Training CNNs on speckled optical dataset for edge detection in SAR images

Chenguang Liu, Florence Tupin, Yann Gousseau

Training CNNs on speckled optical dataset for edge detection in SAR images

Chenguang Liu, Florence Tupin and Yann Gousseau

***Abstract*—Edge detection in SAR images is a difficult task due to the strong multiplicative noise inherent to this imaging modality. Many researches have been dedicated to edge detection in SAR images but very few try to address the most challenging situations, namely edge detection in 1-look real SAR images. Motivated by the success of Convolutional Neural Networks (CNNs) for the analysis of natural images, we study the applicability of a classical CNN edge detector to SAR images, especially for edge detection in 1-look real SAR images. One crucial factor that contributes to the success of CNNs is a training dataset with labeled ground truth. In order to avoid the tedious job of annotating a large amount of SAR images, we simulate a SAR dataset leveraging the optical dataset BSDS500 [1] to train CNN models. Under the hypothesis that both optical and SAR images can be divided into piecewise constant areas, the main gap between simulated SAR images and real SAR images is that the possibility of mean intensity values for homogeneous areas is different. Therefore, we propose to train the CNN model on the gradient magnitude fields of the SAR images. The motivation behind this choice is that, provided a suitable definition of the gradient for SAR images, the gradient magnitude fields of homogeneous areas follow the same distribution regardless of their mean intensity values and the gradient distribution of two homogeneous areas across boundaries depends only on the ratio of their mean intensity values. The proposed stategy yields a detector with an approximately constant false alarm rate (CFAR property). We propose to train a state-of-the-art edge detector based on deep convolutional networks, namely HED (Holistically-Nested Edge Detector) [2], [3], on a Ratio-based Gradient operator (GR) on SAR images. The proposed edge detector, GRHED, achieves excellent performances in all our simulations including several 1-look synthetic edge images with different edge contrasts, two hundred 1-look optical images with synthetic noise, which are simulated from BSDS500, one synthetic SAR image and two 1-look real SAR images. In all these situations, the proposed edge detector outperforms concurrent approaches.**

***Index Terms*—edge detection, 1-look SAR image, speckled optical dataset, CNNs, gradient distribution, GRHED**

I. Introduction

Edges are important features in Synthetic Aperture Radar (SAR) images. They can be used as low level features for many applications like line segment detection [4], [5], SAR image segmentation [6], [7], coastline detection [8], [9], image registration [10], [11] and even SAR image despeckling [12]. Due to the strong multiplicative speckle noise in SAR images, methods developed for optical images, which are usually based on pixel value differences, produce more false edges in brighter areas and thus are not suitable for SAR images. Many researches have been dedicated to edge detection in SAR images in the past years. In [13], the Ratio of Average (ROA) was proved to have a constant false alarm rate for SAR images (CFAR property). The ratio operator is applied along four directions and the minimum normalized ratio is used to compute the gradient magnitude. The direction corresponding to the minimum normalized ratio is regarded as the edge orientation. A threshold determined by a given probability of false alarm rate and a morphological operator are then applied to obtain a binary thin edge map. However, ROA is optimal only for isolated step edges. An efficient multiedge detector, ROEWA, was proposed afterwards in [14]. The Ratio of Exponentially Weighted Average was shown to be optimal in terms of minimum mean square error (MMSE) under the hypothesis of a stochastic multi-edge model. The method is based on ROA, but averages are weighted by a decreasing exponential function, allowing a better detection of multiple edges close to each other. Besides, instead of computing the ratio along four directions, the ratio is computed along the horizontal and vertical directions and the normalized ratios are considered as the horizontal and vertical components of the gradient magnitude. A modified watershed algorithm is then used to threshold the Edge Strength Map and a region merging algorithm is used to eliminate the false edge pixels. Edge detectors using different shape of window functions were introduced later in [15] and [16]. Non-maxima suppression [17] and hysteresis thresholding are applied to obtain the binary thin edge map. An edge compensation strategy was also introduced in [16] to extract weak edge pixels. To reduce the influence of isolated strong bright points in real SAR images, an Anisotropic Morphological Directional Ratio (AMDR) [18] was proposed by replacing the weighted average filter with the weighted median filter. The edge localization accuracy in the Edge Strength Map (ESM) and the Edge Direction Map (EDM) was then improved by a multiplicative spatial and directional matching filter. By introducing the improved ESM and EDM into the routine of Canny edge detector [17], the resulting edge detector is able to obtain a binary thin edge map. The connectivity of edges is finally improved by an edge remedy strategy. However, the performances of the edge detectors developed for SAR images are still not fully satisfying, especially in the very noisy 1-look situation.

On the other hand, convolutional neural networks (CNNs) have proven to be very succesful for edge detection in natural images, with approaches such as Deepedge [19], DeepContour [20], HED [2], [3], CEDN [21], AMH-Net [22] or

Chenguang Liu, Florence Tupin and Yann Gousseau are with Télécom Paris, Institut Polytechnique de Paris, France (email: chenguangl@whu.edu.cn; florence.tupin@telecom-paris.fr; yann.gousseau@telecom-paris.fr). This work is supported by the China Scholarship Council (File No. 201606270202), the National Natural Science Foundation of China (NSFC) (Grant No.61771014) and Télécom Paris, Institut Polytechnique de Paris.

RCF [23], [24], which have permitted to improve significantly traditional edge detectors like Sobel [25], Canny [17], Statistical Edge [26], Pb [27], gPb [1] and Structured Edge [28], at least on databases similar to the training sets. Motivated by this fact, we study the possibility to apply CNN-based methods to SAR images. We will pay a special attention on developing methods that are as CFAR as possible.

One crucial factor that contributes to the success of CNNs is the availability of training datasets with ground truth. A first difficulty for the present work is that, to the best of our knowledge, there is still no available training dataset for edge detection in SAR images. Considering that edges mostly correspond to changes in local brightness and textures (and color for color images), we assume that the problem of edge detection should not rely much on the specific content of the image. In this case, datasets of natural images used to train CNN could be used to train a SAR edge detector, provided images are modified using a reasonable noise model. On this basis, we propose to simulate a SAR-like training set by multiplying by speckle noise datasets of natural images. The natural images multiplied by speckle are called speckled optical images in the following.

Another difficulty in applying CNN-based method to SAR images is their specific dynamic ranges. The range of values is indeed much higher for SAR images than for natural images. To cope with this problem and ease the training of the network, we propose in this paper to apply a pre-processing step to the SAR images by applying a first low-level edge detection step. This step is done using the Gradient by Ratio (GR) proposed in [29]. These pre-processed features are then given as input to the network to be trained (HED [2], [3]). This strategy will be called GRHED in the following.

This paper is organized as follows: in Section II we give details about the original HED method and explain the building of the speckled optical dataset. In section III, we describe our gradient-based strategy and the proposed method, GRHED. In section IV, we demonstrate the efficiency of GRHED with experiments on several 1-look simulated edge images, 200 speckled optical images simulated from BSDS500 [1], one 1-look realistically simulated SAR image, and two 1-look real SAR images.

II. Original HED algorithm and training dataset

A. Description of HED

The HED method, introduced in [2], [3], relies on a fully convolutional network [30], [31], which is trained end-to-end to perform image-to-image prediction. The network architecture of HED is shown in table I. The HED network is trimmed from the VGG-16 [32] net by discarding the last max-pooling layer and the 3 fully connected layers. Motivated by the deeply-supervised nets [33], five side outputs are added to the convolutional layers just before the five max-pooling layers in the original VGG-16 net. The loss function of the HED network is composed of the loss function from the side layers and the loss function from the fused output. The final output of HED is an average of the side outputs and the fused output. In-network bilinear interpolation [30], [31] is used to upsample the side outputs so that they all have the same size as the edge ground truth.

TABLE I: The network architecture of HED

conv3-64 conv3-64	→ side output 1
MAX-POOLING	
conv3-128 conv3-128	→ side output 2
MAX-POOLING	
conv3-256 conv3-256 conv3-256	→ side output 3
MAX-POOLING	
conv3-512 conv3-512 conv3-512	→ side output 4
MAX-POOLING	
conv3-512 conv3-512 conv3-512	→ side output 5

B. Speckled optical dataset

Since the aim of this paper is to train CNNs for edge detection in SAR images, the training dataset is of crucial importance for the performances of the edge detector. In order to avoid the tedious job of manual edge labeling, we leverage the available optical dataset for edge detection in natural images.

In order to simulate a SAR dataset, it is very important to take into account the statistics of SAR images. SAR data represents a complex backscattered electro-magnetic field z whose modulus $|z|$ is the amplitude and the square of the modulus $|z|^2$ the intensity data. Due to the coherent imaging system, SAR data present the well known speckle phenomenon. Following Goodman model [34] of fully developed speckle, it can be shown that the amplitude of the backscattered electromagnetic field of a homogeneous area with mean intensity $\langle I \rangle$ follows a Nakagami distribution :

$$f(t|\langle I \rangle) = \frac{2}{\Gamma(L)} \left(\frac{L}{\langle I \rangle} \right)^L t^{2L-1} e^{-(Lt^2/\langle I \rangle)}, \tag{1}$$

L being the number of looks of the image. For images with the best resolution, $L = 1$ and the amplitude of a physically homogeneous area follows a Rayleigh distribution. Another way of modeling 1-look data is the multiplicative model : $t = \sqrt{\langle I \rangle} s$, s representing the speckle noise and following the Rayleigh distribution given in eq. (1) with $\langle I \rangle = 1$ and $L = 1$.

Using the multiplicative noise model it is therefore easy to generate speckled data by multiplying an image by s. Although not truly verified (for real data the speckle is spatially correlated and the fully developed model of Goodman is verified only for rough surfaces [34]), this model is widely used to generate simulated data.

In this paper, we use the Berkeley Segmentation Data Set 500 (BSDS500) [1] with the same data augmentation as in

HED [3] for training and validation, to simulate a SAR dataset. Specifically, each image in the training and validation set is rotated by 16 angles, flipped horizontally, and rescaled to 50%, 100% and 150% of its original size. There is no data augmentation for the testing set. The resulting speckled optical dataset, which we will call BSDS500-speckled in the following, is obtained by multiplying the grayscale intensity channel of each color image with 1-look speckle noise. It contains $300 \times 16 \times 2 \times 3 = 28800$ images for training (80%) and validation (20%) and 200 images for testing. After creating a simulated dataset for training, the next question is how to train HED using this dataset and how to enable the trained model to be applied real SAR images. We will describe the way to tackle the gap between the training and testing data in the next section.

III. GRHED, AN EDGE DETECTOR ADAPTED TO SAR IMAGES

Even though speckled optical images and real SAR images are both contaminated by speckle noise, they still largely differ. Specifically, the range of pixel values differs a lot between natural images coded on 8-bits and thus having a maximum value of 255 and real SAR data. Real SAR data are known to have heavy-tailed distributions with very strong values backscattered by trihedral configurations. Besides the range of the data is between 0 and several thousands (for instance between 0 and 37 000 for Sentinel-1 amplitude images).

In order to deal with the differences in the range of pixel values between speckled optical images and real SAR images, but also to ease the training of the network, we propose to train HED on the gradient magnitude fields of the training images and to apply the trained network to the gradient magnitude fields of images at test time. The main reason for computing the gradient is that the gradient distribution of speckled optical images and SAR images are similar for the usual ratio based gradient computation methods [13], [14], [29]. The interest of this strategy will be discussed in the following and be demonstrated in the experimental part, Section IV.

The gradient computation is achieved by the Gradient by Ratio method [29] and the HED network is trained on the gradient magnitude fields computed by GR. We therefore call GRHED the resulting architecture. In the following we will describe the details about the GR and the benefits of computing GR.

A. Gradient by Ratio (GR)

For a given pixel located at position (x, y) in the image u, the horizontal and vertical gradient components (GR) are defined as [29]:

$$G^h(x,y) = \log(R^h(x,y)),$$

$$G^v(x,y) = \log(R^v(x,y)).$$

where $R^h(x,y)$ and $R^v(x,y)$ are the ratio of exponentially weighted average in the opposite side windows of pixel located at (x, y), along the horizontal and vertical directions respectively. In the horizontal direction, $R^h(x,y)$ can be computed as

$$R^h(x,y) = \frac{m_1^h(x,y)}{m_2^h(x,y)},$$

where

$$\begin{aligned} m_1^h(x,y) &= \sum_{x'=-W}^{W} \sum_{y'=1}^{W} u(x+x', y+y') \times e^{-\frac{|x'|+|y'|}{\alpha}}, \\ m_2^h(x,y) &= \sum_{x'=-W}^{W} \sum_{y'=-W}^{-1} u(x+x', y+y') \times e^{-\frac{|x'|+|y'|}{\alpha}}, \end{aligned}$$

and where W is the upper integer part of $\log(10) \times \alpha$. The quantity $R^v(x,y)$ can be computed in a similar way.

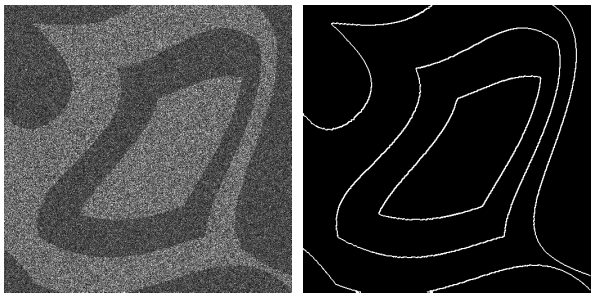
The magnitude $|GR(x,y)|$ and orientation $ang(GR(x,y))$ of GR at position (x, y) are defined as

$$\begin{aligned} |GR(x,y)| &= \sqrt{G^h(x,y)^2 + G^v(x,y)^2}, \\ ang(GR(x,y)) &= \arctan \frac{G^v(x,y)}{G^h(x,y)}. \end{aligned}$$

B. Interest of using GR to feed the network

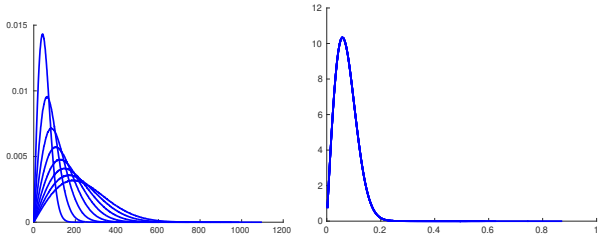
An image, whatever optical or SAR, can be roughly divided into two types of regions: homogeneous areas and boundaries (boundaries exist between two homogeneous areas). If we train HED directly on the images, the aim of training can be summarized as follows: 1) HED should not produce edge pixels in homogeneous areas; 2) HED should be able to highlight the boundaries between any two homogeneous areas. In this case, HED is trained to process samples drawn from many different distributions and the way of HED to process those samples depends on their corresponding distributions. In 1-look SAR images, the amplitude of all homogeneous areas follows a Rayleigh distribution depending on their mean intensity values. Therefore, even for homogeneous areas, there are many different distributions that HED has to learn to process. For two homogeneous areas across boundaries, their distribution depends on the mean intensity values of both homogeneous areas. Computing the ratio based gradient thus ease the training process, reducing the number of situations to learn, since for all homogeneous areas, their gradient magnitude fields follow the same distribution, regardless of their mean intensity values, and the gradient distribution computed over two homogeneous areas across boundaries depends only on the ratio of the mean intensity values. What's more important, The main gap between speckled optical images and real SAR images is that SAR images may contain homogeneous areas with very high mean intensity values. Since the gradient distribution will not be influenced by the mean intensity values of any homogeneous areas, feeding the gradient features enables the network trained on speckled optical dataset to work well in real SAR images. Besides, it ensures a CFAR property for the network result since only CFAR features are given as input.

In order to give a better explanation on the benefits of computing GR, we compare the distribution of the data before



(a) 1-look image (b) GT

Fig. 1: A 1-look synthetic image with ratio contrast 1.5 and the associated ground truth. The size of the image is 512×512 pixels.



(a) histograms of images (b) histograms of GR fields

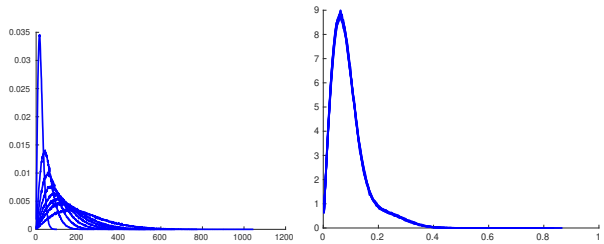
Fig. 2: Histograms of the eight 1-look pure noise images and histograms of their gradient magnitude fields computed by GR with $\alpha = 4$.

and after computing the gradient for both homogeneous areas and two homogeneous areas across boundaries. We simulate eight 1-look pure noise images of size 4096×4096 pixels and eight 1-look synthetic images of size 512×512 pixels with amplitude ratio contrast 1.5. The square root of the mean intensity values of the homogeneous images (proportional to an amplitude value) are 60, 90, 120, 150, 180, 210, 240 and 270. For two homogeneous areas across boundaries, the amplitude ratio is chosen as 1.5 for mean values (for the smallest value along the edge) of $20, 50, 70, 90, 110, 130, 150, 200$. One example of the synthetic edge image can be found in Fig. 1-(a).

The histograms of the eight 1-look pure noise images and the histograms of their gradient magnitude fields computed by GR with weighting parameter $\alpha = 4$ can be found in Fig. 2. The histograms of eight 1-look synthetic edge images as well as the histograms of their gradient magnitude fields can be found in Fig. 3. From Fig. 2 and Fig. 3 we can deduce that the number of distributions that has to be learned by HED when trained on the gradient magnitude fields is much smaller than that of a direct training on the images. Since the amount of data is unchanged, the amount of training data for each distribution is therefore increased.

C. The resulting GRHED edge detection method

Since the gradient distribution of GR is the same for homogeneous areas with all possible mean intensity values, and the gradient distribution in two homogeneous areas across



(a) histograms of images (b) histograms of GR fields

Fig. 3: Histograms of the eight 1-look synthetic edge images and histograms of their gradient magnitude fields computed by GR with $\alpha = 4$.

boundaries is influenced by the ratio and not by their mean intensity values, we assume that the possible distributions in the gradient feature space of real SAR images are included in those of the training dataset. Therefore, we propose to train HED on the gradient magnitude field of GR. This can be seen as the addition of a hand-crafted layer before the usual HED layers. The resulting GRHED has the following advantages:

- the distribution of feature map values obtained by GR depends only on the ratio of the mean intensity values on the opposite side windows of each pixel (ratio being 1 for pixels located in homogeneous areas), therefore, homogeneous areas with very high mean intensity values in real SAR images will not impair the performances of GRHED;
- constant false alarm rate (CFAR) is ensured for SAR images because of the ratio operation;
- by using multiple weighting parameter α values in GR [29] we can combine diverse informations from the image by concatenating together gradient magnitude fields produced by GR with different α values. It has been studied in [35] that GR with different α values can capture complementary informations.

IV. Experiments

In this section we study the performances of the method we propose, GRHED, using several 1-look synthetic edge images, two hundred 1-look speckled optical images in BSDS500-speckled, one 1-look realistically simulated SAR image and two 1-look real SAR images. In order to show the efficiency of GRHED, we compare it with the original HED algorithm and to the result of training HED on the logarithm of images, as defined below, an algorithm that we will cal HED-log. Observe that for GRHED, HED and HED-log, the convolutional layers that need to be trained are the same, the difference between these method being the input of the network: gradient magnitude fields of images for GRHED, plain images for HED and logarithm of these for HED-log. To take into account the distribution of the training set, the SAR data are first normalized to have a similar distribution as explained in section IV-A. The training strategy for those convolutional layers is as follows: we use Adam optimizer to train the nework from scratch on the speckled optical dataset BSDS500-speckled. The number of iterations for training is

10000 and the learning rate is 0.001. The size of the batch is 10. The preprocessing step before feeding the data into the learnable layers is global mean substraction, as done in VGG. Since the outputs of all methods are probability edge maps, they should be processed further to obtain the binary edge maps. In order to obtain the binary edge map, we use the same Non-maxima Suppression procedure as the one in Structured Edge [28] and use a threshold to discard pixels with low values in the probability edge map.

The following methods will be used in the comparison of the next sections:

- GR with $\alpha = 4$ (this choice of $\alpha = 4$ being adapted to GR in 1-look situations, as discussed in [35]). For the GR magnitude field, we use the same postprocessing steps as for the magnitude field produced by HED, HED-log and GRHED;
- HED: HED is trained on the original speckled images and tested on normalized images (see IV-A);
- HED-log: HED is trained on the logarithm of the images in BSDS500-speckled, and is applied to the logarithm of testing images after normalization (see IV-A);
- GRHED with multiple α values, $\alpha = 2, 3, 4, 5$: HED is trained on the gradient feature maps which are obtained by concatenating the gradient magnitude fields produced by GR with different α values.

A. Normalization of the SAR data

As already mentioned, speckled optical data have a much narrower dynamic range than real SAR images. This is a problem at test time if we want the method to be efficient on real SAR images. We therefore need to define a strategy to normalize images at test time, ensuring a relatively stable dynamic range. In order to do so, for a given image (potentially a real SAR image), we divide its amplitude values by the square root of their mean intensity value, before multiplying them by the square root of the mean intensity value of the optical (training) dataset. Let $\langle I_{opt} \rangle$ be the mean intensity of the global training set and $\langle I_{SAR} \rangle$ the mean intensity value of the SAR data, the normalization formula of the SAR amplitude values u is the following :

$$\hat{u} = \frac{u.\sqrt{\langle I_{opt} \rangle}}{\sqrt{\langle I_{SAR} \rangle}} \tag{2}$$

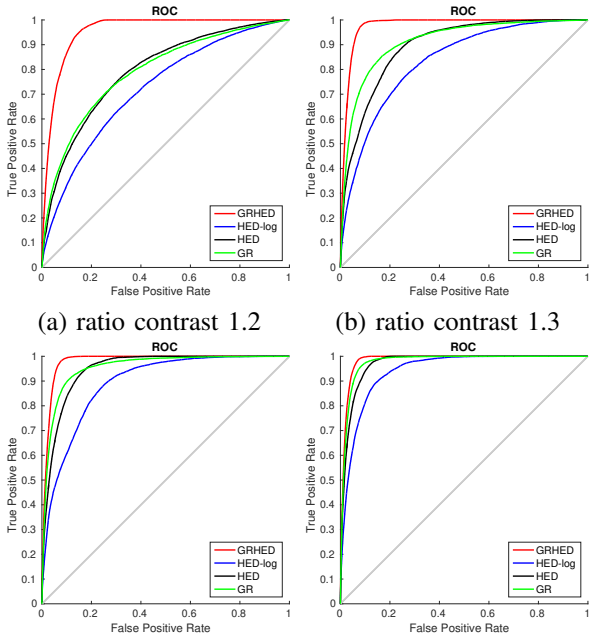
It is easy to check that the the mean value of $\hat{u}$ in intensity is equal to $\langle I_{opt} \rangle$ and thus corresponds to the one of the training distribution.

We use maximum likelihood estimator to estimate $\langle I_{opt} \rangle$, while for $\langle I_{SAR}$, to avoid being too much influenced by the strong backscattered values in real SAR images (especially for urban areas), we propose to use a robust estimate of $\langle I_{SAR} \rangle$, assuming a global Rayleigh distribution of the scene. Under this hypothesis it can easily be shown that the link between the mode of the distribution A_m and the mean intensity is the following:

$$\langle I_{SAR} \rangle = 2A_m^2.$$

In the following, for each test image, A_m is computed to estimate $\langle I_{SAR} \rangle$ and the data are normalized using equation (2).

B. Comparison in 1-look synthetic edge images



(a) ratio contrast 1.2 (b) ratio contrast 1.3

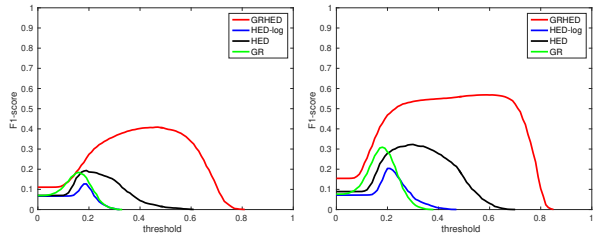
(c) ratio contrast 1.4 (d) ratio contrast 1.5

Fig. 4: ROC curves computed from the magnitude field produced by GR, HED, HED-log and GRHED in 1-look simulated images with contrast 1.2, 1.3, 1.4 and 1.5. The size of the images is 512×512 pixels.

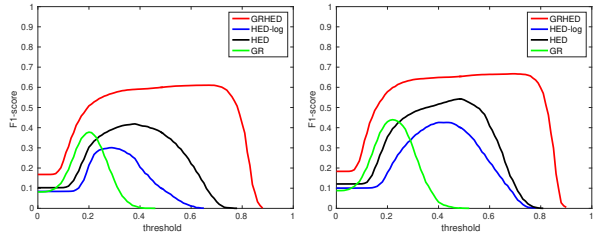
In order to give a fair comparison between the different methods, we compare them in terms of ROC curves and F1-score curves in 1-look synthetic edge images with different ratio contrasts.

One example of the simulated 1-look edge images with contrast 1.5 and the corresponding ground truth can be found in Fig. 1. The ROC curves computed in the magnitude fields produced by GR, HED, HED-log and GRHED in 1-look simulated images with ratio contrast 1.2, 1.3, 1.4 and 1.5 can be found in Fig. 4. It can be seen from Fig. 4 that GRHED yields the best performances in all situations, especially in 1-look and low contrast situations. The performances of GR and HED appear comparable in terms of ROC curves, and both are better than HED-log.

In order to give a clearer comparison between the different methods, we apply the same Non-maxima suppression step for all methods and vary the threshold from 0.0 to 0.9 with step 0.01 for each method. The F1-score is computed for each threshold and the corresponding F1-score curves for each method in the 4 simulated edge images can be found in Fig. 5. From Fig. 5 we can see that the F1-score curves of GRHED are above the F1-score curves of the other methods. In addition, flat areas of F1-score curves of GRHED indicate greater stability regarding the threshold choice than with other approaches. Although the best F1-scores that can be obtained



(a) ratio contrast 1.2 (b) ratio contrast 1.3



(c) ratio contrast 1.4 (d) ratio contrast 1.5

Fig. 5: F1-score curves obtained by GR, HED, HED-log and GRHED in synthetic edge images with ratio contrast 1.2, 1.3, 1.4 and 1.5. The F1-score curves are obtained by varying the threshold from 0.0 to 0.9 with step 0.01.

by GR and HED are comparable, the performances of HED are less sensitive to the choice of the threshold.

The optimal edge maps obtained by GR, HED, HED-log and GRHED in those synthetic images with contrast 1.2 and 1.4 (using the threshold which gives the best F1-score) can be found in Fig. 6 and Fig. 7. From Fig. 6 and Fig. 7 we can see that the ability of GRHED to preserve true edge pixels and suppress noise pixels is clearly greater on these images than when using the other methods; the GRHED method detects most true edge pixels while having the least number of false detections. It should be noted that although these edge maps are obtained using the optimal threshold, the edge maps obtained by HED and GRHED will remain reasonable with the threshold in a relatively large range, as can be deduced from Fig. 5, especially for GRHED.

C. Comparison of different algorithms on the speckled optical images in BSDS500-speckled

In order to give a more comprehensive comparison in more general situations between different methods, we compare GR, HED, HED-log and GRHED in the two hundred 1-look speckled optical images in BSDS500-speckled. For GR, we use $\alpha = 2$ and $\alpha = 4$. For GRHED, GRHED with a single α value and GRHED combining multiple α values are all provided. Three criterions are used to compare different algorithms: ODS F1 (fixed contour threshold for 200 images), OIS F1 (best threshold for each image), and average precision (AP). The quantitative comparison can be found in table II. From table II we can see that compared to GR, the CNN-based methods give much higher values for all three criterions. In addition, GRHED gives at least comparable or even better performances than HED and HED-log in the 200 1-look images, especially when combining multiple α values. We also notice that GRHED combining multiple α values yields better performances than GRHED using a single α value, which is probably due to the richer information which is provided to HED convolutional layers.

The edge maps obtained by GR (with $\alpha = 2$), GR (with $\alpha = 4$), HED, HED-log, GRHED (with $\alpha = 2$) and GRHED (combining $\alpha = 2, 3, 4, 5$) on one speckled optical image can be found in Fig. 8. For each method, the threshold is chosen to be the one corresponding to the ODS F1, which gives the best results in the 200 images. Specifically, the threshold used for each method can be found in table III. It should be noted that using the threshold corresponding to the ODS F1, GRHED (combining $\alpha = 2, 3, 4, 5$) will obtain near-optimal edge maps in the simulated edge images as shown in Fig. 5. The ground truth of this image is the one provided by 1 labeler (there are usually five labelers for each image in BSDS500). From Fig. 8 we can see that GRHED detects slightly more true edge pixels than HED, and that both preserve more true edge pixels while detecting less false detections than GR. Using HED as a postprocessing of GR field, the GRHED is able to strengthen true edge pixels and suppress false detections.

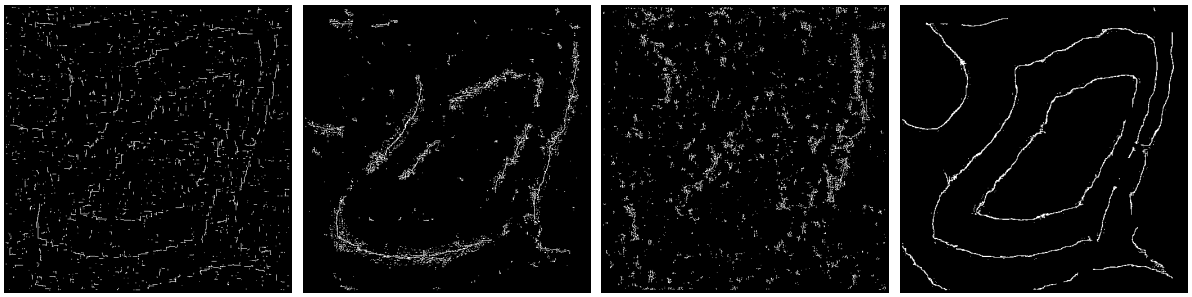
TABLE II: The performances of different methods over the 200 speckled optical images in BSDS500-speckled.

methods	ODS (F1)	OIS (F1)	AP
GR ($\alpha = 2$)	0.5658	0.5852	0.5094
GR ($\alpha = 4$)	0.5894	0.6151	0.5286
HED	0.6461	0.6671	0.6981
HED-log	0.6258	0.6466	0.6838
GRHED ($\alpha = 1$)	0.6427	0.6523	0.6949
GRHED ($\alpha = 2$)	0.6603	0.6762	0.7208
GRHED ($\alpha = 3$)	0.6570	0.6783	0.7018
GRHED ($\alpha = 4$)	0.6552	0.6729	0.7050
GRHED ($\alpha = 5$)	0.6492	0.6687	0.6897
GRHED ($\alpha = 6$)	0.6463	0.6660	0.6893
GRHED ($\alpha = 2, 3, 4, 5$)	0.6643	0.6826	0.7109
GRHED ($\alpha = 1, 2, 3, 4, 5, 6$)	0.6643	0.6832	0.7070

D. Comparison of different methods in 1-look SAR images

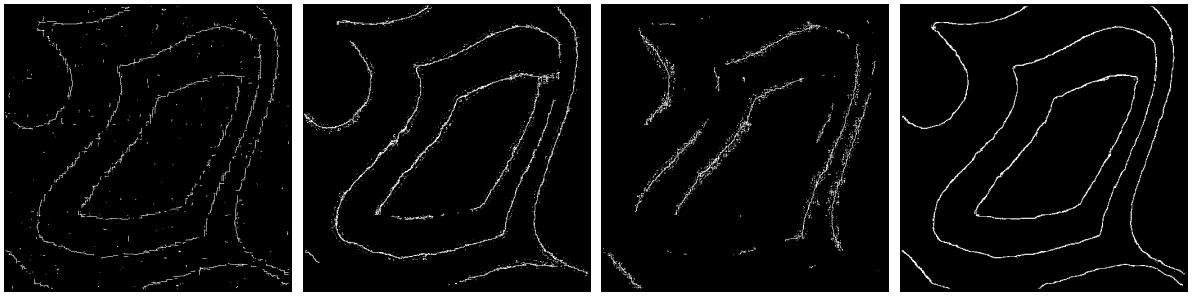
Though the efficiency of GRHED has been demonstrated in both simulated edge images and speckled optical images, demonstrating its ability to detect edges in real SAR images is the most important point in practice.

1) Setting thresholds according to a given probability of false alarm rate: Usually, detection thresholds for SAR images are set according to a chosen probability of false alarms (pfa). This is only possible when the detectors have a constant false alarm rate (CFAR). Therefore, it is important to study whether HED, HED-log and GRHED have CFAR for SAR images. The CFAR property can be checked experimentally by ploting the histograms of their gradient magnitude fields computed in noise images having different mean values. A method is considered to have CFAR if the histograms of its gradient magnitude fields computed in all images overlap well. We plot the histograms of the magnitude fields produced by



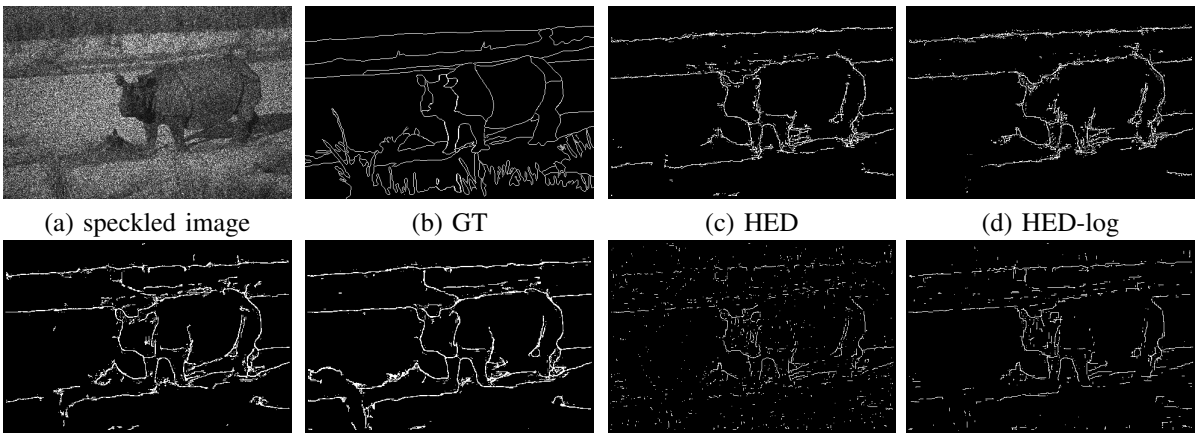
GR, contrast 1.2 HED, contrast 1.2 HED-log, contrast 1.2 GRHED, contrast 1.2

Fig. 6: Optimal edge maps obtained by GR, HED, HED-log and GRHED in 1-look simulated edge images with contrast 1.2. For each method, the chosen threshold gives the best F1-score in this image. The size of the images is 512×512 pixels.



GR, contrast 1.4 HED, contrast 1.4 HED-log, contrast 1.4 GRHED, contrast 1.4

Fig. 7: Optimal edge maps obtained by GR, HED, HED-log and GRHED in 1-look simulated edge images with contrast 1.4. For each method, the chosen threshold gives the best F1-score in this image. The size of the images is 512×512 pixels.



(a) speckled image (b) GT (c) HED (d) HED-log

(e) GRHED ($\alpha = 2$) (f) GRHED ($\alpha = 2, 3, 4, 5$) (g) GR ($\alpha = 2$) (h) GR ($\alpha = 4$)

Fig. 8: Comparison of different edge detectors on 1-look speckled optical images.

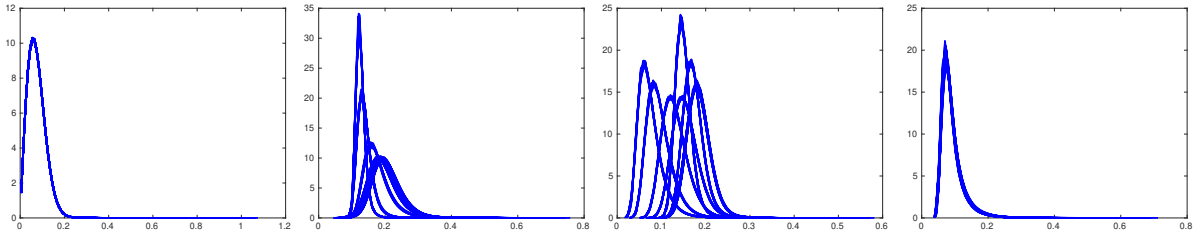
GR, HED, HED-log and GRHED in 160 1-look noise images of size 1024×1024 pixels in Fig. 9. The square root of the mean intensity values (proportional to the amplitude) of these noise images can take the following 8 values: 50, 80, 120, 150, 180, 200, 230 and 250. For each possible mean value, there are 20 random realizations. From Fig. 9 we can see that GR and GRHED both hold CFAR, but HED and HED-log do not.

Next, we set the detection thresholds according to a given pfa (envethough HED and HED-log do not have CFAR for SAR images). The threshold corresponding to a given probability of false alarm rate for GR, HED, HED-log and GRHED is estimated from 160 noise images as shown in table IV. Notice that for different methods, we will use the testing threshold corresponding to the pfa (10^{-5}, fifth column in table IV).

2) Comparison of different algorithms in synthetic realistic SAR images: It is usually difficult to annotate the edges in real SAR images due to very strong multiplicative noise. In order to give a quantitative evaluation on the performances of HED, HED-log, GRHED, and GR in images with targets similar to those in real SAR images, we use the ground truth which is obtained by applying HED-clean (HED trained on the grayscale images converted from clean natural images in

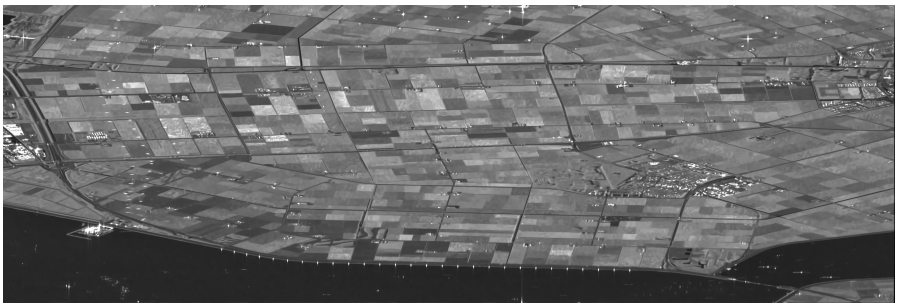
TABLE III: The threshold corresponding to the ODS F1-score for each method.

methods	HED	HED-log	GRHED ($\alpha = 2$)	GRHED ($\alpha = 2, 3, 4, 5$)	GR ($\alpha = 2$)	GR ($\alpha = 4$)
threshold (ODS F1)	0.5666	0.5306	0.5686	0.5516	0.3137	0.2745

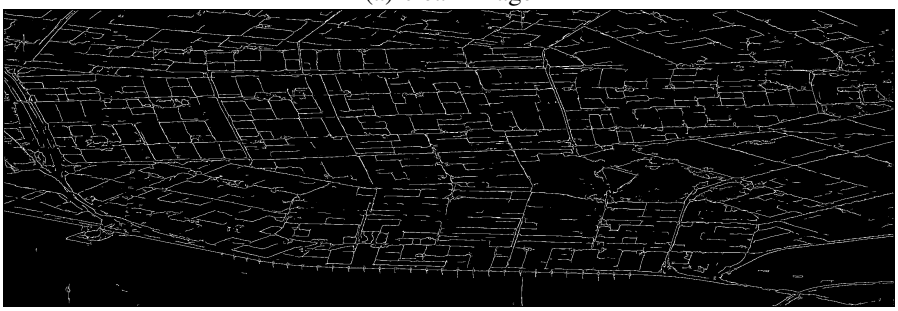


(a) GR ($\alpha = 4$) (b) HED (c) HED-log (d) GRHED ($\alpha = 2, 3, 4, 5$)

Fig. 9: The histograms of the gradient magnitude fields produced by GR, HED, HED-log and GRHED in 160 1-look pure noise images of size 1024×1024 pixels. There are 8 possibilities for the square root of the mean intensity values of these images, namely, 50, 80, 120, 150, 180, 200, 230 and 250. For noise images with each mean intensity value, there are 20 random realizations.



(a) clean image



(b) GT

Fig. 10: A denoised multi-look image (Leystad, Sentinel 1) and its ground truth. The size of the image is 1024×3072 pixels.

TABLE IV: The threshold corresponding to a given pfa for each method.

pfa	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}
threshold (GR)	0.18	0.22	0.26	0.29	0.31
threshold (HED)	0.3	0.38	0.48	0.57	0.63
threshold (HED-log)	0.23	0.27	0.3	0.34	0.39
threshold (GRHED)	0.17	0.26	0.37	0.45	0.52

BSDS500) to a SAR image with very little speckle noise. This SAR image with reduced speckle is obtained by averaging a large amount of well registered Sentinel-1 images (equivalent to temporal multi-looking). Although not justified in changing areas, this temporal multi-looking allows a strong reduction of the speckle in stable ones like roads, urban areas, etc. To overcome the varying residual noise, a final despeckling step is applied [36]. The image is then converted to an 8-bits image using a clipping between $[0, 255]$ with a threshold given by the mean value of the image plus three times its standard deviation. This image is very similar to a natural image and an "edge ground truth" is obtained by using HED-clean on it. The multi-temporal despeckled SAR image and its associated ground truth can be found in Fig. 10.

F1-score curves computed for GR, HED, HED-log and

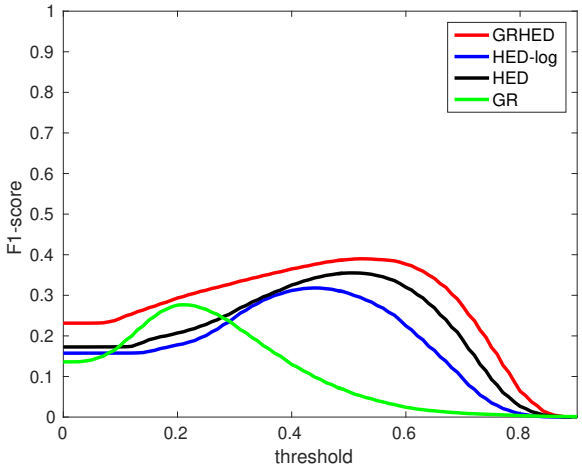


Fig. 11: F1-score curves computed for GR, HED, HED-log and GRHED on a 1-look synthetic SAR image of size 1024×3072 pixels. The 1-look synthetic SAR image is obtained by multiplying the clean SAR image in Fig. 10 with 1-look speckle noise.

GRHED for the synthetic realistic 1-look SAR image can be found in Fig. 11, where the 1-look SAR image is obtained by multiplying the clean SAR image with 1-look speckle noise. What can be seen from Fig. 11 is that the F1-score curve of GRHED is above the F1-score curves of all the other methods. We observe also that the performances of GRHED are relatively robust to the choice of theb threshold. In addition, both HED and HED-log are shown to be more powerful than GR to detect edges in complex situations. The edge maps obtained with the threshold corresponding to pfa (10^{-5}) for different methods (for GR, the threshold is corresponding to pfa 10^{-3} in order to preserve more true edges) can be found in Fig. 12. From Fig. 12 we can see that GRHED detects more true edge pixels than the other methods and the F1-scores of GRHED is higher than with other methods. In addition, HED also provides very competitive edge detection results. GR detects more false detections probably because of the threshold chosen from a higher pfa, but it should be noted that the F1-score of GR is very close to its optimal value, according to Fig. 11. The method HED-log yields significantly more flase detection, possible as a result of not having CFAR. Due to the poor performances of HED-log, we do not use it for comparison in the following.

3) Further comparisons on real SAR images: In this part, we test the efficiency of GRHED with two 1-look real SAR images. We first have tested the efficiency of the state-of-art edge detectors AMDR [18] in SAR images, but we found it not suitable for such complex and noisy situations and did not investigate its behavior further. We then compare HED and GRHED with GR on a 1-look real SAR image (Leystad, Sentinel 1) of size 1024×3072 pixels as shown in Fig. 13. From Fig. 13 we can see that though GR is able to detect many true edges, it is not able to provide a good separation between true edge pixels and noise pixels. Therefore, there are also many false detections in the edge maps produced by GR. In comparison, both HED and GRHED detect many true edge pixels with clearly less false detections. However, HED produces spurious detections in very bright areas (maybe because of the lack of such zone in the training set). On the other hand, GRHED does not suffer from a similar problem. On this experiment for GRHED, the number of false detections is smaller and more evenly distributed, and edges appear to be better connected.

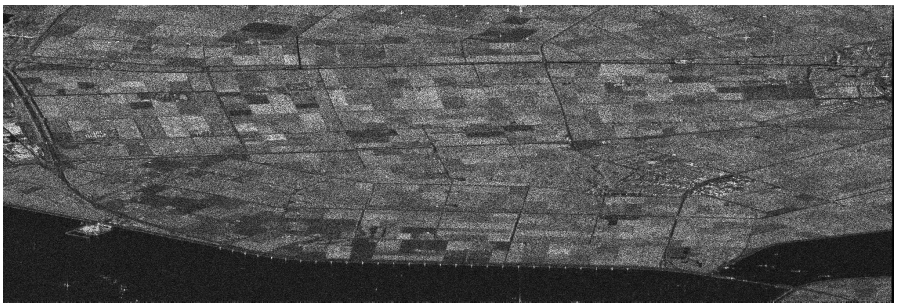
We compare GR, HED, and GRHED in another 1-look real SAR image (San Francisco, TerraSAR-X) of size 2048×2048 pixels. The edge maps obtained by these methods can be found in Fig. 14. Again, the weakness of GR is that it is not able to separate the true edge pixels and noise pixels efficiently. For HED, since many homogeneous areas with very high mean intensity values exist in the image, and these kinds of areas do not exist in the training images, it is very difficult for HED to perform effective edge detection on them. Therefore, the edge map outputed by HED has spurious detections in many bright areas. In comparison, those very bright areas do not cause troubles to GRHED because the gradient distribution computed by GR in homogeneous areas will not be influenced by their mean intensity values. What is more important, GRHED is able to separate true edge pixels and false edge pixels efficiently.

V. Conclusion

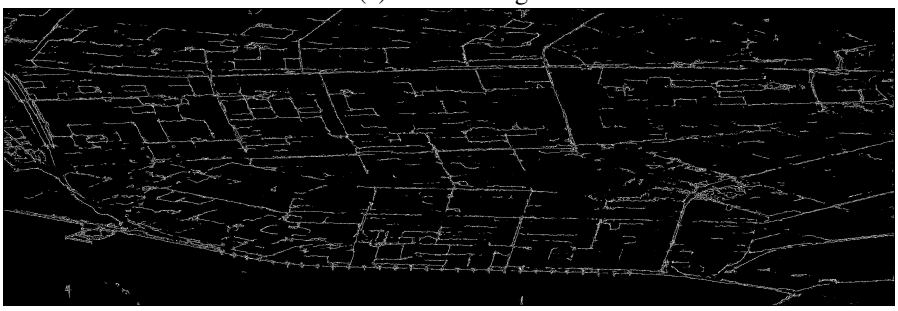
In this paper we addressed the challenging task of edge detection in 1-look real SAR images. Leveraging the available optical dataset, we proposed to develop a CNN-based edge detector for SAR images by training the CNN models on the gradient magnitude fields of speckled optical images. By introducing the fixed hand-crafted layer (GR) instead of a learnable one, the proposed CNN edge detector GRHED is much less influenced by the differences in the range of pixel values between speckled optical and real SAR images. GR ensures that all kinds of gradient distributions of real SAR images are very likely to be included in those of speckled optical images, but it should be pointed out that there may be some gradient distributions that do not exist in those of the training dataset. For example, when some edges with extremely high contrast appear in real SAR images, the ratio between the mean intensity values of the homogeneous areas on the opposite side of the edges is too high.

From all the experiments we can see that GRHED is able to obtain stable and reliable detection results. This is especially interesting in the case of one-look images.

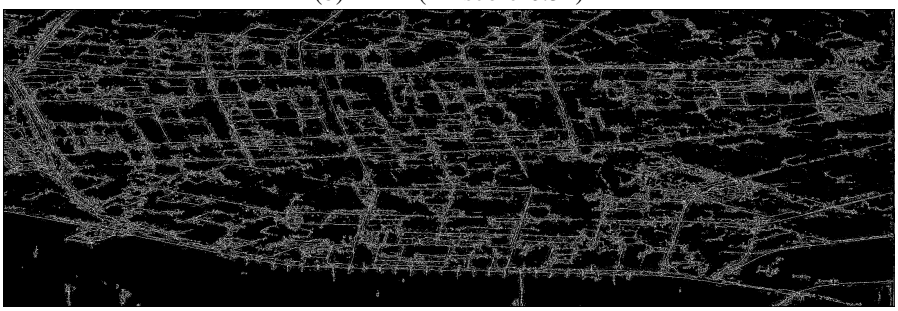
Among the points that have not been addressed in this work and will be addressed in further works, we can mention the followings. First, the spatial correlation of the noise on real images has not been addressed and probably leads to a decreasing of the CNN performances. The method of [37] could be an interesting approach to take it into account. Secondly, the specific features of SAR images like bright points and lines due to strong backscatterings of diedral or triedral structures do not exactly correspond to edges. Therefore dedicated detectors should be developed for these specific structures to be combined with edge detectors.



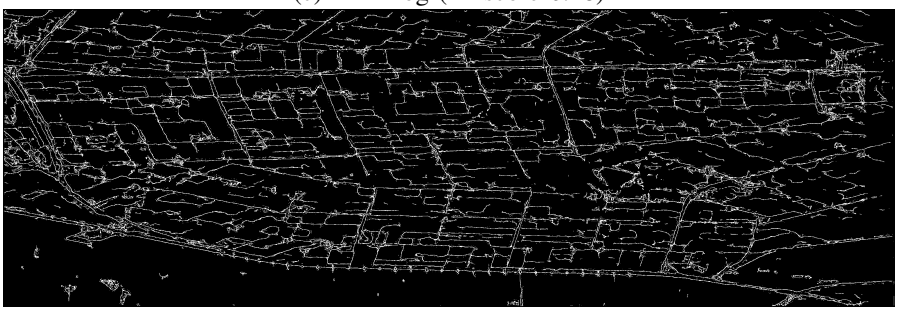
(a) 1-look image



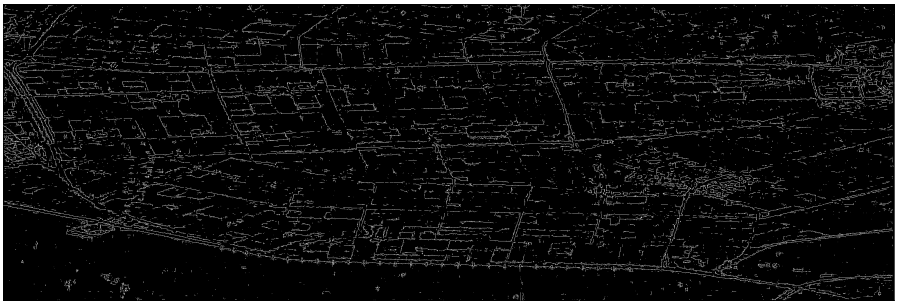
(b) HED (F1-score 0.34)



(c) HED-log (F1-score 0.28)

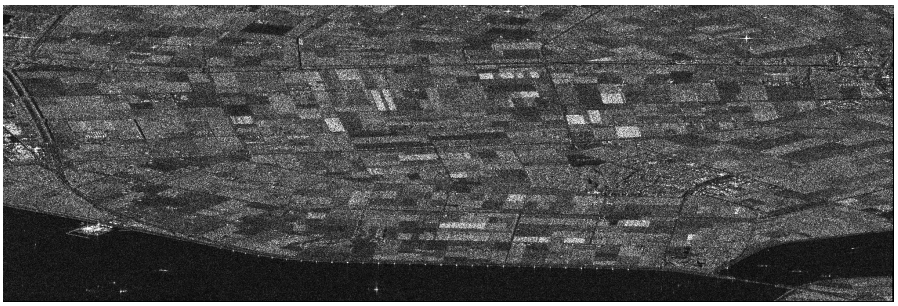


(d) GRHED (F1-score 0.38)

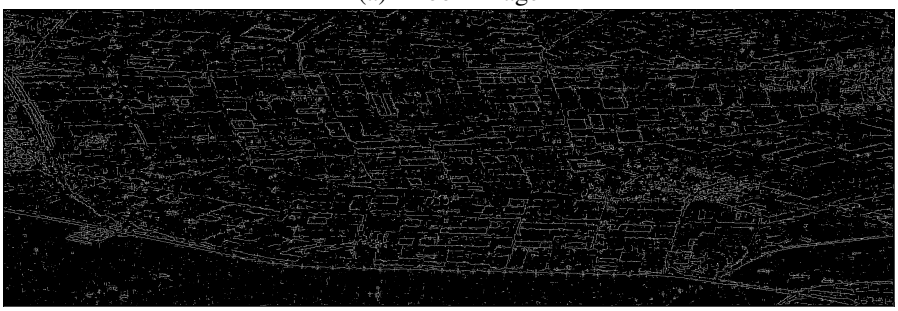


(e) GR (F1-score 0.28)

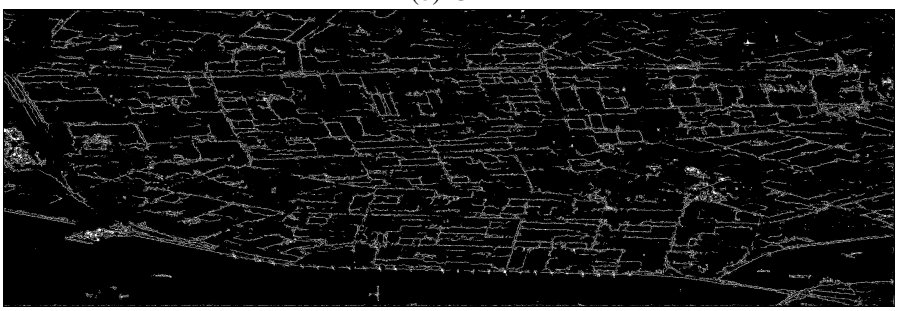
Fig. 12: Edge maps computed with a threshold corresponding to pfa (10^{-5}) on a synthetic realistic 1-look image (Leystad, Sentinel 1) for different methods. For GR, we use the threshold corresponding to pfa (10^{-3}), which is very close to the threshold corresponding to the best F1-score.



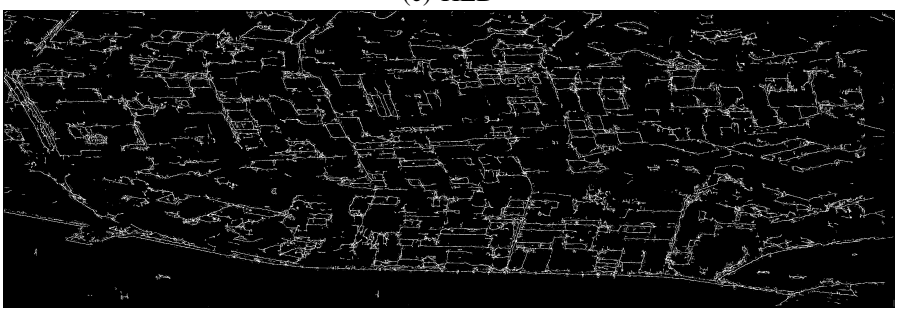
(a) 1-look image



(b) GR



(c) HED



(d) GRHED ($\alpha = 2, 3, 4, 5$)

Fig. 13: Comparison of different methods on a 1-look real SAR image (Leystad, Sentinel 1). The size of the image is 1024×3072 pixels.

REFERENCES

[1] Pablo Arbelaez, Michael Maire, Charless Fowlkes and Jitendra Malik, "Contour Detection and Hierarchical Image Segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, pp. 898–916, 2011.

[2] Saining Xie and Zhuowen Tu, "Holistically nested Edge Detection," in *2015 IEEE International Conference on Computer Vision*, 2015, pp. 1395–1403.

[3] ——, "Holistically-Nested Edge Detection," *International Journal of Computer Vision*, vol. 125, pp. 3–18, 2017.

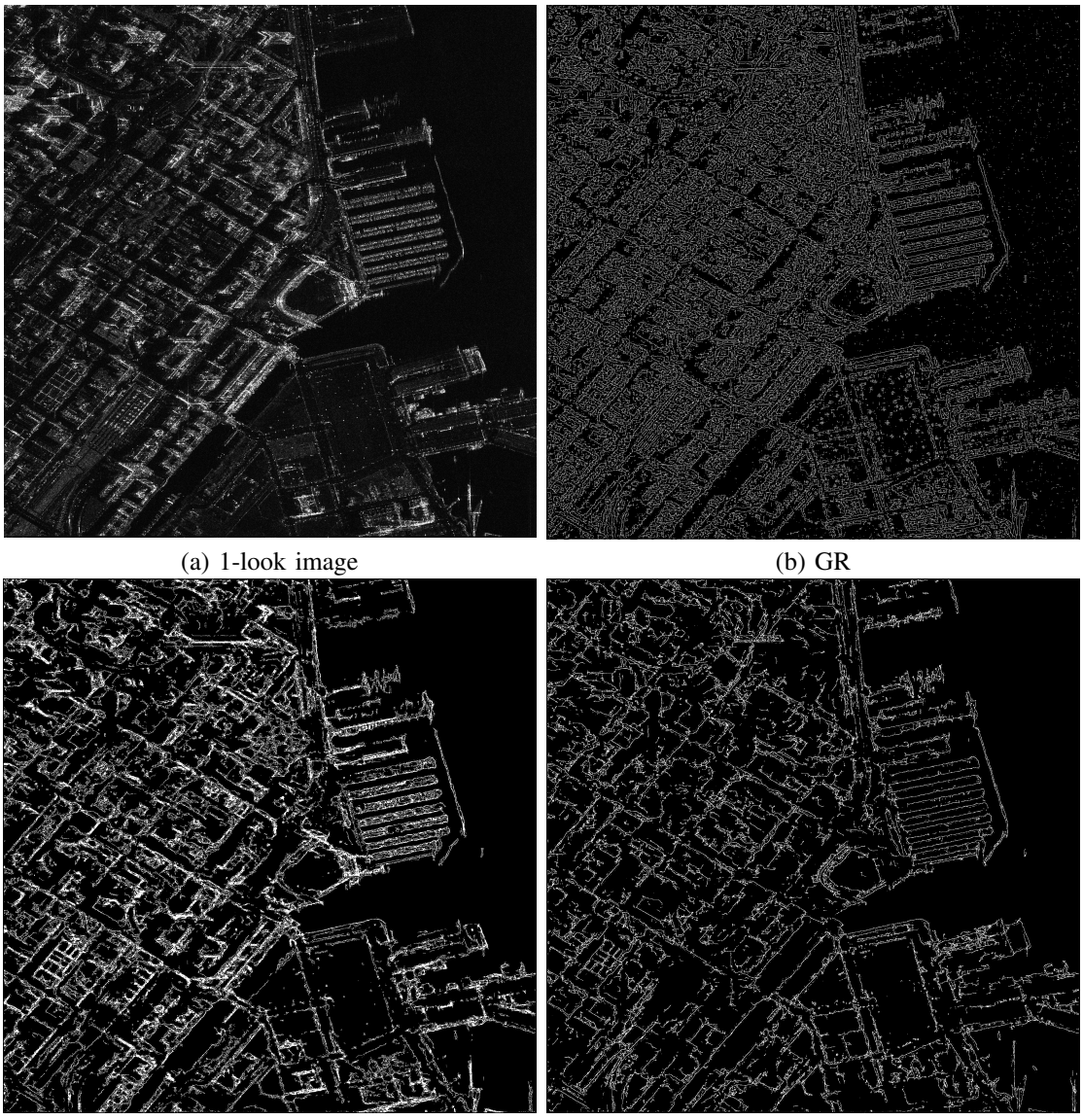
[4] Qian-Ru Wei and Da-Zheng Feng, "Extracting Line Features in SAR images Through Image Edge Fields," *IEEE Geoscience and Remote Sensing Letters*, vol. 13, pp. 540–544, 2016.

[5] Qian-Ru Wei, Da-Zheng Feng, Wei Zheng and Jiang-Bin Zheng, "Rapid Line Extraction Method for SAR Images Based on Edge-Field Features," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, pp. 1865–1869, 2017.

[6] Peter Yu, A.K. Qin and David A. Clausi, "Unsupervised Polarimetric SAR image Segmentation and Classification Using Region Growing with Edge Penalty," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, pp. 1302–1317, 2012.

[7] Huihui Song, Bo Huang and Kaihua Zhang, "A Globally Statistical Active Contour Model for Segmentation of Oil Slick in SAR Imagery," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 6, pp. 2402–2409, 2013.

[8] JONG-SEN LEE and IGOR JURKEVICH, "Coastline Detection and Tracing in SAR images ," *IEEE Transactions on Geoscience and Remote*



(a) 1-look image (b) GR

(c) HED (d) GRHED ($\alpha = 2, 3, 4, 5$)

Fig. 14: Comparison of different methods on a 1-look real SAR image (San Francisco, TerraSAR-X). The size of the image is 2048×2048 pixels.

Sensing, vol. 28, pp. 662–668, 1990.

[9] Chun Liu, Yingying Xiao and Jiao Yang, "A Coastline Detection Method in Polarimetric SAR images Mixing the Region-Based and Edge-Based Active Contour Models," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, pp. 3735–3747, 2017.

[10] Tianze Chen, Limin Chen and Yi Su, "A SAR Image Registration Method Based on Pixel Migration of Edge-Point Features," *IEEE Geoscience and Remote Sensing Letters*, vol. 11, pp. 906–910, 2014.

[11] Han Zhang, Weiping Ni, Weidong Yan, Junzheng Wu and Sha Li, "Robust SAR image Registration Based on Edge Matching and Refined Coherent Point Drift," *IEEE Geoscience and Remote Sensing Letters*, vol. 12, pp. 2115–2119, 2015.

[12] Min Dai, Cheng Peng, A.K. Chan and D. Loguinov, "Bayesian wavelet shrinkage with edge detection for SAR image despeckling," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, pp. 1642–1648, 2004.

[13] R. Touzi, A. Lopès, and P. Bousquet, "A statistical and geometrical edge detection for SAR images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 26, pp. 764–773, 1988.

[14] R. Fjørtoft, A. Lopès, P. Marthon, and E. Cubero-Castan, "An optimal multiedge detector for SAR image segmentation," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 36, pp. 793–802, 1998.

[15] P.-L. Shui and D. Cheng, "Edge Detector of SAR images Using Gaussian-Gamma-Shapped Bi-Windows," *IEEE Geoscience and Remote Sensing Letters*, vol. 9, pp. 846–850, 2012.

[16] Qian-Ru Wei, Da-Zheng Feng and Hu Xie , "Edge Detector of SAR Images using Crater-Shaped Window with Edge Compensation Strategy," *IEEE Geoscience and Remote Sensing Letters*, vol. 13, pp. 38–42, 2016.

[17] J. Canny, "A Computational Approach to Edge Detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-8, pp. 679–698, 1986.

[18] PENGLANG SHUI and SHUCHEN FAN, "SAR Image Edge Detection Robust to Isolated Strong Scatterers Using Anisotropic Morphological Directional Ratio Test," *IEEE Access*, vol. 6, pp. 37 272–37 285, 2018.

[19] Gedas Bertasius, Jiangbo Shi and Lorenzo Torresani, "Deepedge: A multi-scale bifurcated deep network for top-down contour detection," in *2015 IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 4380–4389.

[20] Wei Shen, Xinggang Wang, Yan Wang, Xiang Bai and Zhijiang Zhang, "DeepContour: A deep convolutional feature learned by positive-sharing loss for contour detection," in *2015 IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 3982–3991.

[21] Jimei Yang, Brian Price, Scott Cohen, Honglak Lee and Ming-Hsuan Yang, "Object Contour Detection with a Fully Convolutional Encoder-Decoder Network," in *2016 IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 193–202.

[22] Dan Xu, Wanli Ouyang, Xavier Mameda-Pineda, Elisa Ricci, Xiaogang Wang and Nicu Sebe, "Learning Deep Structured Multi-Scale Features using Attention-Gated CRFs for Contour Prediction," in *2017 Conferece on Neural Information Processing Systems (NIPS 2017)*, 2017.

[23] Yun Liu, Ming-Ming Cheng, Xiaowei Hu, Kai Wang, Xiang Bai, "Richer Convolutioonal Features for Edge Detection," in *2017 IEEE Conference on Computer Vision and Pattern Recognition*, 2017.

[24] Yun Liu, Ming-Ming Cheng, Xiaowei Hu, Jia-Wang Bien, Le Zhang, Xiang Bai, Jinhui Tang, "Richer Convolutional Features for Edge Detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 41, pp. 1939–1946, 2019.

[25] J Kittler, "On the accuracy of the Sobel edge detector," *Image and Vision Computing*, vol. 1, pp. 37–42, 1983.

[26] Scott Konishi, Alan L. Yuille, James M. Coughlan and Songchun Zhu, "Statistical edge detection: Learning and Evaluating Edge Cues," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, pp. 57–74, 2003.

[27] David R. Martin, Charless C. Fowlkes and Jitendra Malik, "Learning to detect natural image boundaries using local brightness, color, and texture cues," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, pp. 530–549, 2004.

[28] Piotr Dollar and C. Lawrence Zitnick, "Fast Edge Detection Using Structured Forests," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, pp. 1558–1570, 2015.

[29] F. Dellinger, J. Delon, Y. Gousseau, J. Michel, and F. Tupin, "SAR-SIFT: A SIFT-Like Algorithm for SAR Images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, pp. 453–466, 2015.

[30] Jonathan Long, Evan Shelhamer and Trevor Darrel, "Fully convolutional networks for semantic segmentation," in *2015 IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 3431–3440.

[31] Evan Shelhamer, Jonathan Long and Trevor Darrell, "Fully Convolutional Networks for Semantic Segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, pp. 640–651, 2017.

[32] Simonyan, K. and Zisserman, A., "Very Deep Convolutional Networks for Large-Scale Image Recognition," in *International Conference on Learning Representations*, 2015.

[33] Chen-Yu Lee, Saining Xie, Patrick Gallagher, Zhengyou Zhang and Zhuowen Tu, "Deeply-supervised nets," in *International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2015.

[34] J. Goodman, *Statistical properties of laser speckle patterns*. Laser Speckle and Related Phenomena, 1975, vol. ch. 2.

[35] Chenguang Liu, Rémy Abergel, Yann Gousseau and Florence Tupin, "LSDSAR, a Markovian a contrario framework for line segment detection in SAR images," *Pattern Recognition*, vol. 98, 2020.

[36] C. Deledalle, L. Denis, S. Tabti, and F. Tupin, "MuLoG, or how to apply gaussian denoisers to multi-channel sar speckle reduction?" *IEEE Transactions on Image Processing*, vol. 26, no. 9, pp. 4389–4403, Sept 2017.

[37] A. Lapini, T. Bianchi, F. Argenti, and L. Alparone, "Blind speckle decorrelation for SAR image despeckling," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 2, 2014.