generator aims at producing accurate oil spill segmentation, while the regressor characterizes discriminative distributions with respect to true and generated oil spill segmentations. It is the co-play between the generator net and the regressor net against each other that achieves a minimal of the maximum lower bound for the f -divergence. The adversarial strategy enhances the representational powers of both the generator and the regressor and avoids requesting large amounts of labelled data for training the deep network parameters. In addition, the trained generator net enables automatic oil spill detection that does not require manual initialization. Benefiting from the comprehensiveness of f -divergence for characterizing diversified distributions, our framework can accurately segment variously shaped oil spills in noisy SAR images. Experimental results validate the effectiveness of the proposed oil spill segmentation framework.

Index Terms—Oil spill segmentation, f -divergence minimization, adversarial learning, synthetic aperture radar (SAR) image processing.

I. INTRODUCTION

OIL spills, which are leaked from ships or drilling platforms, may cause disastrous consequences to the environment and social economy. Synthetic aperture radar (SAR) is regarded as a powerful tool for observing environment and targets [31][15][30] due to its advantages of all-weather and all-time operations. SAR also provides an important means for monitoring marine oil spills [9][29]. Detecting oil spills through SAR images timely is vital for damage assessment and oil spread control. In this paper, we aim at developing intelligent algorithms for automatically segmenting marine oil spill regions from SAR images.

This work was supported in part by National Natural Science Foundation of China (Project No. 61671481), Qingdao Applied Fundamental Research (Project No. 16-5-1-11-jch), and The Fundamental Research Funds for the Central Universities under Project 18CX05014A. (Corresponding Author: Peng Ren).

X. Yu and P. Ren are with College of Information and Control Engineering, China University of Petroleum (East China), Qingdao 266580, China (e-mail: xingruiyu_upc@hotmail.com; pengren@upc.edu.cn).

H. Zhang and C. Luo are with College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter EX4 4QF, UK (e-mail: hz298@exeter.au.uk; c.luo@exeter.ac.uk).

H. Qi is with College of Engineering, The University of Tennessee, Knoxville 37996, USA (e-mail: hqi@utk.edu).

In the literature, the majority of oil spill analysis based on SAR data lies in the investigation of the physical characteristics of oil spills for reflecting different types of electromagnetic waves. One general method for oil spill observation via SAR is based on the non-Bragg scattering phenomena caused by oil spills. The capillary and short gravity waves give rise to Bragg scattering that is sensed by SAR. On the other hand, the oil spills on the ocean surface damp out the Bragg scattering, resulting in dark patches in the SAR images. The non-Bragg scattering regions provide indications for observing the oil spill regions in SAR images.

Lately, more sophisticated oil and electromagnetic wave relationships have been explored, and especially the polarimetric characteristics of oil spills have been comprehensively investigated. In this regards, researchers including Migliaccio et al. [32][4], Ricci et al. [3][2], Minchew et al. [20][6] and Brekke et al. [6][8] have conducted studies that represent state-of-the-art oil spill observation research based on polarimetric SAR data. The polarimetry based strategies enhance oil spill observations in images through polarimetric analysis such that basic image processing techniques such as thresholding [25] and K -means clustering [4] are normally applied to detect oil spills in the enhanced representations. State of the art methods [7] [19] in the geoscience and remote sensing literature focus on investigating various physics based features which enable comprehensive observations of oil spills.

On the other hand, researchers mainly from the image processing and machine learning community have started working on developing more sophisticated oil spill segmentation methods for accurately detecting oil spill regions in images. Most of the recent image processing techniques for oil spill segmentation are formulated in terms of energy minimization. An energy functional measures oil spill fitness and similarity characteristics such that it guides the oil spill contours to evolve towards minimum energy. Xia et al. [33] developed a modified continuous energy functional and employed level sets for detecting oil spill contours. Mdakane et al. [18] incorporated a region-based signed pressure force function into a continuous energy function for detecting oil slicks from moving vessels. Ren et al. presented a dual smooth higher order energy function and applied the graph cut algorithm to obtain optimal segmentation [27]. These strategies directly perform the segmentation procedure according to the oil intrinsic characteristics and do not require training a model. However, it is observed that their segmentation accuracy heavily relies on the prior knowledge in terms of manual labellings for initializing the energy minimization in segmen-

tation [34]. Though manual initialization in terms of coarsely observing and capturing oil spill regions is commonly accepted in practice, it is neither efficient nor reliable for accurate segmentation. In order to avoid the manual initialization in the segmentation procedure and render an automatic oil spill segmentation scheme, a strategy that trains a segmentation model based on labelled images and infers the oil spill regions without requiring manual initialization in the segmentation procedure is expected.

In this paper, we present a deep learning oriented oil spill segmentation framework in terms of minimizing the f -divergence between the ground-truth and the model generated segmentations. The f -divergence minimization is achieved by adversarial training a regressor and a generator. The adversarial strategy is highly motivated by generative adversarial network (GAN) [10] and its variants [21] [23], which train two models to play against each other for increasing model representational power. However, there exists a significant difference between GAN and our work. GAN approximates the target sample distribution via Kullback-Leibler divergence. In contrast, our framework characterizes oil spill distributions in terms of the general f -divergence that enables more comprehensive modeling capability. This merit makes our framework easily address irregular oil spill shapes as well as adverse influences caused by various noises in SAR images. Furthermore, we exploit the relationship between f -divergence and surrogate loss [22] and develop an efficient computation method for minimizing the f -divergence in the task of oil spill segmentation. Specifically, we exploit two different deep neural networks for implementing the generator and the regressor that adversarial optimize the minimal of the tight lower bound of the f -divergence. In the adversarial learning process, the generator aims to produce accurate oil spill segmentation for a given SAR image. The regressor is fed by either ground-truth or generated oil spill segmentation and regresses variational representations for the segmentation distributions. In this scenario, generated segmentations tend to exhibit large f -divergence. At the convergence of the adversarial training, the generator has the ability of fooling the regressor to give small f -divergence and is qualified to produce accurate oil spill segmentation for SAR images. Our method provides an automatic segmentation approach which can be considered as a post processing procedure after the oil spill observation. Therefore, we believe that our method provides an effective post-processing procedure for the state of the art physics based oil spill observation strategies [19] [32] [4].

There are several reasons for us to develop the adversarial training strategy rather than trying to directly minimize f -divergence. First, the adversarial training which characterizes the tight lower bound of f -divergence renders a more tractable computation scheme than straightforward f -divergence minimization. Second, the surrogate loss exploited in the adversarial training has the ability of characterizing various types of divergences and thus increases the model comprehensiveness. More importantly, the adversarial trained generator can be employed to perform automatic oil spill segmentation without manual initialization. Last but not the least, benefiting the

adversarial strategy, training our model requires only a small collection of ground-truth segmented oil spill data, which provides an economical scheme for real world oil spill segmentations.

The main contributions of this paper are summarized as follows:

- We propose a novel oil spill segmentation method in terms of minimizing the f -divergence between the ground-truth and generated segmentations, which not only extends the theoretical implication of f -divergence but also exhibits robustness in practice.
- We formulate the objective in terms of minimizing the lower bound of f -divergence based on surrogate loss, which results in a tractable optimization scheme for segmenting irregular oil spills in noisy SAR images.
- We develop a deep learning strategy that adversarial learns the f -divergence and establishes an oil spill segmentation framework without manual initialization.

Furthermore, experimental evaluations validate that our framework outperforms the state-of-the-art GAN and level set method and achieves accurate segmentation in the situations of irregular oil spill shapes and noise interferences.

II. f -DIVERGENCE MINIMIZATION FOR OIL SPILL SEGMENTATION

In this section, we commence by introducing the principle of f -divergence based oil spill segmentation. Then, we describe how to reformulate it via surrogate loss. Finally, we exploit deep neural networks for developing variational representations for f -divergence.

A. f -Divergence for Oil Spill Segmentation

In this subsection, we formulate oil spill segmentation in terms of f -divergence minimization. We aim to train a segmentation model to minimize the distributional disagreement between the generated oil spill segmentation and the ground-truth. In this scenario, we formulate the oil spill segmentation problem as that of minimizing an f -divergence, in which the f -divergence measures the probabilistic difference between the generated segmentation and the ground-truth.

Supposing oil spill segmentations are distributed in the domain Ω , we denote P_S and $P_{\hat{S}}$ as the probabilistic distributions of the ground-truth segmentation S and the generated segmentation \hat{S} , respectively. The f -divergence (1) between P_S and $P_{\hat{S}}$ is given as follows:

$$D_f(P_S \| P_{\hat{S}}) = \int_{\Omega} f\left(\frac{p_S(x)}{p_{\hat{S}}(x)}\right) p_{\hat{S}}(x) dx, \quad (1)$$

where x is a distribution variable representing one segmentation sample, and $f(\cdot)$ is a function relating to a specific type of distribution divergence. $p_S(x)$ and $p_{\hat{S}}(x)$ are density functions of segmentation distribution P_S and $P_{\hat{S}}$, respectively.

Our goal is to train a model that produces the generated segmentation \hat{S} which is as accurate as the true segmentation S for a SAR image. Specifically, in the training procedure, the model is optimized by SAR images along with their

segmentation maps in terms of the f -divergence (1). In the segmentation procedure, given an unknown SAR image, the trained model is expected to automatically generate accurate oil spill segmentation.

B. Reformulating f -Divergence via Surrogate Loss

The f -divergence provides a comprehensive characterization of distribution disagreement because various divergence types can be adopted in (1). However, straightforwardly minimizing the f -divergence is always intractable. The reason for this shortcoming is two-fold. First, the segmentation domain, which may exhibit arbitrary forms, is difficult to characterize in practice. Second, even if some tricky technique enables computing one type of divergence, it does not guarantee that various divergences arising from the general f -divergence can be addressed in the same way.

To render an effective and tractable solution, we turn to minimizing the tight lower bound of f -divergence. This is effected by formulating the divergence function f (2) in terms of a surrogate loss function ϕ as follows:

$$f(u) := -\inf_v [\phi(-v) + \phi(v)u]. \quad (2)$$

Here u and v are two variational variables related to the distribution variable x . In the light of (1), the variable u has a straightforward relationship with the distribution variable x such that $u = \frac{p_S(x)}{p_{\hat{S}}(x)}$. Additionally, we characterize the relationship between v and x via deep neural networks, which will be presented in the next subsection.

The employment of surrogate loss provides a different route for achieving the f -divergence minimization. Specifically, we represent the f -divergence by taking the surrogate loss characterized $f(\cdot)$ in (1), and have the derivations as follows:

$$\begin{aligned} & D_f(P_S || P_{\hat{S}}) \\ &= \int_{\Omega} f\left(\frac{p_S(x)}{p_{\hat{S}}(x)}\right) p_{\hat{S}}(x) dx \\ &= \int_{\Omega} \left\{ -\inf_v \left[\phi(-v) + \phi(v) \frac{p_S(x)}{p_{\hat{S}}(x)} \right] \right\} p_{\hat{S}}(x) dx \quad (3) \\ &\geq -\inf_v \left[\int_{\Omega} \phi(-v) p_{\hat{S}}(x) dx + \int_{\Omega} \phi(v) p_S(x) dx \right] \\ &= -\inf_v \mathbb{E}_{x \sim P_S} [\phi(v)] + \mathbb{E}_{x \sim P_{\hat{S}}} [\phi(-v)]. \end{aligned}$$

The inequality follows Jensen's inequality [5] and characterizes the lower bound of the f -divergence.

We then reformulate the f -divergence minimization into a lower bound minimization problem as follows:

$$\min_{P_{\hat{S}}} \left\{ -\min_v \mathbb{E}_{x \sim P_S} [\phi(v)] + \mathbb{E}_{x \sim P_{\hat{S}}} [\phi(-v)] \right\}, \quad (4)$$

where we replace variational infimum sign \inf by the minimization sign \min for practical implementation. We further remove the negative sign by replacing the second \min by \max in (4) and reformulate the objective (4) into

$$\min_{P_{\hat{S}}} \max_v \mathbb{E}_{x \sim P_S} [\phi(v)] + \mathbb{E}_{x \sim P_{\hat{S}}} [\phi(-v)]. \quad (5)$$

We observe that the \max operation in (5) practically constrains the lower bound obtained in (3) to be tight and confirms a rigorous condition for minimizing f -divergence.

The f -divergence minimization in terms of (5) is achieved in a two-fold manner. First, we seek the tighter lower bound of the f -divergence in the \max -step. Then we minimize the obtained lower bound in the \min -step. It is through iterating the \max and \min operations that achieves the minimal of the f -divergence.

C. Deep Neural Nets for Variational Approximation

In this subsection, we describe how to develop deep neural networks (DNNs) for approximating the variational variables v of the oil spill segmentation f -divergence introduced in Section II-B. Specifically, we establish two DNNs representing a generator G and a regressor R , separately. Given an oil spill SAR image I , G aims at mapping it to its segmentation \hat{S} . The regressor R regresses a variational representation for a ground-truth or generated segmentation distribution.

Two basic components for constructing the DNNs are the decoder (Fig. 1 (a)) and the encoder (Fig. 1 (b)). An encoder consists of a convolutional layer, a batch normalization layer [12] and an LReLU activation layer [16]. A decoder consists of a deconvolutional layer [35], a batch normalization layer and an LReLU activation layer. The detailed architectures of the generator G and the regressor R are illustrated in Fig. 1 (c) and (d), respectively.

The generator G takes the SAR image I as input. As illustrated in Fig. 1 (c), the architecture of G is adapted from that in [26] and reformed in the shape of "U-Net" [28]. It is an n encoder-decoder stacked network with skip connections [11] added between mirrored layers (i.e., the encoder i and the decoder $n - i$). In this case, more low-level features are sent from encoders to decoders, which improves the reconstruction ability of the decoders and the representation ability of DNNs. The generator G produces a segmentation map \hat{S} for I as the final output.

The regressor R takes the original SAR image I and its segmentation map (either ground-truth or generated) as inputs. The regressor R is a convolutional neural network composed of convolutional layers followed by one LReLU layer and three encoders. The regressor architecture is finalized by a Tanh layer after the convolutional layer, and results in a variational representation for a segmentation distribution as the ultima output.

We use the generator net G and regressor net R thus obtained to characterize the relationship between variational variable v in the surrogate loss (2) and the segmentation distribution variable x in (1). In order to regularize the unconstrained output of DNNs and fully explore the representational power of our framework, we exploit a transformation function $a_f(\cdot)$ for scaling the variational representation regressed from R . We use θ_r and θ_g to denote the parameter sets for R and G , respectively, and hence reframe (5) as follows:

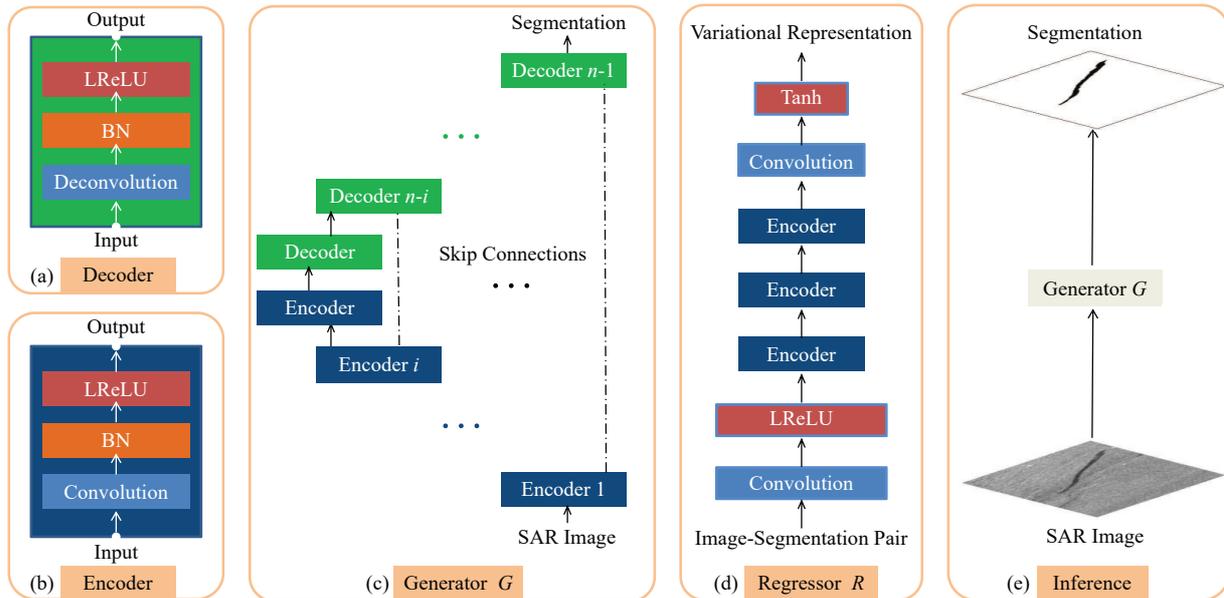


Fig. 1. DNN architectures and inference procedure. The basic components of our framework are (a) decoder and (b) encoder. The decoder is composed of a deconvolutional layer followed by a BN layer and an LReLU layer. The encoder (b) shares a similar architecture with the decoder, with the deconvolutional layer being replaced by the convolutional layer. (c) Generator is composed of several encoders and decoders, with the dashed-dotted line indicating the skip connections. (d) Regressor is composed of a convolutional layer and three encoders followed by one convolutional layer and one tanh layer. The regressor is fed by a SAR image and its segmentation to generate the variational representation of the segmentation map. (e) At the inference stage, given an unknown SAR image, the trained generator G tries to produce an accurate oil spill segmentation map without manual initialization.

TABLE I
THE OPTIONS OF f -DIVERGENCE FUNCTION $f(\cdot)$, SURROGATE LOSS $\phi(\cdot)$ AND TRANSFORMATION $a_f(\cdot)$.

f -divergence	$f(u)$	$\phi(v) (v > 0)$	$a_f(w)$
Pearson χ^2 (P)	$(u - 1)^2$	$-\frac{3}{4}v^2 + 4v - \frac{8}{3}$	$-w$
Squared Helinger (SH)	$(\sqrt{u} - 1)^2$	$e^v - 1$	$1 - e^{-w}$
Total Variation (TV)	$\frac{1}{2} u - 1 $	e^{-v}	$\frac{1}{2} \tanh(w)$
Capacity Discrimination Distance (CDD)	$u \log(u) - (u + 1) \log(u + 1)$	$\log(1 + e^{-v})$	$-\log(1 + e^{-w})$
Symmetric Kullback-Leibler (SKL)	$u \log(u) - \log(u)$	$e^{-v} - v - 1$	$-\log(1 + e^{-w})$

$$\min_{\theta_g} \max_{\theta_r, \theta_d} \mathbb{E}_{S \sim P_S, I \sim P_I} [\phi(a_f(R(S, I)))] + \mathbb{E}_{I \sim P_I} [\phi(-a_f(R(G(I), I)))] \quad (6)$$

The original GAN [10] structure consists of a generator and a discriminator. The GAN and our method share the same architecture for the generator. On the other hand, the discriminator of GAN and our regressor have different structures and play different roles. One major distinction is that the discriminator ends up with a sigmoid layer and the output of our regressor is scaled by the transformation function $a_f(\cdot)$. The advantages of our method are two-fold. Firstly, the sigmoid layer confines the discriminator to the discrete labeling scenario, and in contrast our regressor characterizes continuous representations and enables more comprehensive similarity measure. Secondly, our regressor is comprehensively trained subject to the f -divergence such that the regressor output indicates the lower bound of f -divergence. The intrinsic relationship between f -divergence and surrogate loss described in [22] enables various specific distribution divergences to be employed for

training our DNNs. The function $\phi(\cdot)$ takes an exact form results in a specific surrogate loss function. The surrogate loss function along with a specific transformation function $a_f(\cdot)$ concretizes the f -divergence. The relationships regarding a concrete f -divergence and its corresponding $f(\cdot)$, $\phi(\cdot)$ and $a_f(\cdot)$ in specific forms are listed in Table I. This provides a more general probabilistic characterization than the GAN that only admits the Kullback-Leibler divergence, which is just one specific case in the family of f -divergence. In this scenario, our framework is a general strategy that addresses a broad category of oil spill segmentation tasks with respect to various probabilistic divergence measurements.

III. ADVERSARIAL TRAINING FOR MINIMIZING f -DIVERGENCE

We first introduce the basic principles for characterizing f -divergence by adversarial training the regressor and the generator. We then describe how to practically implement the training procedures.

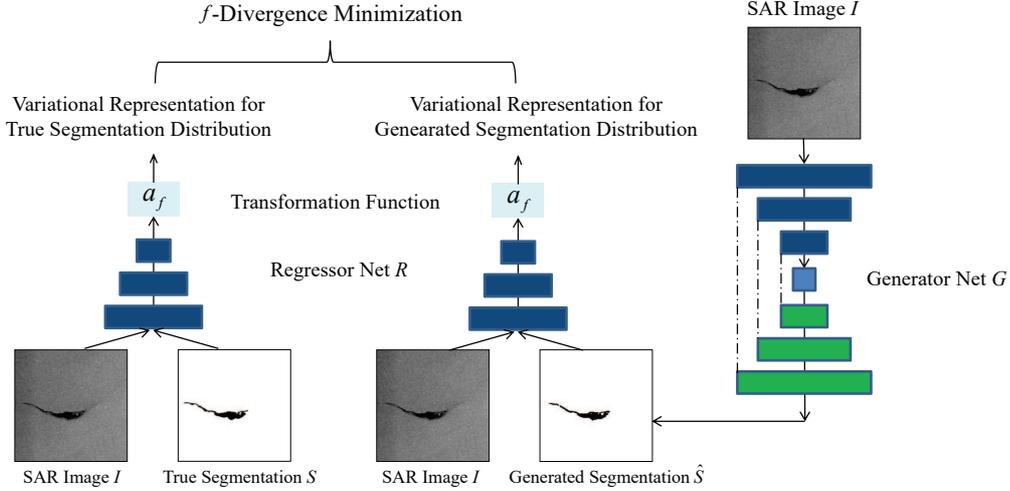


Fig. 2. The pipeline of the proposed oil spill segmentation framework via f -divergence minimization.

Algorithm 1: The training procedure of the proposed oil spill segmentation method.

Input: A training dataset consisting of original SAR images and their corresponding ground-truth oil spill segmentation maps.

Output: The trained parameter set θ_g for G .

for number of training iterations **do**

1) Min-step optimization:

for k steps **do**

- Sample a minibatch of m examples $\{(I_1, S_1), \dots, (I_m, S_m)\}$ from the training dataset;
- Update the regressor by ascending its stochastic gradient as: $\theta_r \leftarrow \theta_r + \eta \nabla_{\theta_r} \mathcal{L}_R$.

end

2) Max-step optimization:

- Sample another minibatch of m examples $\{(I_1, S_1), \dots, (I_m, S_m)\}$ from the training dataset;
- Update the generator by descending its stochastic gradient as: $\theta_g \leftarrow \theta_g - \eta \nabla_{\theta_g} \mathcal{L}_G$.

end

A. Adversarial Training Strategy

The regressor R and the generator G play different roles in optimizing the objective (6). The requirement for the regressor R is that the regressed variational representations for generated segmentations are supposed to exhibit smaller f -divergence than those for ground-truth. Therefore, R plays a role of discriminating the segmentation generated by G from ground-truth as much as possible in the max operation. This max operation also enables the identification of the tight lower bound for the f -divergence. Whereas, the generator G tries its best to fool R by generating the segmentation map \hat{S} as similar as possible to the true segmentation map S for I . Therefore, G aims at minimizing the f -divergence.

It is obvious that the minimization of the tight lower bound of the f -divergence gives rise to an adversarial training strategy for the regressor net R and the generator net G . It is the co-play between R and G against each other that achieves a minimal of the maximum lower bound for the f -divergence. The adversarial training strategy provides a tractable manner for minimizing the f -divergence which normally exhibits in a sophisticated form. Furthermore, it

enhances the representational powers of both the generator and the regressor and avoids requesting large amounts of labelled data for training the deep network parameters. This provides a practical advantage for oil spill segmentation, because it is difficult to access a large number of oil spill SAR images for model training.

The overall pipeline of training the DNNs via optimizing (6) is illustrated in Fig. 2. In contrast to the energy minimization based segmentation methods such as level sets [14] and graph cuts [24], the trained generator net enables automatic oil spill detection that does not require manual initialization. Detailed implementations of the adversarial training are presented in the following subsection.

B. Implementations of the Adversarial Training

In (6), the maximization operates with respect to both R and G , whereas the minimization operates with respect to just G . We tackle the objective (6) in terms of the *max*-step as:

$$\begin{aligned} \max_{\theta_r, \theta_g} \mathbb{E}_{S \sim P_S, I \sim P_I} [\phi(a_f(R(S, I)))] \\ + \mathbb{E}_{I \sim P_I} [\phi(-a_f(R(G(I), I)))] \end{aligned} \quad (7)$$

and the *min*-step as:

$$\min_{\theta_g} \mathbb{E}_{S \sim P_S, I \sim P_I} [\phi(-a_f(\|G(I) - S\|_1))], \quad (8)$$

where $\|\cdot\|_1$ is the L_1 norm utilized to further penalize the dissimilarity between the generated segmentation $G(I)$ and the ground-truth S .

We discard R and replace the original $\phi(-a_f(R(G(I), I)))$ by $\phi(-a_f(\|G(I) - S\|_1))$ in (8) for two reasons. First, the gradients transferred from R decrease dramatically in back propagation and thus tend to be insufficient to train G well in practice. Second, in order to enhance the generative power of G , training G by comparing it with ground-truth segmentation is much more effective than that interfered with R . Therefore, it is ineffective to train G as the input of R and we use (8) instead as in the *min*-step.

Given a finite training set, we approximate the expectations in (7) and (8) by using a minibatch of samples. The loss functions of the regressor and the generator are given by

$$\mathcal{L}_R = \frac{1}{m} \sum_{i=1}^m [\phi(a_f(R(S_i, I_i))) + \phi(-a_f(R(G(I_i), I_i)))], \quad (9)$$

and

$$\mathcal{L}_G = \frac{1}{m} \sum_{i=1}^m [\phi(-a_f(\|G(I_i) - S_i\|_1))], \quad (10)$$

where m is the minibatch size.

In this case, we alternatively optimize the deep network parameters of R and G with respect to their loss functions until convergence. In contrast to most deep learning methods that are effected by training models based on labelled big data, our framework does not require a large number of labelled segmentation maps for training the f -divergence model. This advantage is meaningful especially in the field of automatic oil spill segmentation, where the ground-truth data are very limited. In practice, we validate that around 20 ground-truth segmentation maps for oil spill image patches are enough for training an effective f -divergence model. The capability of training with small data benefits from the adversarial strategy which trains the generator to produce a number of generated segmentation maps. The generated segmentation maps in fact augment the size of training dataset and increase the data variability. Therefore, the adversarial strategy enables training a comprehensive f -divergence model with a small number of labelled oil spill segmentation maps. The detailed training procedure of our segmentation framework is given in Algorithm 1.

IV. EXPERIMENTAL RESULTS

In this section, experimental evaluations of different oil spill segmentation methods are performed on both synthetic SAR images and real SAR images. We empirically compare the proposed segmentation method with its closely related GAN [10], neural networks and the state-of-the-art initialization dependent method [14]. All experiments are performed on a PC server with a NVIDIA Tesla K80 GPU and 64GB memory.

A. Experimental Settings

1) *Parameters*: We use the same hyper-parameter setting for evaluating GAN and the proposed oil spill segmentation method. We train the DNN models using Adam optimizer [13] with $\beta_1 = 0.5$, $\beta_2 = 0.999$. The learning rate is fixed as $\eta = 0.0002$ and the minibatch size is set as 1.

2) *Evaluation Metrics*: In our experiments, both the generated oil spill segmentation maps and the ground-truth ones are binary images, in which zero value pixels indicate the oil spills. To quantitatively evaluate the performance of alternative oil spill segmentation methods, we compute the *accuracy* of the generated segmentation map \hat{S} as follows:

$$\text{Accuracy}(\hat{S}) = \frac{|\hat{S} \cap S \cap \mathbb{O}|_{card}}{|S|_{card}},$$

where $|\cdot|_{card}$ denotes an operator that computes the cardinal of a set. \mathbb{O} is a set composed of zero values, and shares the same cardinal with S .

In addition, the performance of the generated oil spill region \hat{S} is also evaluated in terms of the *region fitting error* (RFE) [17] as follows:

$$\text{RFE}(\hat{S}) = \frac{|\hat{S} \cup S|_{card} - |\hat{S} \cap S|_{card}}{|S|_{card}}.$$

In this case, a smaller RFE indicates a more accurate segmentation. In particular, when the generated segmentation tightly follows the ground-truth, the RFE is close to zero.

B. Segmentation on Synthetic SAR Images

To evaluate the segmentation methods with respect to different interferences, we perform experimental comparisons between GAN and the proposed method on synthetic SAR images with various oil spill shapes and noises.

Synthetic SAR images are of 256×256 pixels and obtained by contaminating original images with various noises. The original image is a clean synthetic image with no noise. In this experiment, three original images containing one, two and three objects are employed separately to simulate irregular (i.e., continuous and/or discontinuous) oil spill regions.

To approximately simulate the noises that exist in real world SAR images, we adopt two types of noise, i.e. (a) the salt and pepper additive noise and (b) weighted multiplicative noise, for interfering original images. (a) For generating salt and pepper noisy synthetic SAR images, a proportion ρ_1 of pixels are randomly selected and their intensities are reset to be random values in the interval $[0, 255]$. The parameter ρ_1 is set to be $\{0.0001, 0.0005, 0.01, 0.05, 0.1, 0.2, 0.3, 0.4\}$, separately, for generating training images and testing images. (b) Let I_p denote an original image. The multiplicative noisy synthetic SAR images are simulated in terms of $I_p + \rho_2 \times P(z) \times I_p$, where ρ_2 is the weight parameter and $P(z)$ is the probabilistic distribution of Gamma, Rayleigh or Log-normal (as given in Table II) based noises. The weight parameter ρ_2 is set to be $\{0.03, 0.05, 0.07, 0.09\}$ and $\{0.02, 2, 200\}$ for generating training images and testing images, respectively. Specifically, we employ Gamma biased noise under $j = 1.0$ and $\theta = 1.0$,

TABLE II
PROBABILISTIC DISTRIBUTIONS OF MULTIPLICATIVE NOISES ($E[\cdot]$ AND $Var[\cdot]$ ARE THE EXPECTATION AND VARIANCE).

Distribution	Probabilistic density function	Parameters
Gamma	$P(z) = z^{j-1} \frac{e^{-\frac{z}{\theta}}}{\theta^j \Gamma(j)}$	$j = \frac{E^2[z]}{Var[z]}, \theta = \frac{E[z]}{Var[z]}$
Rayleigh	$P(z) = \frac{z}{\sigma^2} e^{-\frac{z^2}{2\sigma^2}}$	$\sigma^2 = \frac{2}{4-\pi} Var[z]$
Log-normal	$P(z) = \frac{1}{\sigma z \sqrt{2\pi}} e^{-\frac{(\ln(z)-\mu)^2}{2\sigma^2}}$	$\mu = \ln\left(\frac{E^2[z]}{\sqrt{Var[z]+E^2[z]}}\right), \sigma^2 = \ln\left(\frac{Var[z]}{E^2[z]} + 1\right)$

Rayleigh biased noise under $\sigma = 1$, Log-normal biased noise under $\mu = 0$ and $\sigma = 0.5$, separately. We totally generate 63 multiplicative noisy synthetic SAR images, with 36 images for training and the rest for testing. We formulate the training objectives in terms of all f -divergences in Table I, and train our framework with the parameters designed in the subsection IV-A1.

We commence by qualitatively visualizing the segmentation results in different noisy situations. In Fig. 3, we show the segmentation results of synthetic SAR images with different salt and pepper noise. Specifically, the images containing one, two and three oil spill regions under varying salt and pepper noises are illustrated in Fig. 3 (a), (b) and (c), respectively. For both GAN and the proposed method, the segmentation results get worse with the increasing proportion of contaminated pixels. Especially in Fig. 3, GAN cannot accurately segment both continuous and discontinuous oil spill regions under heavy salt and pepper noise.

We then quantitatively evaluate the performance of GAN and the proposed segmentation method on the synthetic SAR images. For salt and pepper noise, the accuracy and RFE in terms of varying proportion ρ_1 are shown in Fig. 4 and Fig. 5, respectively. Though our f -divergence framework overwhelmingly outperforms the GAN segmentation method in the additive noisy scenario, there are slightly performance differences across the specific divergences in our framework, which can be observed from the zoomed out curves in Fig. 4 (b) and Fig. 5 (b). Figs. 4 (a) and 5 (a) reveal that the TV divergence are the inferior among all alternatives within f -divergence for segmenting one oil spill region. One possible reason for this comparative ineffectiveness arises from the computation of absolute value (see Table I). Compared with the two and three oil spill region scenarios, the one oil spill region exhibits a comparatively more continuous form. The absolute value in the TV divergence results in abrupt changes when the variational variable is around one, and thus cannot properly characterize the one continuous oil spill region. Furthermore, Figs. 4 (b), (c) and 5 (b), (c) reveal that the CDD divergence performs worst among all alternatives within f -divergence for segmenting two and three oil spill regions. One major reason for this comparative shortcoming is that CCD has a very complicated form (see Table I), which may give rise to overfitting in learning oil spill regions. Therefore, in the two and three oil spill region scenarios, the discontinuous and complicated regions may be over learned in terms of the CCD divergence such that biases arise in segmenting new SAR

images.

The detailed accuracy and RFE of segmenting synthetic noisy SAR images containing one, two and three regions are given in Table III and Table IV. It is observed that our framework (based on different specific divergences) achieves better performance (both in terms of accuracy and RFE) than GAN in the cases of one oil spill region. Especially, TV and SKL achieve the best performance for all noise types, with approximately 3% accuracy increase and 35% RFE decrease. For the two oil spill region cases, our models are better than GAN, e.g. P results in 15% accuracy increase and 25% RFE decrease for the Gamma noise. Although our method and GAN have indiscriminative accuracy performance in three oil spill region cases, our method is superior to GAN in terms of RFE, with decreases for Gamma noise ranging from 25.5% to 33%, for Rayleigh noise ranging from 46.5% to 50.7%, and for Log-normal noise ranging from 49% to 51%. The weight parameter ρ_2 has little effect on the performance for different noises and oil spill regions. This reveals that our framework is robust with respect to different specific divergences in the multiplicative noisy scenario.

The overall segmentation performance of GAN is inferior to our proposed method in terms of accuracy and RFE. For both methods, the accuracy decreases and the RFE increases as the increasing of noise proportion ρ_1 and ρ_2 . On the other hand, the proposed method is more robust than GAN in various noisy situations.

C. Segmentation on Real SAR Images

In this subsection, we evaluate the proposed oil spill segmentation method on real SAR images with VV polarization in NOWPAP database¹. We mainly use three types of SAR images in this experiment, i.e. C-band SAR images from ERS-1, ERS-2 satellites and C-band ASAR images from Envisat-1 satellite. These images containing different types of oil spills are captured in separate time by different sensors. The sources of SAR images and properties of sensors collected the data are illustrated in Table V and Table VI, respectively. In the two tables, “-” means unknown information.

From the SAR images in the NOWPAP database, we extract SAR image patches that cover oil spill regions. One training sample is a pair of a SAR image patch and its corresponding ground-truth segmentation. We randomly select 20 SAR images for training and use the rest 23 images for testing. SAR

¹<http://cearac.poi.dvo.ru/en/db/>

TABLE III
ACCURACY OF SEGMENTING SYNTHETIC SAR IMAGES WITH WEIGHTED MULTIPLICATIVE NOISES.

Synthetic SAR Image with		One Oil Spill Region			Two Oil Spill Regions			Three Oil Spill Regions		
ρ_2		0.02	2	200	0.02	2	200	0.02	2	200
Gamma	GAN	0.9313	0.9347	0.9321	0.8422	0.8508	0.8516	0.8471	0.8443	0.8442
	P	0.9565	0.9558	0.9566	0.9560	0.9556	0.9561	0.8781	0.8809	0.8820
	SH	0.9554	0.9548	0.9522	0.9349	0.9343	0.9341	0.8665	0.8798	0.8778
	TV	0.9570	0.9569	0.9570	0.9363	0.9361	0.9346	0.8869	0.8879	0.8877
	CDD	0.9564	0.9560	0.9566	0.9336	0.9335	0.9338	0.8795	0.8812	0.8814
	SKL	0.9559	0.9563	0.9567	0.9311	0.9328	0.9329	0.8537	0.8623	0.8629
Rayleigh	GAN	0.9353	0.9361	0.9362	0.8628	0.8606	0.8597	0.8481	0.8544	0.8548
	P	0.9560	0.9556	0.9560	0.9337	0.9327	0.9344	0.8781	0.8809	0.8785
	SH	0.9551	0.9541	0.9547	0.9354	0.9353	0.9354	0.8807	0.8814	0.8820
	TV	0.9568	0.9561	0.9564	0.9356	0.9353	0.9357	0.8799	0.8824	0.8815
	CDD	0.9564	0.9560	0.9566	0.9336	0.9335	0.9338	0.8795	0.8812	0.8814
	SKL	0.9567	0.9560	0.9562	0.9338	0.9338	0.9338	0.8735	0.8768	0.8792
Log-normal	GAN	0.9292	0.9328	0.9334	0.8610	0.8610	0.8622	0.8483	0.8499	0.8489
	P	0.9561	0.9558	0.9565	0.9339	0.9345	0.9348	0.8855	0.8861	0.8865
	SH	0.9547	0.9526	0.9534	0.9357	0.9363	0.9360	0.8872	0.8874	0.8877
	TV	0.9570	0.9569	0.9570	0.9363	0.9361	0.9364	0.8869	0.8879	0.8877
	CDD	0.9573	0.9575	0.9571	0.9346	0.9341	0.9341	0.8871	0.8879	0.8875
	SKL	0.9574	0.9573	0.9575	0.9345	0.9342	0.9342	0.8839	0.8850	0.8821

TABLE IV
RFE OF SEGMENTING SYNTHETIC SAR IMAGES WITH WEIGHTED MULTIPLICATIVE NOISES.

Synthetic SAR Image with		One Oil Spill Region			Two Oil Spill Regions			Three Oil Spill Regions		
ρ_2		0.02	2	200	0.02	2	200	0.02	2	200
Gamma	GAN	0.0173	0.0166	0.0194	0.0348	0.0345	0.0364	0.0337	0.0343	0.0321
	P	0.0120	0.0108	0.0125	0.0268	0.0267	0.0264	0.0220	0.0225	0.0240
	SH	0.0125	0.0115	0.0128	0.0272	0.0276	0.0281	0.0245	0.0225	0.0236
	TV	0.0117	0.0112	0.0116	0.0269	0.0267	0.0279	0.0235	0.0220	0.0229
	CDD	0.0122	0.0115	0.0125	0.0266	0.0258	0.0273	0.0217	0.0202	0.0208
	SKL	0.0118	0.0110	0.0118	0.0260	0.0263	0.0265	0.0251	0.0212	0.0227
Rayleigh	GAN	0.0182	0.0174	0.0174	0.0309	0.0320	0.0313	0.0363	0.0299	0.0324
	P	0.0122	0.0118	0.0114	0.0262	0.0273	0.0272	0.0194	0.0200	0.0211
	SH	0.0121	0.0115	0.0111	0.0266	0.0278	0.0270	0.0186	0.0197	0.0210
	TV	0.0117	0.0111	0.0117	0.0260	0.0265	0.0270	0.0193	0.0198	0.0203
	CDD	0.0121	0.0117	0.0115	0.0268	0.0263	0.0262	0.0179	0.0186	0.0204
	SKL	0.0112	0.0109	0.0109	0.0257	0.0270	0.0264	0.0182	0.0192	0.0202
Log-normal	GAN	0.0168	0.0181	0.0185	0.0309	0.0320	0.0313	0.0352	0.0329	0.0343
	P	0.0111	0.0112	0.0114	0.0265	0.0268	0.0260	0.0179	0.0186	0.0185
	SH	0.0115	0.0115	0.0115	0.0268	0.0259	0.0258	0.0179	0.0191	0.0191
	TV	0.0112	0.0113	0.0113	0.0264	0.0264	0.0259	0.0179	0.0185	0.0184
	CDD	0.0115	0.0115	0.0114	0.0257	0.0259	0.0253	0.0171	0.0177	0.0178
	SKL	0.0111	0.0111	0.0110	0.0252	0.0256	0.0255	0.0179	0.0181	0.0178

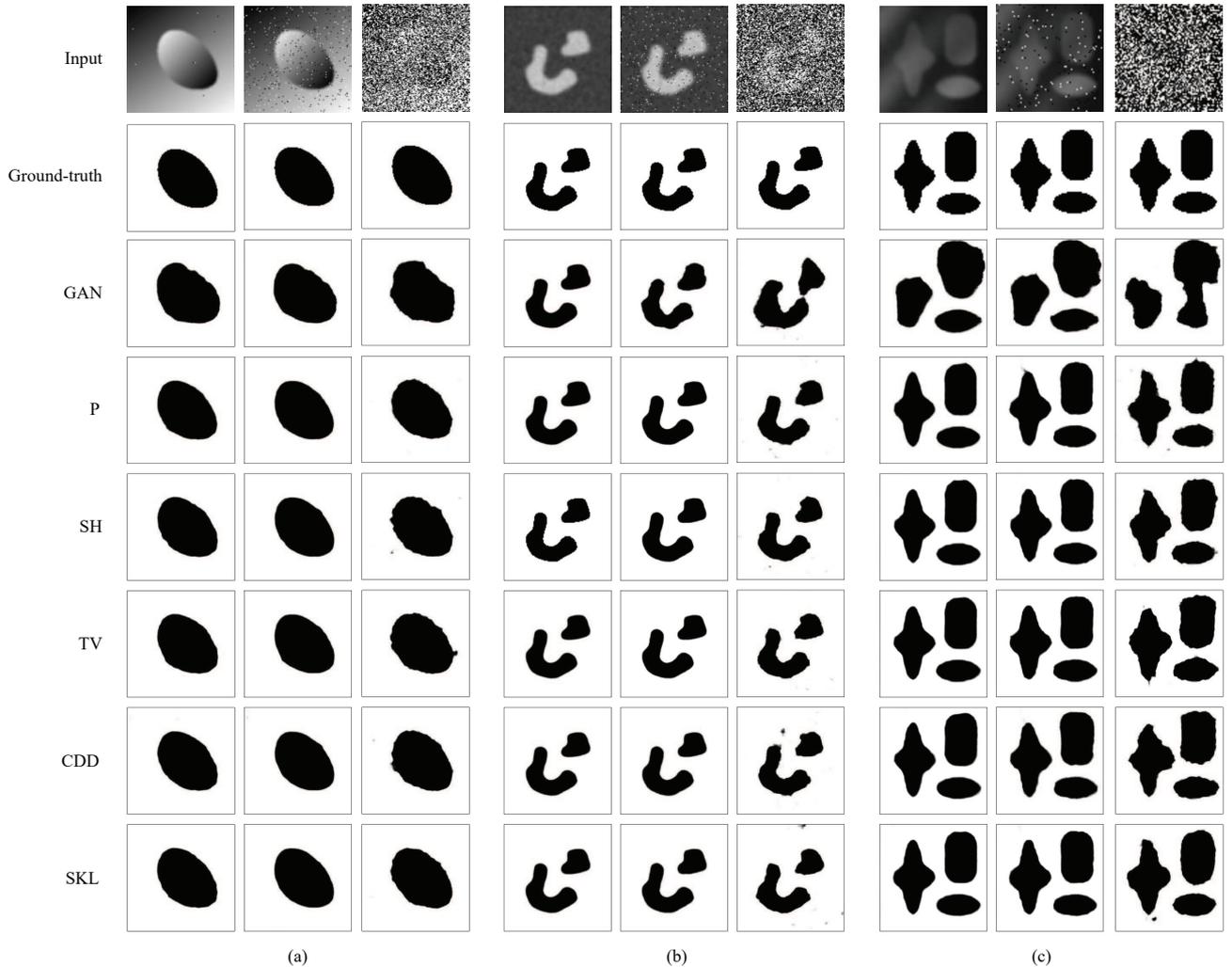


Fig. 3. Oil spill segmentation on synthetic SAR images contaminated by different salt and pepper noises. (a)–(c) Images containing one, two, and three oil spill regions, respectively. For each kind of synthetic SAR image, the noise proportion ρ_1 varies in terms of $\{0.0001, 0.05, 0.4\}$.

TABLE V
SAR IMAGES FROM THE NOWPAP DATABASE.

Capture Time	Sensor	Type of Oil Spills
08.11.1993 01:46:10	ERS-1 SAR	-
19.06.1995 02:30:12	ERS-1 SAR	-
02.09.1996 02:00:55	ERS-2 SAR	-
20.07.1997 02:14:26	ERS-2 SAR	Ship Spill
22.09.1997 02:00:22	ERS-2 SAR	-
27.09.1999 02:01:50	ERS-2 SAR	-
17.12.1997 01:57:42	ERS-2 SAR	-
30.09.2000 02:05:08	ERS-2 SAR	-
15.08.2007 13:04:01	Envisat ASAR	Ship Spill

images in different sizes are resized into 256×256 and the pixel values are normalized into $[-1, 1]$ before training.

We train the proposed method by minimizing all f -divergences listed in Table I and compare the segmentation performance with GAN. In Fig. 6, we show the segmentation

TABLE VI
SAR SENSORS.

Sensor	Wavebands	Spatial Resolution	Noise Floor
ERS-1,2 SAR	C-band	$30 \text{ m} \times 30 \text{ m}$ (3 looks)	-
Envisat ASAR	C-band	$150 \text{ m} \times 150 \text{ m}$	-20~22 dB

results on four real SAR images. For each SAR image, five different f -divergences (i.e. P, SH, TV, CDD and SKL) based segmentation models produce more accurate segmentation than GAN. GAN generates the segmentation maps with larger incorrect areas (marked by dotted boxes) than our method.

To quantitatively evaluate the performance of GAN and the proposed segmentation method on real SAR images, we provide comparison of accuracy and RFE in Table VII and Table VIII, respectively. The overall performance of the proposed method is superior to GAN in terms of both accuracy and RFE. For segmenting continuous regions, the P divergence outperforms other divergences, while the SH divergence performs inferior results. For segmenting discontinuous oil spills,

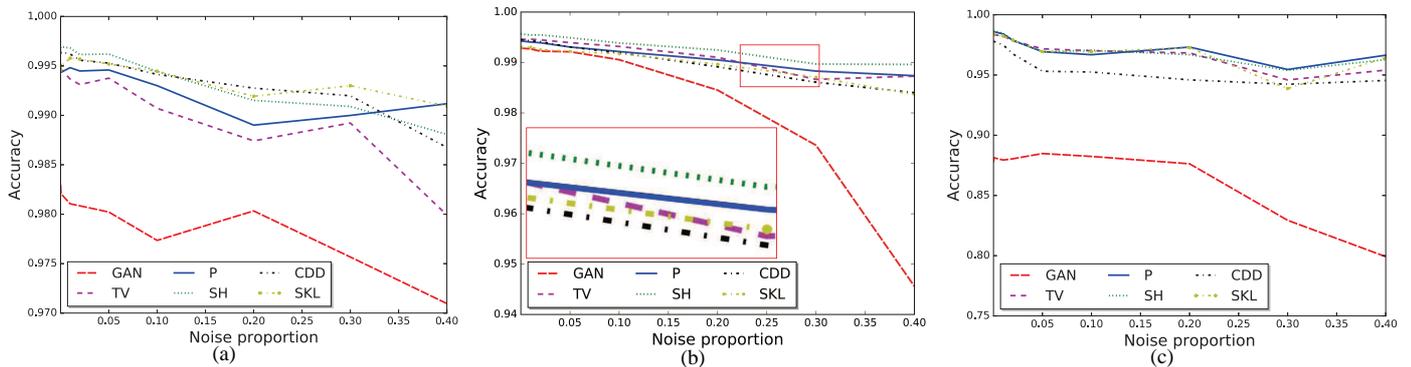


Fig. 4. ACC of segmenting synthetic SAR images with respect to varying salt and pepper noises. The noise proportion ρ_1 is set to be $\{0.0001, 0.0005, 0.01, 0.05, 0.1, 0.2, 0.3, 0.4\}$. Figures (a), (b) and (c) represent the synthetic images containing one, two and three oil spill regions, respectively. The performance differences between different f -divergence forms are comparatively little, which can be observed from the zoomed out curves (red solid box in (b)).

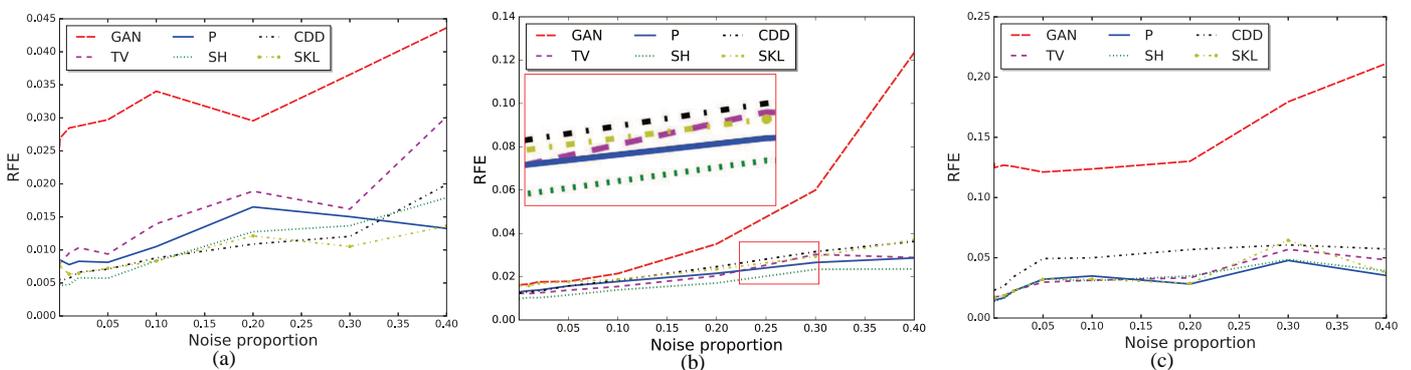


Fig. 5. RFE of segmenting synthetic SAR images with respect to varying salt and pepper noises. The noise proportion ρ_1 is set to be $\{0.0001, 0.0005, 0.01, 0.05, 0.1, 0.2, 0.3, 0.4\}$. Figures (a), (b) and (c) represent the synthetic images containing one, two and three oil spill regions, respectively. The performance differences between different f -divergence forms are comparatively little, which can be observed from the zoomed out curves (red solid box in (b)).

TABLE VII
ACCURACY OF SEGMENTING OIL SPILLS FROM SAR IMAGES IN FIG. 6.

	Fig. 6 (a)	Fig. 6 (b)	Fig. 6 (c)	Fig. 6 (d)
GAN	0.9958	0.9942	0.9988	0.9863
P	0.9984	0.9958	0.9991	0.9883
SH	0.9983	0.9952	0.9989	0.9867
TV	0.9985	0.9954	0.9988	0.9881
CDD	0.9980	0.9947	0.9990	0.9882
SKL	0.9984	0.9959	0.9990	0.9883

TABLE VIII
RFE OF SEGMENTING OIL SPILLS FROM SAR IMAGES IN FIG. 6.

	Fig. 6 (a)	Fig. 6 (b)	Fig. 6 (c)	Fig. 6 (d)
GAN	0.2908	0.2447	0.1363	0.2490
P	0.1107	0.1781	0.0940	0.2128
SH	0.1164	0.2014	0.1252	0.2410
TV	0.1071	0.1941	0.1293	0.2163
CDD	0.1361	0.2211	0.1087	0.2137
SKL	0.1107	0.1735	0.1099	0.2120

the P, TV and SKL divergences are more suitable, whereas the CDD divergence is inapplicable.

We further evaluate the proposed method on segmenting wild images, which are obtained from different sources rather than the training image source. We evaluate the performance of our trained model (with P divergence) on several wild SAR images, as shown in Fig. 7: one C-band Envisat ASAR image (Fig. 7 (a)), with VV polarization and a spatial resolution of 150 m; two airborne L-band UAVSAR images (Fig. 7 (b) (c)), with HH, VV and HV polarization and a spatial resolution of 6 m; and a Space Shuttle X-SAR image (Fig. 7 (d)), with VV polarization and a spatial resolution of 30 m. The segmentation results are illustrated in Fig. 7. From Fig. 7, we observe that our model trained on C-band VV polarization images from the NOWPAP database is able to generate accurate segmentation for the C-band Envisat ASAR image (Fig.7 (a)), L-band data from UAVSAR (Fig.7 (b) (c)) and Space Shuttle X-SAR data from the NOWPAP database (Fig.7 (d)). In addition, our model trained on the single-polarization data is able to generate segmentation for multi-polarization data (Fig.7 (b) (c)).

We also evaluated the trained model on MODIS images. Fig. 8 shows the segmentation results on four MODIS images using our trained model (with P divergence). Although our method is only trained on single-polarization SAR images, our method is able to segment oil spill regions in images

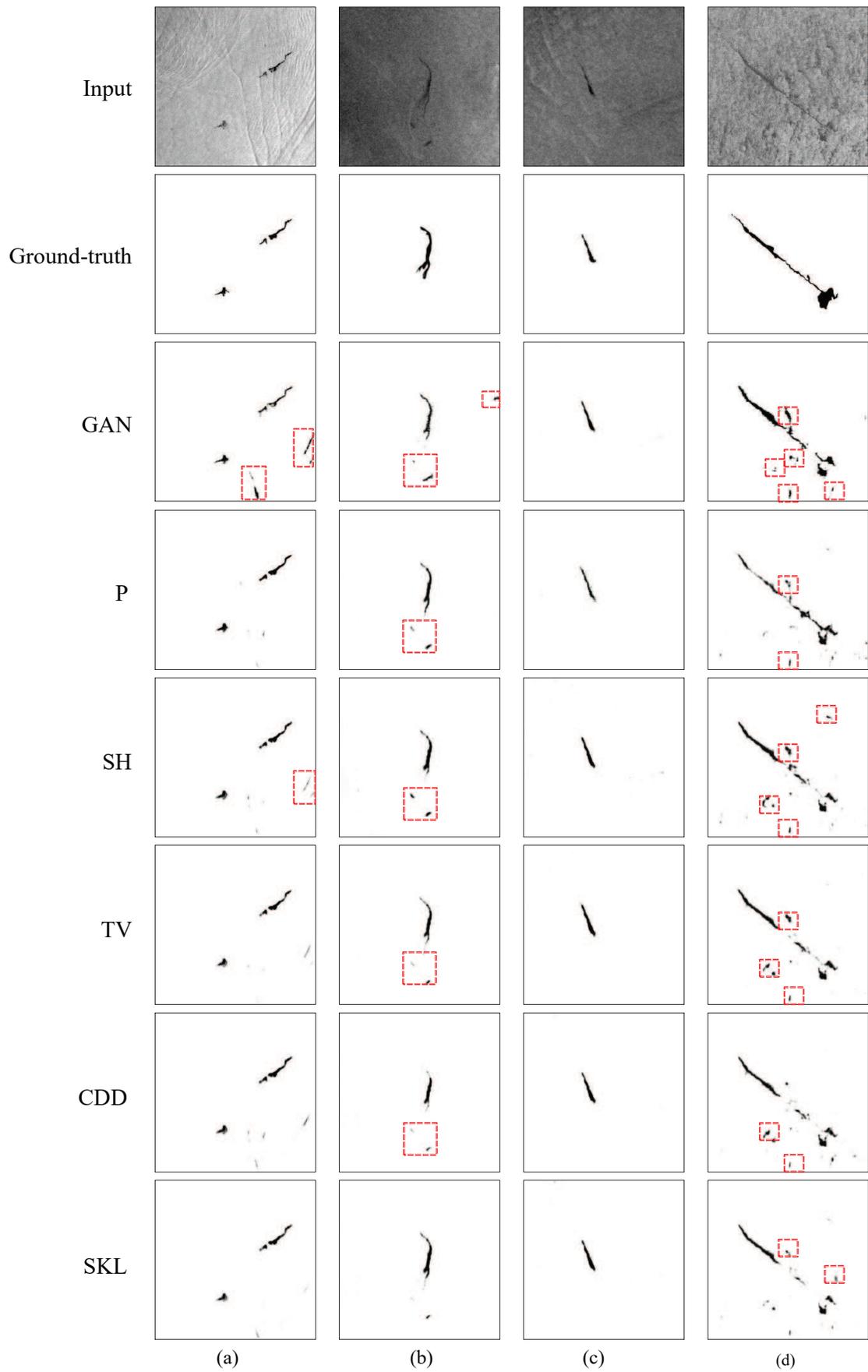


Fig. 6. Oil spill segmentation on real SAR images using learning based methods. The incorrect segmentation areas are marked by dashed boxes.

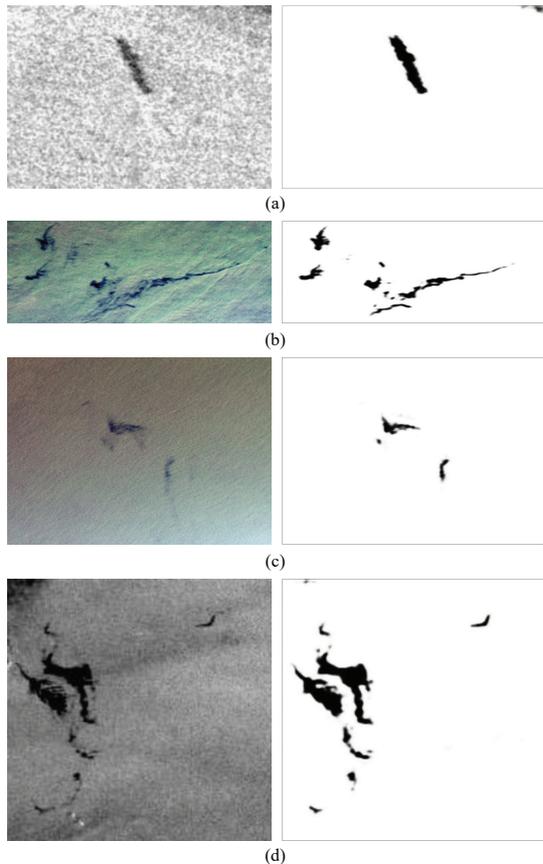


Fig. 7. Segmentation results on wild SAR images using the proposed method with P divergence. (a) is a ASAR image from Envisat-1 satellite, with a size of 220×154 pixels. (b) (c) are airborne L-band images from UAVSAR, with a size of 1800×700 and 2800×2000 pixels respectively. (d) is a Space Shuttle X-SAR image, with a size of 180×190 pixels.

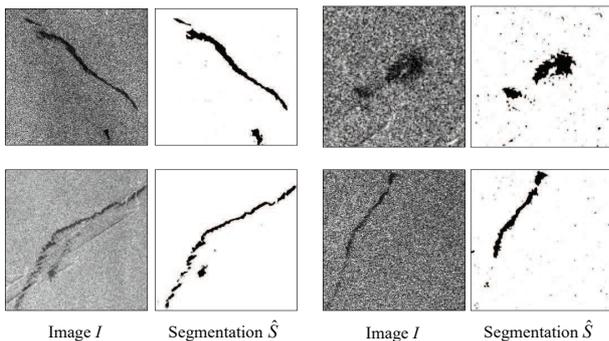


Fig. 8. Segmentation results on wild MODIS images using the proposed method with P divergence.

with different polarization mode (both single-polarization and multi-polarization), which are captured by different sensors (e.g. SAR, ASAR, UAVSAR, X-SAR and MODIS) in different bands (C-band and L-band). In addition, our method is able to generate segmentation for images in different spatial resolution, from a lower resolution (150 m of ASAR, and 30 m of SAR and X-SAR) to a higher resolution (6 m of UAVSAR).

To have a better understanding of the convergence property of GAN and the proposed segmentation method, we show

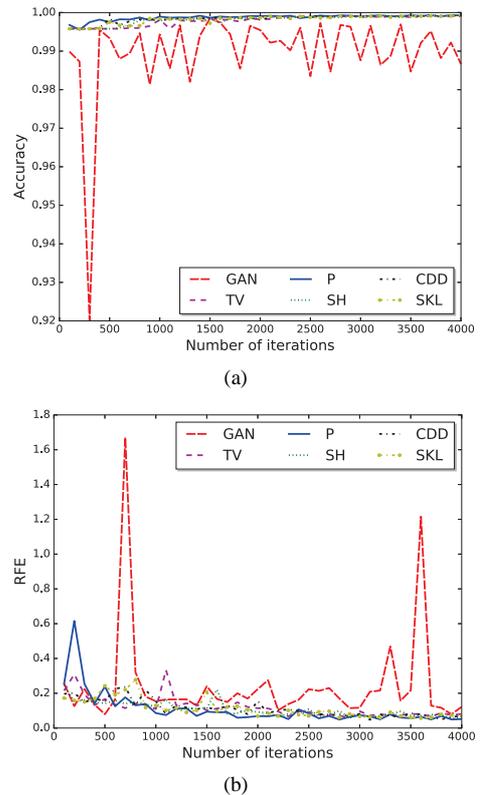


Fig. 9. The accuracy and RFE training curves of different segmentation methods with respect to the number of iterations (x-axis).

the training curves of accuracy and RFE with respect to the number of training iterations in Fig. 9 (a) and (b), respectively. It can be seen that the accuracy and RFE curves of training GAN vary drastically at the early stage and fluctuate slightly during the convergence progress. On the contrary, the curves of our method reflect the learning and converging process are more stable and robust. The stable and efficient convergence for training our framework benefits from the comprehensiveness of f -divergence. As pointed in [1], the Kullback-Leibler and Jensen-Shannon divergences used in GAN are not capable of characterizing nonoverlapping distributions. Therefore, turbulence inevitably occurs in training GAN when distributions vary from separate to overlapping. However, our f -divergence framework characterizes more comprehensive divergences that effectively avoid the shortcomings of Kullback-Leibler and Jensen-Shannon divergences. Therefore, our f -divergence framework provides a more stable and efficient training process.

D. Comparison with Neural Networks

In this part, we evaluate the performance of the proposed method by comparing its segmentation results with that of the classic neural networks. The classic method trains a neural network to segment oil spills by optimizing a L_1 loss between the generated segmentation \hat{S} and the ground-truth segmentation S , i.e. $\|\hat{S} - S\|$. We use two different types of neural networks, one of which has the same architecture as the generator (Fig. 1 (c)), and the other is a much simpler

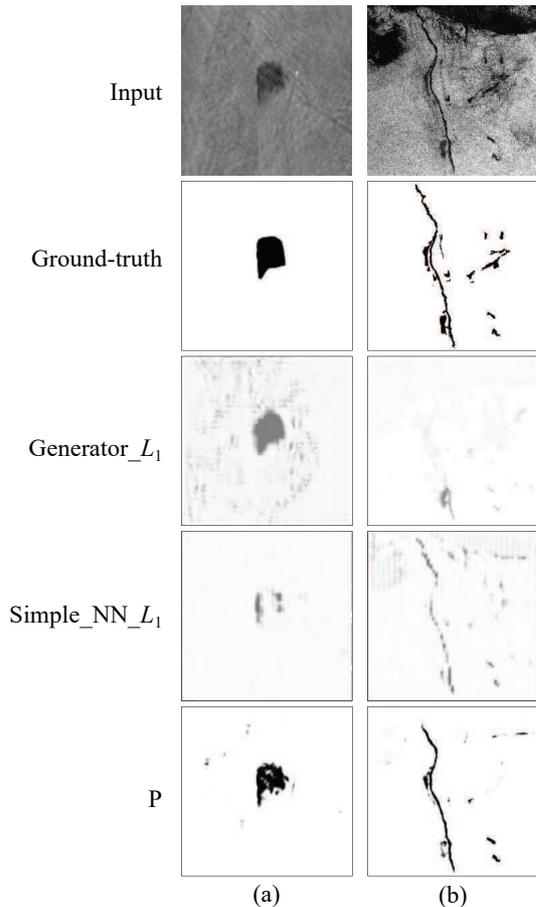


Fig. 10. Comparison between two different types of neural networks and our trained model (with P divergence).

neural network consisting of one convolutional layer and one deconvolutional layer. The segmentation results are shown in Fig. 10 as ‘Generator_ L_1 ’ and ‘Simple_NN_ L_1 ’, respectively. For a fair comparison, we train the both neural networks on same training data and parameter setting as the real SAR experiment. In Fig. 10, we only show the segmentation results of our trained model (with P divergence), since there is only a slight difference between our models with respect different divergences. It is observed that the classic method for oil spill segmentation using the both neural networks cannot generate a clear segmentation at all. However, our trained model (with P divergence) generates clear segmentations, even in the case of having non-oil dark spots on the SAR image (Fig. 10 (b)).

E. Comparison with Initialization Dependent Oil Spill Segmentation

In this part, we evaluate the performance of the proposed method by comparing its segmentation results with level set segmentation which is initialization dependent. We use the state-of-the-art region scalable fitness (RSF) level set segmentation method [14] for comparison. RSF requires a manual initialization by setting initial oil spill contours. It should be noted that different initialization usually leads to different segmentation performance.

TABLE IX
ACCURACY OF SEGMENTING REAL SAR IMAGES IN FIG. 11.

	Fig. 11 (a)	Fig. 11 (b)
RSF ₁	0.9202	0.9621
RSF ₂	0.9159	0.9620
P	0.9099	0.9670
SH	0.9134	0.9678
TV	0.9188	0.9683
CDD	0.9183	0.9679
SKL	0.9114	0.9658

TABLE X
RFE OF SEGMENTING REAL SAR IMAGES IN FIG. 11.

	Fig. 11 (a)	Fig. 11 (b)
RSF ₁	0.6960	0.9651
RSF ₂	0.7282	0.9816
P	0.0763	0.0900
SH	0.1062	0.1836
TV	0.1311	0.2302
CDD	0.1310	0.1726
SKL	0.0911	0.2195

We compare RSF and the proposed method on segmenting two SAR images with slender oil spills in Fig. 11, which are difficult for segmentation in terms of their irregular shapes. We conduct two different level set manual settings for each SAR image, which are marked by the dashed boxes on the input images. The segmentation results of RSF related to different initialization are shown below the ground-truth segmentation. Specifically, RSF₁ is on the left side and RSF₂ is on the right side in Fig. 11 (a) and (b), respectively.

The segmentation results of RSF₁ and RSF₂ are quite different to each other. Neither of them can accurately detect the oil spill regions, especially for the slender parts as shown in Fig. 11. On the other hand, the proposed method (especially with P and CDD divergence) is able to generate more accurate segmentation maps for the irregular and slender oil spills.

We also compare the quantitative results in terms of accuracy and RFE in Tables IX and X for RSF and proposed method, respectively. On the one hand, the proposed method achieves comparable segmentation accuracy as RSF method with different initialization. On the other hand, the RFE of the proposed model is much lower than that of RSF. For instance, the RFE of P divergence based segmentation model is almost ten times lower than that of RSF₁ and RSF₂, which reflects the effectiveness of our method. More importantly, the proposed method does not require manual initialization in the segmentation procedure. Therefore, it enables segmentation in an automatic and effective manner, which requires no human guidance and provides a more practically efficient route to real world oil spill segmentation.

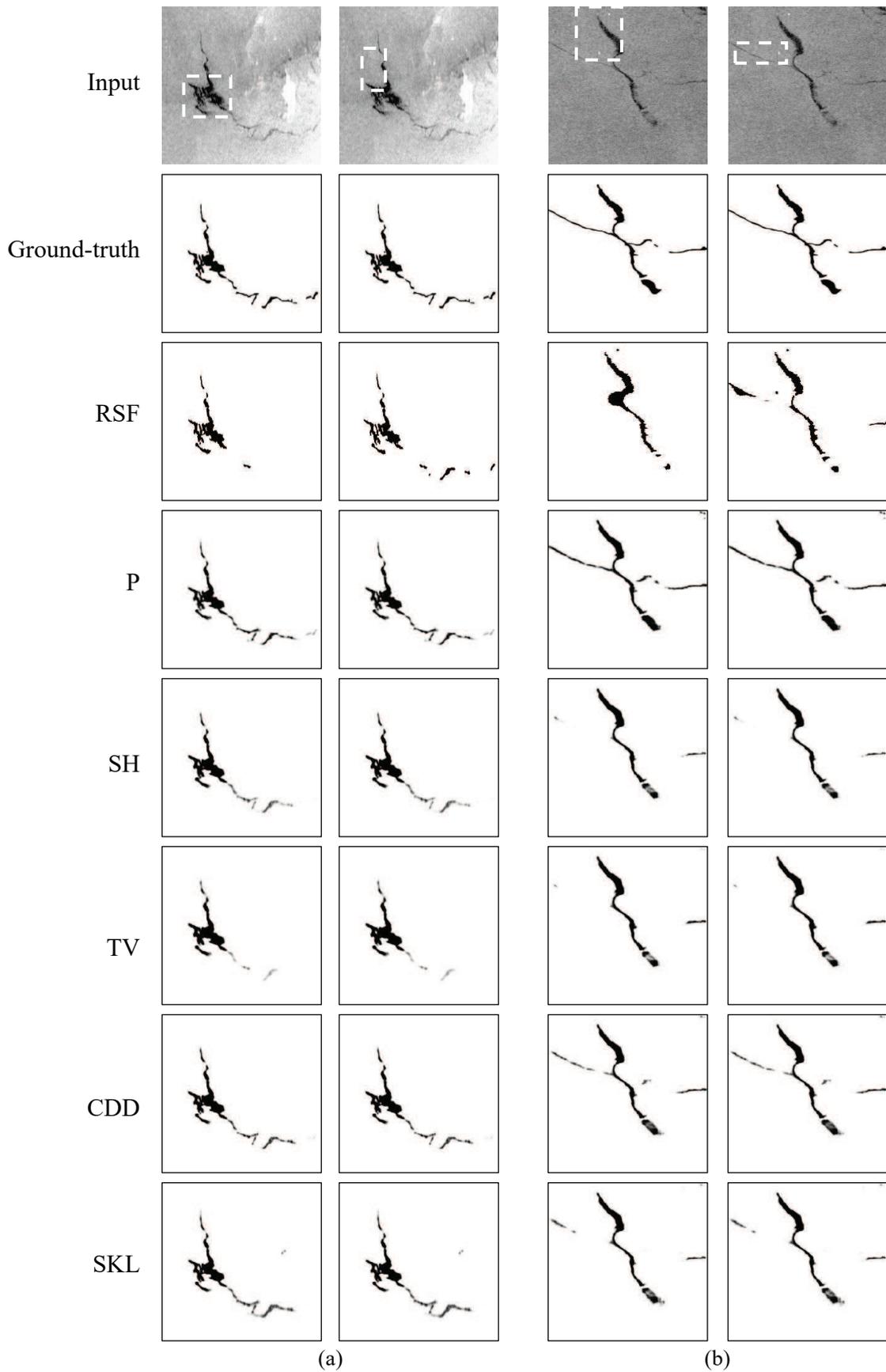


Fig. 11. Comparison between the initialization dependent method and the f -divergence minimization based method. The dashed boxes indicate manual initialization for the RSF level set method.

V. CONCLUSION AND FUTURE WORK

We have presented an automatic oil spill segmentation method for SAR image based on f -divergence minimization. Specifically, we aim at minimizing the f -divergence of a true segmentation distribution and a generated segmentation distribution for the purpose of accurately detecting oil spill regions in a SAR image. To practically achieve this goal, we have reformulated the f -divergence minimization in terms of seeking the minimal of the f -divergence tight lower bound and have adversarial trained two deep neural networks for variational approximation. There are several advantages of our method. First, it does not require a large number of training samples because the adversarial training procedure generates segmentation samples and enhance the representation power of the model. Second, our method performs accurate segmentation for irregular oil spills even in very noisy conditions because the comprehensiveness of f -divergence enables it to have the capability of tackling rigorous situations. Third, our segmentation method is totally automatic in detecting oil spills in an unknown image and does not require manual initialization.

There are several issues worth investigation in our future work. First, we will consider how to develop a model that automatically segment oil spill images with different sizes and resolutions. Second, we will incorporate the useful information such as multi-polarization features into our framework for lifting its potential in discriminating oil spills from lookalikes. In addition, we will keep investigating the effectiveness of our framework subject to more sophisticated scattering and contrast ratio models.

REFERENCES

- [1] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein generative adversarial networks," in *Proc. Int. Conf. Mach. Learning*, vol. 70, Aug. 2017, pp. 214–223.
- [2] F. Bandiera, A. Masciullo, and G. Ricci, "A bayesian approach to oil slicks edge detection based on sar data," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 5, pp. 2901–2909, Jul. 2014.
- [3] F. Bandiera and G. Ricci, "Slicks detection on the sea surface based upon polarimetric sar data," *IEEE Geosci. Remote Sens. Lett.*, vol. 2, no. 3, pp. 342–346, Jul. 2005.
- [4] A. Buono, F. Nunziata, M. Migliaccio, and X. Li, "Polarimetric analysis of compact-polarimetry sar architectures for sea oil slick observation," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 10, pp. 5862–5874, Jun. 2016.
- [5] D. Chandler, *Introduction to modern statistical mechanics*. New York: Oxford University Press, 1987.
- [6] M. J. Collins, M. Denbina, B. Minchew, C. E. Jones, and B. Holt, "On the use of simulated airborne compact polarimetric sar for characterizing oil water mixing of the deepwater horizon oil spill," *IEEE J. Sel. Topics Appl. Earth Observations and Remote Sens.*, vol. 8, no. 3, pp. 1062–1077, Mar. 2015.
- [7] F. Del Frate, A. Petrocchi, J. Lichtenegger, and G. Calabresi, "Neural networks for oil spill detection using ers-sar data," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 5, pp. 2282–2287, Sep. 2000.
- [8] M. M. Espeseth, S. Skrunes, C. E. Jones, C. Brekke, B. Holt, and A. P. Doulgeris, "Analysis of evolving oil spills in full-polarimetric and hybrid-polarity sar," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 7, pp. 4190–4210, Jul. 2017.
- [9] B. Fiscella, A. Giancaspro, F. Nirchio, P. Pavese, and P. Trivero, "Oil spill detection using marine sar images," *Int. J. Remote Sens.*, vol. 21, no. 18, pp. 3561–3566, Nov. 2010.
- [10] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. C. Courville, and Y. Bengio, "Generative adversarial networks," in *Proc. Neural Inform. Process. Syst.*, Jul. 2014, pp. 2672–2680.
- [11] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Comput. Vis. Pattern Recog.*, Jun. 2016, pp. 770–778.
- [12] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *Proc. Int. Conf. Mach. Learning*, vol. 37, Jul. 2015, pp. 448–456.
- [13] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *Int. Conf. Learning Representations*, May. 2015.
- [14] C. Li, C. Kao, J. C. Gore, and Z. Ding, "Minimization of region-scalable fitting energy for image segmentation," *IEEE Trans. Image Process.*, vol. 17, no. 10, pp. 1940–1949, 2008.
- [15] H. Li, J. Wu, W. Perrie, and Y. He, "Wind speed retrieval from hybrid-pol compact polarization synthetic aperture radar images," *IEEE J. Ocean. Eng.*, vol. PP, no. 99, pp. 1–12, Jul. 2017.
- [16] A. L. Maas, A. Y. Hannun, and A. Y. Ng, "Rectifier nonlinearities improve neural network acoustic models," in *Proc. ICML Workshop on Deep Learning for Audio, Speech and Language Processing*, 2013.
- [17] R. C. P. Marques, F. N. Medeiros, and J. S. Nobre, "Sar image segmentation based on level set approach and G_A^0 model," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 10, pp. 2046–2057, Dec. 2012.
- [18] L. W. Mdakane and W. Kleynhans, "An image-segmentation-based framework to detect oil slicks from moving vessels in the southern african oceans using sar imagery," *IEEE J. Sel. Topics Appl. Earth Observations and Remote Sens.*, vol. 10, no. 6, pp. 2810–2818, Jun. 2017.
- [19] M. Migliaccio, A. Gambardella, and M. Tranfaglia, "Sar polarimetry to observe oil spills," *IEEE Trans. Geos. Remote Sens.*, vol. 45, no. 2, pp. 506–511, Feb. 2007.
- [20] B. Minchew, C. E. Jones, and B. Holt, "Polarimetric analysis of backscatter from the deepwater horizon oil spill using l-band synthetic aperture radar," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 10, pp. 3812–3830, Mar. 2012.
- [21] M. Mirza and S. Osindero, "Conditional generative adversarial nets," *Neural Inform. Process. Syst. Deep Learning and Representation Learning Workshop*, pp. 2672–2680, Dec. 2014.
- [22] X. L. Nguyen, M. J. Wainwright, and M. I. Jordan, "On surrogate loss functions and f-divergences," *Ann. Stat.*, vol. 37, no. 2, pp. 876–904, 2009.
- [23] S. Nowozin, B. Cseke, and R. Tomioka, "f-gan: Training generative neural samplers using variational divergence minimization," in *Proc. Neural Inform. Process. Syst.*, 2016, pp. 271–279.
- [24] S. Pelizzari and J. Bioucas-Dias, "Oil spill segmentation of sar images via graph cuts," in *Proc. IEEE Int. Geos. Remote Sens. Symp.* IEEE, 2007, pp. 1318–1321.
- [25] M. Poodanchi, G. Akbarizadeh, E. Sobhanifar, and K. Ansari-Asl, "Sar image segmentation using morphological thresholding," in *Proc. Conf. Inform. Knowledge Technology (IKT)*, May 2014, pp. 33–36.
- [26] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," *Int. Conf. Learning Representations*, Dec. 2016.
- [27] P. Ren, M. Di, H. Song, C. Luo, and C. Grecos, "Dual smoothing for marine oil spill segmentation," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 1, pp. 82–86, Nov. 2016.
- [28] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Medical Image Computing and Computer-Assisted Intervention*, Oct. 2015, pp. 234–241.
- [29] A. H. S. Solberg, C. Brekke, and P. O. Husoy, "Oil spill detection in radarsat and envisat sar images," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 3, pp. 746–755, Feb. 2007.
- [30] T. Soukissian, F. Karathanasi, and P. Axaopoulos, "Satellite-based offshore wind resource assessment in the mediterranean sea," *IEEE J. Ocean. Eng.*, pp. 1–14, Jun. 2016.
- [31] D. Velotto, C. Bentes, B. Tings, and S. Lehner, "First comparison of sentinel-1 and terrasars-x data in the framework of maritime targets detection: South italy case," *IEEE J. Ocean. Eng.*, vol. 41, no. 4, pp. 993–1006, Apr. 2016.
- [32] D. Velotto, M. Migliaccio, F. Nunziata, and S. Lehner, "Dual-polarized terrasars-x data for oil-spill observation," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 12, pp. 4751–4762, Dec. 2011.
- [33] G. Xia, G. Liu, W. Yang, and L. Zhang, "Meaningful object segmentation from sar images via a multiscale nonlocal active contour model," *IEEE Trans. Geos. Remote Sens.*, vol. 54, no. 3, pp. 1860–1873, Oct. 2016.
- [34] M. Xu, Y. Yu, F. Chen, X. Jiang, P. Ren, and E. Yang, "Level sets with one dot fuzzy initialization for marine oil spill segmentation," in *Proc. IEEE OCEANS*, 2017.

- [35] M. D. Zeiler, G. W. Taylor, and R. Fergus, "Adaptive deconvolutional networks for mid and high level feature learning," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2011, pp. 2018–2025.



Xingrui Yu received his BEng in Electronic Information Engineering from China University of Petroleum (East China), Qingdao, China, in 2015. He is currently working toward the MEng in Information and Communication Engineering at China University of Petroleum (East China), Qingdao, China.

His current research interests include machine learning and deep learning (particularly adversarial learning), with applications to remote sensing.



He Zhang received his BEng in Electronic Information Engineering from the China University of Petroleum (East China), Qingdao, China, in 2016. He is currently working toward the PhD degree in Computer Science at University of Exeter, UK.

His research interests are in spatial and temporal information prediction using deep learning methods.



Chunbo Luo (M'11) is a Lecturer in Computer Science at the University of Exeter, UK. His research interests focus on developing model based and machine learning algorithms to solve practical problems such as wireless networking and remote sensing. He received his PhD degree from the University of Reading, UK, for his work on the study of high performance cooperative wireless networks in 2011. He then joined the EPSRC SUAAVE project within the University of Ulster. He later took a Lecturer position at the University of the West of Scotland

in 2013 and then moved to University of Exeter in 2015. He is a Fellow of the Higher Education Academy, an IEEE member and a BCS member. He is an invited reviewer of a group of IEEE Transactions, Elsevier journals and International Journal series.



Hairong Qi (F'17) received the BS and MS degrees in computer science from Northern JiaoTong University, Beijing, China, in 1992 and 1995, respectively, and the PhD degree in computer engineering from North Carolina State University, Raleigh, NC, in 1999. She is currently a Professor in the Department of Electrical Engineering and Computer Science at the University of Tennessee, Knoxville, TN. Her current research interests are in advanced imaging and robust collaborative processing in resource-constraint distributed environment and information

unmixing. She is the recipient of the NSF CAREER Award. She also received the Best Paper Awards at the 18th International Conference on Pattern Recognition and the third ACM/IEEE International Conference on Distributed Smart Cameras. She recently receives the Highest Impact Paper from the IEEE Geoscience and Remote Sensing Society. She is Fellow of the IEEE.



Peng Ren (SM'16) received his BEng and MEng both in Electronic Engineering from Harbin Institute of Technology, China, and his PhD in Computer Science from the University of York, UK. He is currently a Professor with College of Information and Control Engineering, China University of Petroleum (East China). He received K. M. Scott Prize from the University of York in 2011, and was one coauthor of the Eduardo Caianiello Best Student Paper Award at 18th International Conference on Image Analysis and Processing (ICIAP) in 2015. His major interests include remote sensing and machine learning. He has published over academic 40 papers in these research fields. He is a senior member of IEEE.