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

Chenguang Liu, Florence Tupin

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

Chenguang Liu and Florence Tupin

***Abstract*—Edge detection in SAR images remains a difficult task due to the strong multiplicative noise. Many researches have been dedicated to edge detection in SAR images but very few of them try to address the most challenging situation, namely the edge detection in 1-look real SAR images. Motivated by the success of Convolutional Neural Networks (CNNs) in natural images, we study the applicability of the usual 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 the training dataset with labeled ground truth. Instead of doing the tedious job of annotating plenty of SAR images, we simulate a SAR dataset leveraging the optical dataset BSDS500 [1] to train CNN models because edges are mainly corresponding to changes in brightness and textures in grayscale images. In order to cope with the differences in the range of pixel values between SAR and optical images, we propose to train CNN models on the gradient magnitude fields of images because the differences in the gradient distribution between speckled optical images and real SAR images are tiny and insignificant. In the gradient feature space, the gradient magnitude fields of homogeneous areas follow exactly the same distribution regardless of their mean intensity values, and the distribution of gradient magnitude fields for two homogeneous areas across boundaries depends only on the ratio of their mean intensity values. The proposed CNN edge detector GRHED achieves excellent performances in all our simulations including several 1-look synthetic edge images with different ratio contrasts, two hundred 1-look optical images, which are simulated from BSDS500, one synthetic SAR image and two 1-look real SAR images. In addition, it exceeds the existing edge detectors in SAR images a lot.**

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

I. Introduction

Edges are very important features in Synthetic Aperture Radar (SAR) images. The edge features can be used as low level features for many applications like line segment detection [2], [3], SAR image segmentation [4], [5], coastline detection [6], [7], image registration [8], [9] and SAR image despeckling [10]. 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 [11], the Ratio of Average (ROA) was proved to have a constant false alarm rate for SAR images. The ratio operator was performed along four directions and the minimum normalized ratio was used to compute the gradient magnitude. The direction corresponding to the minimum normalized ratio was regarded as the edge orientation. A threshold determined by a given probability of false alarm rate and the morphological operator were 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 [12]. 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 multiedge model. The average was 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 was computed along the horizontal and vertical directions and the normalized ratios were considered as the horizontal and vertical components of the gradient magnitude. The modified watershed algorithm was then used to threshold the Edge Strength Map and a region merging algorithm was used to eliminate the false edge pixels. Edge detectors using different shape of window functions were introduced later in [13] and [14]. Non-maxima Suppression [15] and hysteresis thresholding were then applied to obtain the binary thin edge map. An edge compensation strategy was also introduced in [14] to extract weak edge pixels. To reduce the influence of isolated strong bright points in real SAR images, an Anisotropic morphological directional ratio (AMDR) [16] 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 [15] edge detector, 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 far from satisfying, especially in the very noisy 1-look situation. Motivated by the success of the convolutional neural networks (CNNs) for edge detection in natural images, such as Deepedge [17], DeepContour [18], HED [19], [20], CEDN [21], AMH-Net [22] and RCF [23], [24], all of which achieve comparable or even better performances than the traditional edge detectors like Sobel edge detector [25], Canny [15], Statistical Edge [26], Pb [27], gPb [1] and Structured Edge [28], we study the possibility to apply CNN-based methods to SAR images.

One crucial factor that contributes to the success of CNNs is the training dataset with labeled ground truth, while there

Chenguang Liu and Florence Tupin are with Télécom Paris, Institut Polytechnique de Paris, France (email: chenguang.liu@telecom-paris.fr; florence.tupin@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.

is still no available training dataset for edge detection in SAR images. Considering that the edges are mainly corresponding to changes in local brightness and textures (and color for color images), the problem of edge detection should rely little on the targets of the images. In this case, the models trained on the optical datasets should be applicable to SAR images. Therefore, we simulate a SAR dataset by multiplying speckle noise with the images in the optical dataset.

Even though the noise models are similar for speckled optical images and real SAR images, the differences in the range of pixel values between them can not be neglected. In a real SAR image, the pixel values can go up to several thousand or much larger, while for all the images in the simulated SAR dataset, the maximum value is approximately 1000. Very large pixel values in real SAR images could exist in homogeneous areas with high mean intensity values (much larger than 255), or be corresponding to the bright targets. Therefore, training CNNs directly on the images is not feasible because the trained model does not learn the way to process those pixels. To deal with this problem, we propose to train CNNs on the gradient magnitude fields of the images and apply the trained model to the gradient magnitude fields of testing images. For the gradient computation methods dedicated to SAR images, the usual ratio based methods [11], [12], [29] have a constant false alarm rate (CFAR) for SAR images, which implies that the gradient magnitude fields of the homogeneous areas computed by the ratio based methods follow the same distribution regardless of their mean intensiy values. The gradient distribution for two homogeneous areas across boundaries does not depend on their mean intensity values, but only on the ratio of them. Therefore, unless some highly contrasted edges exist in real SAR images, there is very minor gap between the gradient magnitude fields of SAR and those of speckled optical images. The gradient computation is done by Gradient by Ratio (GR) [29], and the trainable layers are identical to those in HED [19], [20]. The gradient computation step aims at finding the changes in brightness, and HED layers are responsible for strengthening true edge pixels while suppressing the noise pixels. The proposed edge detector is thus named as GRHED. In the next section, we will give the details about HED and the details about the simulated speckled optical dataset.

II. The details of HED and the training dataset

A. The details of HED

HED is a fully convolutional network [30], [31] and thus it can be trained end-to-end and perform image-to-image prediction. The network architecture of HED is shown in table I. 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 in the conv layers just before the five max-pooling layers in the original VGG-16 net. These side outputs and the fused output obtained from the fusion of them are all supervised by the edge ground truth. The resulting loss function of the HED net is composed of the loss from the side layers and the loss from the fused outputs. 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 all the side outputs have the same size as the edge ground truth. The reason for upsampling is caused by the subsampling of 2 after each max-pooling function.

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. Instead of doing the tedious job of edge labeling, we leverage the available optical dataset for edge detection in natural images motivated by the fact that the edges are mainly corresponding to the changes in brightness and textures so the task of edge detection should not be influenced a lot by the image contents.

In order to simulate a SAR dataset, it is very important to know the statistics of SAR images. Due to the coherent imaging system, SAR data presents the well known speckle phenomenon. Following Goodman model [34] of fully developped speckle, it can be shown that the amplitude of the backscattered electro-magnetic 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 a system parameter. For images with the best resolution, $L = 1$ and amplitude of a homogeneous area follows a Rayleigh distribution. Another way of modeling the 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 $L = 1$. Fig. 1 shows the distribution of s.

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

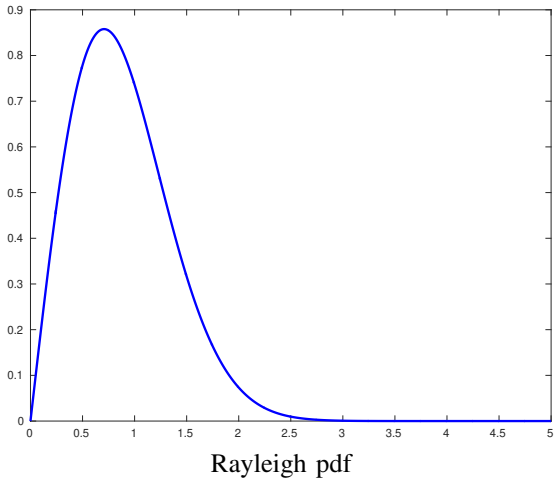


Fig. 1: The pdf of a 1-look speckle noise.

In this paper, we use the Berkeley Segmentation Data Set and Benchmarks 500 (BSDS500) [1] with the same data augmentation as in HED [20] 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 the 50%, 100%, 150% of its original size. There is no data augmentation for the testing images. The resulting speckled optical dataset, which we will call BSDS500-speckled in the following, is formed by multiplying the grayscale images converted from color images in BSDS500 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, especially how to enable the trained model to work well in real SAR images. We will describe the way to tackle the gap between the training and testing data in the next section.

III. GRHED, Introducing a hand-crafted layer before the traditional CNNs

Even though speckled optical images and real SAR images are both contaminated by the speckle noise, the gap between them remains large. Specifically, the range of pixel values differs a lot. The maximum pixel value in a natural image is usually not larger than 255, which means that the mean intensity value of any homogeneous areas in the training dataset will not be larger than 255^2, while in SAR images the mean intensity values of some homogeneous areas could be much larger. Many pixel values in a real SAR image could be out of the range of pixel values in the speckled optical dataset. These pixels could exist in homomgeneous areas with high mean intensity values or be corresponding to bright targets. Therefore, It is quite questionable to apply a model trained directly on the speckled optical images to real SAR images.

In order to deal with the differences in the range of pixel values between speckled optical images and real SAR images, we propose to train HED on the gradient magnitude fields of the training images and apply the trained model to the gradient magnitude fields of testing images. 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 for SAR images [11], [12], [29]. The gradient computation is done by Gradient by Ratio [29] and we train HED on the gradient magnitude fields computed by GR, therefore, the resulting architecture is named as GRHED.

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

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

where $R^h(x,y)$ and $R^v(x,y)$ is the ratio of exponentially weighted average in the opposite side windows of pixel located at (x, y), along the horizontal and vertical directions. 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

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

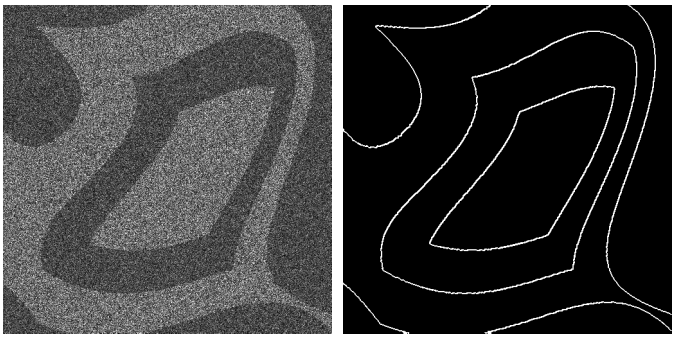
and where W is the upper integer part of $\log(10) \times \alpha$. $R^v(x,y)$ can be computed in a similar way.

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

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

B. Computing GR, fusing different distributions

An image, whatever optical or SAR, can be roughly divided into two parts: 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, the

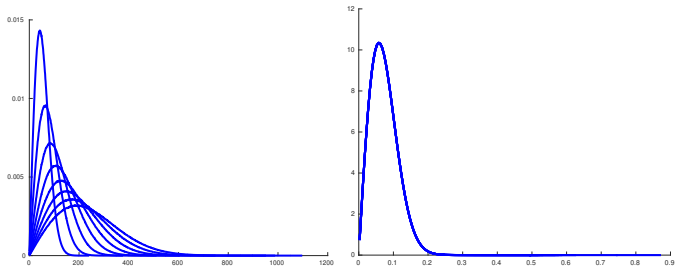


(a) 1-look image (b) GT

Fig. 2: A 1-look edge image with ratio contrast 1.5 and the associated ground truth. The size of the image are 512×512 pixels.

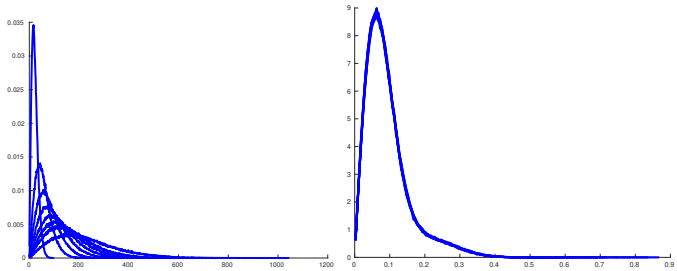
total distribution of them depends on the mean intensity values of both two homogeneous areas. The main gap between speckled optical images and real SAR images is that SAR images may contain homogeneous areas with very high mean intensity values, the mean intensity values of which could be much larger than the maximum possible one in the training dataset. The models trained directly on the speckled optical images do not learn the ability to suppress the pixels in those areas, and they do not learn the way to highlight the edge pixels between those areas and the other areas.

The usual ratio based gradient computation methods [11], [12], [29] ensure a constant false alarm rate for SAR images, which means that for all homogeneous areas, their gradient magnitude fields follow exactly the same distribution, regardless of their mean intensity values. In addition, the gradient distribution computed over two homogeneous areas across boundaries depends only on the ratio of the mean intensity values. Therefore, the distribution of the gradient magnitude fields of speckled optical images and SAR images should be similar except for the situation that some highly contrasted edges exist in real SAR images but not in speckled optical images. In order to give a better explanation about this, we compare the distribution of the data before 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 edge images of size 512×512 pixels with amplitude ratio contrast 1.5. The mean intensity values of the homogeneous areas are 60^2, 90^2, 120^2, 150^2, 180^2, 210^2, 240^2 and 270^2. For two homogeneous areas across boundaries, the relationship between the square root of the mean intensity values and amplitude ratio contrast is as follows: $\frac{20*1.5}{20}$, $\frac{50*1.5}{50}$, $\frac{70*1.5}{70}$, $\frac{90*1.5}{90}$, $\frac{110*1.5}{110}$, $\frac{130*1.5}{130}$, $\frac{150*1.5}{150}$, $\frac{200*1.5}{200}$. One example of the synthetic edge image can be found in Fig. 2-(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. 3. The histograms of eight 1-look synthetic edge images as well as the histograms of their gradient magnitude fields can be found in Fig. 4. From Fig. 3 and Fig. 4 we can deduce that the number of distributions that has to be learned by HED when trained



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

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



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

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

on the gradient magnitude fields is much smaller than that of a direct training on the images, Since the amount of data is unchannged, the amount of training data for each distribution becomes larger. What's more important, the problem caused by the homogeneous areas with very high mean intensity values does not exist since the gradient distribution of these areas is the same as those of homogeneous areas with small mean intensity values. In addition, if a boundary exists between one of these areas and another area, the total distribution depends only on the ratio of their mean intensity values.

C. Justification of training on the gradient magnitude field according to the observation in existing neural networks

1) Fully convolutional neural networks: The main contributions of the paper introducing FCN [30], [31] lie mainly on two points: 1) converting existing networks so that all the layers in the network are convolutional layers, which makes the network can be learned end to end, and can be applied to images of any size; 2) adding the skip connection so that local and global information are combined together to obtain pixel-wise precision segmentation. The local information obtained from shallower layers has better location precision of the object boundaries because less subsampling steps have been done. The global information obtained from high level layers is more semantically meaningful thanks to the increasing size of receptive field so that they are able to take contextual information into account. We can assume that the shallower layers are mainly responsible for capturing the gradient infor-

mation, while the high level layers are trained on the gradient information and learn to highlight, connect the edge pixels while suppressing the false detections.

2) ResNet: One preconditioning in ResNet [35] is that the input and output of a building block (2 or more conv layers) are similar signals, They assume and prove that it is much easier to learn the residual than to learn the unreferenced signal. That is to say, the feature maps passing through residual blocks carry similar information. The higher level layers are trained on the feature maps similar to its output in order to enhance its performance.

3) Deeply supervised nets, HED, and RCF: The deeply supervised nets DSN [33] is motivated by the fact that a classifier trained on highly discriminative features gives better performances than a classifier trained on less discriminative features. Therefore, the transparency of the hidden layer feature maps can be a proxy for the performances of the final classifier. By directly supervising the hidden layer feature maps with the ground truth, the proposed DSN shows superior performances than those without deep supervision. HED [19], [20] leverages the advantages of deeply supervised nets and enforces the feature maps in hidden layers for an accurate prediction of the edges. RCF [23], [24] makes a stronger supervision by supervising the output feature maps of all hidden conv layers with the edge ground truth. Thus, the feature maps of the hidden layers can be assumed to carry the gradient information in the case of edge detection and the high level layers are trained on the gradient feature maps outputed by the low level layers.

D. GRHED: training HED on the gradient feature maps produced by GR

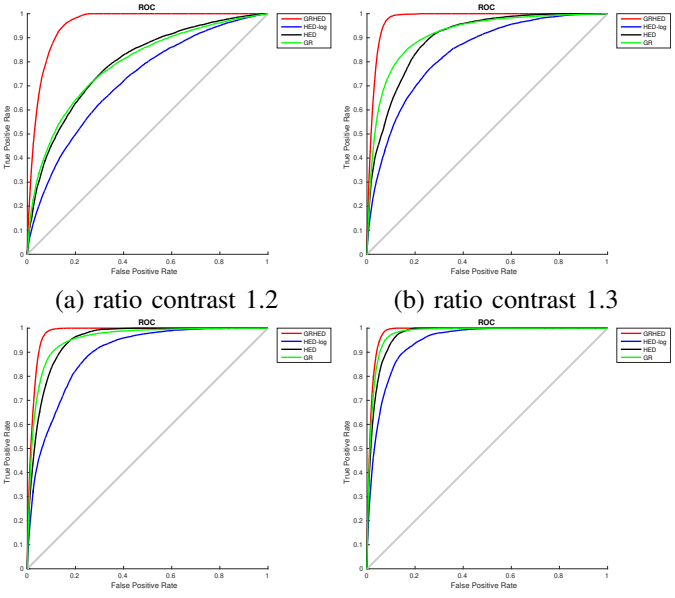
Since the gradient distribution of GR is the same for homogeneous areas with all possible mean intensity values, and the gradient distribution in two homogeneous areas across boundaries is influenced by the ratio and not by their mean intensity values, we assume that the kinds of distributions in the 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 regarding the gradient computation step as a hand-crafted layer before 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 will not cause trouble to 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 information from the image by concatenating together gradient magnitude fields produced by GR with different α values. It has been studied in [36] that GR with different α values can capture complementary information.

IV. Experiments

In this section we are going to demonstrate the efficiency of GRHED in several 1-look synthetic edge images, two hundred 1-look speckled optical images in BSDS500-speckled, one 1-look synthetic SAR image and two 1-look real SAR images. In order to show the efficiency of GRHED, we compare it with HED (trained directly on the images so we call HED directly) and HED-log (trained on the logarithm of the images). Taking logarithm of the images can change multiplicative noise into additive noise and it is a usual way to deal with multiplicative noise in SAR images. For GRHED, HED and HED-log, the conv layers that need to be trained are the same, the difference is just the kind of input data. For GRHED, the gradient magnitude fields of images are the input. For HED, the input are the images themselves and for HED-log the input are the logarithm of images. The training strategy for those conv layers is as follows: we use Adam optimizer to train the conv layers 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 as the one in Structured Edge [28] and use a threshold to discard pixels with low values in the probability edge map.

A. Comparison in 1-look synthetic edge images



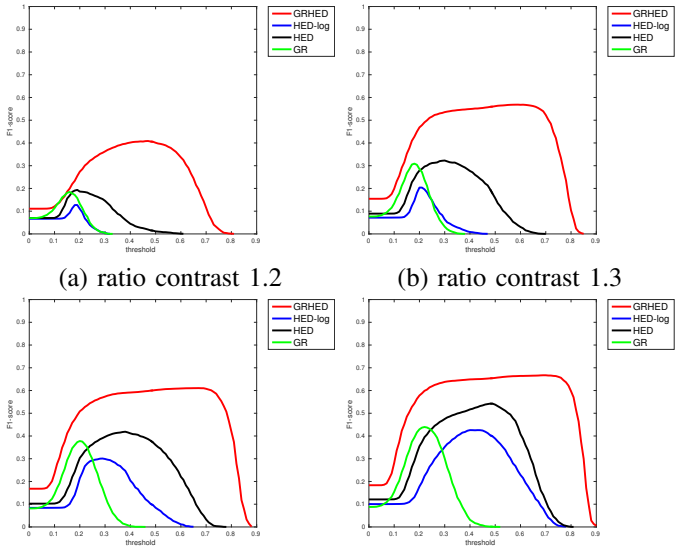
(a) ratio contrast 1.2 (b) ratio contrast 1.3

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

Fig. 5: ROC curves computed in 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 are 512×512 pixels.

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

- GR with $\alpha = 4$: The reason using $\alpha = 4$ is that it seems to be a good choice for GR in 1-look situations as discussed in [36]. For the GR magnitude field, we use the same postprocessing steps as the magnitude field produced by HED, HED-log and GRHED;
- HED: HED is trained on the images and tested on the images;
- HED-log: HED is trained on the logarithm of the images in BSDS500-speckled, and is applied to the logarithm of testing images;
- GRHED with multiple α values, $\alpha = 2, 3, 4, 5$: HED is trained on the gradient feature maps which are constituted by concatenating the gradient magnitude fields produced by GR with different α values.



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

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

Fig. 6: 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.

One example of the simulated 1-look edge images with contrast 1.5 and the corresponding ground truth can be found in Fig. 2. 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. 5. It can be seen from Fig. 5 that GRHED shows the best performances in all situations, especially in 1-look and low contrast situation. The performances of GR and HED appears comparable in terms of ROC curves, and both are better than HED-log.

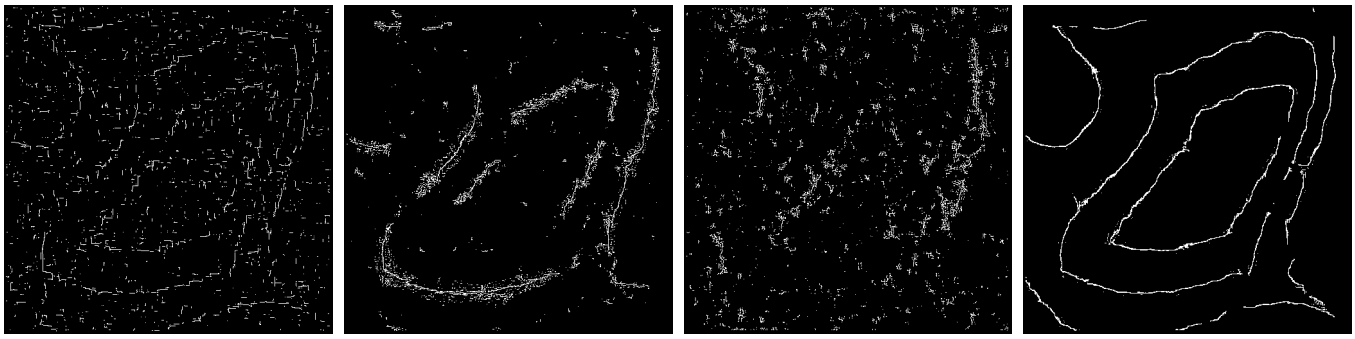
In order to give a more clear comparison of different methods, we apply the same Non-maxima suppression step for all methods and vary the threshold from 0.0 to 0.9 with step 0.01 for each method. The F1-score is computed for each threshold and the corresponding F1-score curves for each method in the 4 simulated edge images can be found in Fig. 6. From Fig. 6 we can see that the F1-score curves of GRHED are above the F1-score curves of all the other methods. In addition, the large flat areas of F1-score curves of GRHED show that the performances of GRHED are very insensitive to the chosen threshold. In addition, it should also be pointed out that although the best F1-scores that can be obtained by GR and HED seem comparable, the performances of HED are less sensitive to the choice of threshold.

The optimal edge maps obtained by GR, HED, HED-log and GRHED in those simulated edges with contrast 1.2 and 1.4 using the threshold which gives the best F1-score can be found in Fig. 7 and Fig. 8. From Fig. 7 and Fig. 8 we can see that the ability of GRHED to preserve true edge pixels and suppress noise pixels are much more powerful than all the other methods. GRHED detects the most true edge pixels while detecting 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 to be reasonable with the threshold in a relatively large range as can be deduced from Fig. 6, especially for GRHED.

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

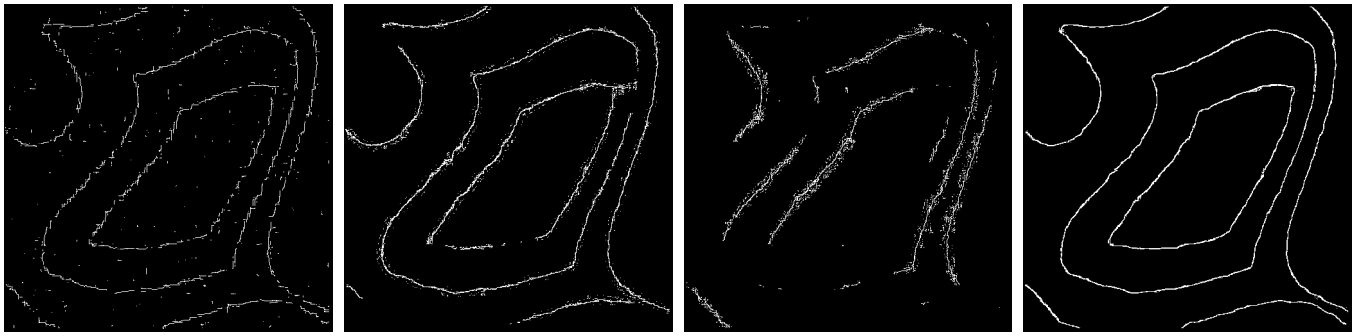
In order to give a more comprehensive comparison in more general situations between different methods, we compare GR, HED, HED-log and GRHED in the two hundred 1-look speckled optical images in BSDS500-speckled. For GR, we use $\alpha = 2$ and $\alpha = 4$ to show its performance. For GRHED, GRHED with a single α value and GRHED combining multiple α values are all provided to give a clearer demonstration of the efficiency for choosing the combination of α values. 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 in 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 in the case of combining multiple α values of GR. It should also be noted that, GRHED combining multiple α values give better performances than GRHED using a single α value, which is probably due to the richer information which is provided to HED conv layers.

The edge maps obtained by GR ($\alpha = 2$), GR ($\alpha = 4$), HED, HED-log, GRHED ($\alpha = 2$) and GRHED ($\alpha = 2, 3, 4, 5$) on one speckled optical image can be found in Fig. 9. 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 ($\alpha = 2, 3, 4, 5$) will obtain near-optimal edge maps in the simulated edge images as shown in Fig. 6. The ground truth of this images is the one provided by 1 labeler (there are usually five labelers for each



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

Fig. 7: 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 are 512×512 pixels.



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

Fig. 8: 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 are 512×512 pixels.

image in BSDS500). From Fig. 9 we can see that GRHED detects a bit more true edge pixels than HED, and both of them preserve more true edge pixels while detecting less false detections. What can also be concluded from the figure is that a direct processing of the GR magnitude field can not obtain satisfying results. Using HED as a postprocessing of GR field, the GRHED is able to strengthen the true edge pixels and suppress the false detections.

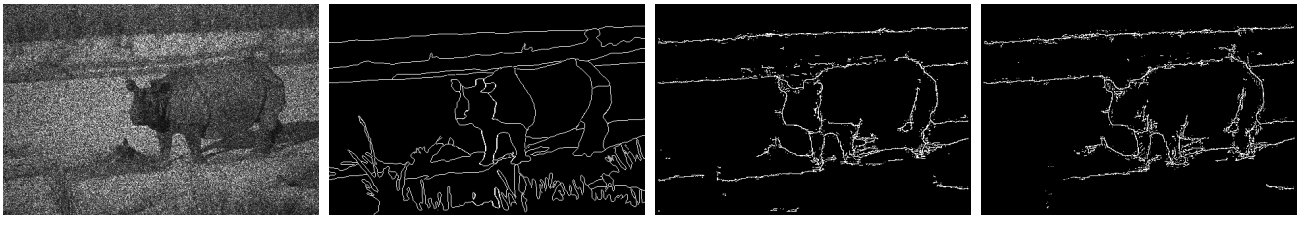
C. Comparison of different methods in 1-look SAR images

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

1) Setting threshold according to a given probability of false alarm rate: For the gradient computation method dedicated to SAR images, only if it has a constant false alarm Rate (CFAR), it is possible to set the threshold according to a chosen probability of false alarm rate (pfa). Therefore, it is important to study whether HED, HED-log and GRHED have CFAR for SAR images. The CFAR property of these methods can be checked experimentally by ploting the histograms of their gradient magnitude fields computed in pure noise images (hommogeneous areas with different mean intensity 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, HED, HED-log and GRHED in 160 1-look pure noise images of size 1024×1024 pixels in Fig. 10. The mean intensity values of these pure noise images have eight possibilities, 50^2, 80^2, 120^2, 150^2, 180^2, 200^2, 230^2 and 250^2. For pure noise images with a certain mean intensity value, there are 20 random realizations. From Fig. 10 we can see

TABLE II: The performances of different methods in 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



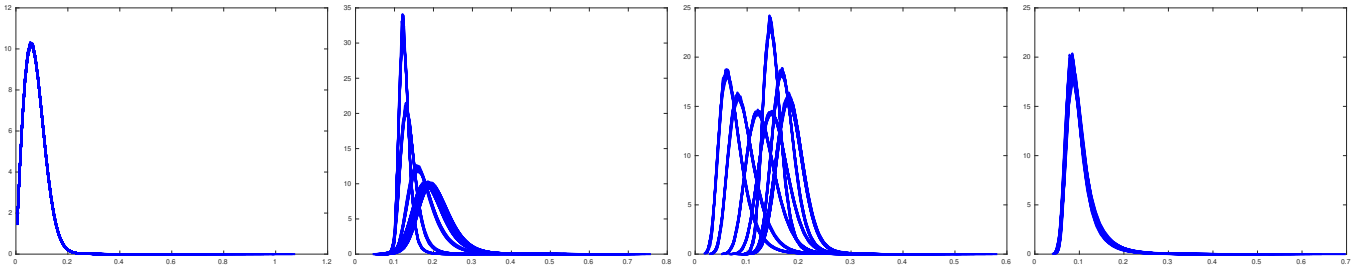
(a) speckled image (b) GT (c) HED (d) HED-log

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

Fig. 9: Comparison of different edge detectors on a 1-look speckled optical images.

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. 10: 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 mean intensity values of these images, namely, 50^2, 80^2, 120^2, 150^2, 180^2, 200^2, 230^2 and 250^2. For noise images with each mean intensity value, there are 20 random realizations.

that GR and GRHED both hold CFAR, but HED and HED-log do not.

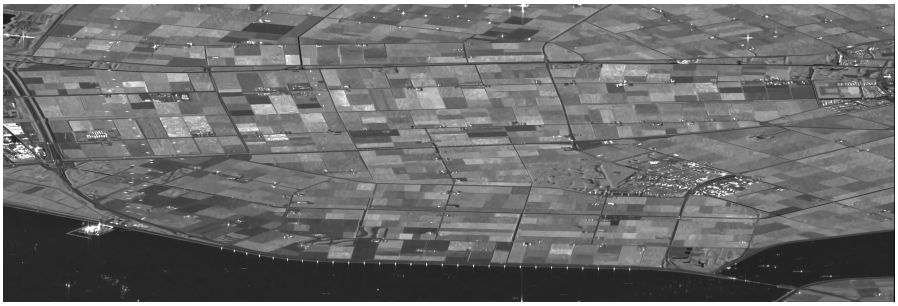
Even though HED and HED-log do not have CFAR for SAR images, we still assume they have CFAR since we desire to set the threshold from a given pfa. The threshold corresponding to a given probability of false alarm rate for GR, HED, HED-log and GRHED is estimated from 160 pure noise images as shown in 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 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 obtain a 'ground truth' by applying HED-clean (HED trained on the grayscale images converted from clean natural images in BSDS500) to a clean SAR image. The clean SAR image is obtained by clipping the denoised multi-look real SAR images (Leystad, Sentinel 1) with the value of its mean value plus three times of its standard deviation and then changing the range of the clipped images to be $[0, 255]$. The clean SAR image with its associated ground truth can be found in Fig. 11.

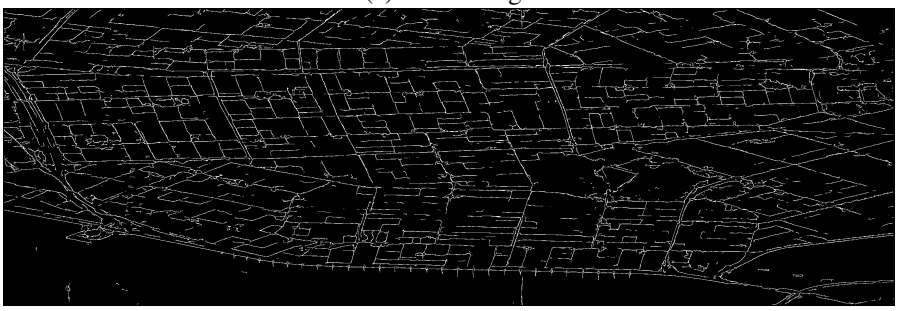
TABLE IV: The threshold corresponding to a given pfa for each method.

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

F1-score curves computed for GR, HED, HED-log and GRHED in the synthetic 1-look SAR image can be found in



(a) clean image



(b) GT

Fig. 11: The denoised multi-look image (Leystad, Sentinel 1) and its ground truth. The size of the image are 1024×3072 pixels.

Fig. 12, where 1-look SAR image is obtained by multiplying the clean SAR image with 1-look speckle noise. What can be seen from Fig. 12 is that the F1-score curve of GRHED is above the F1-score curves of all the other methods and the performances of GRHED is not sensitive to the choice of 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. 13. From Fig. 13 we can see that GRHED detects more true edge pixels than the other methods and the F1-scores of GRHED is higher than all the 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 the optimal one for GR according to Fig. 12. There are too many false detections for HED-log, this could probably be because HED-log is too far from CFAR, the strategy to choose threshold is not suitable for it. Due to the poor performances of HED-log, we do not compare with it in the following.

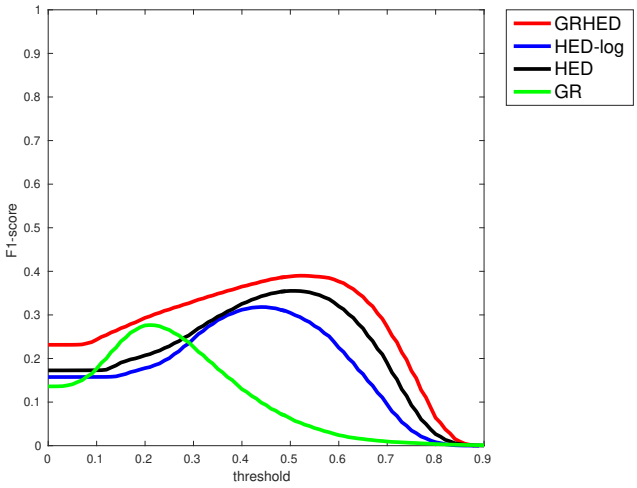
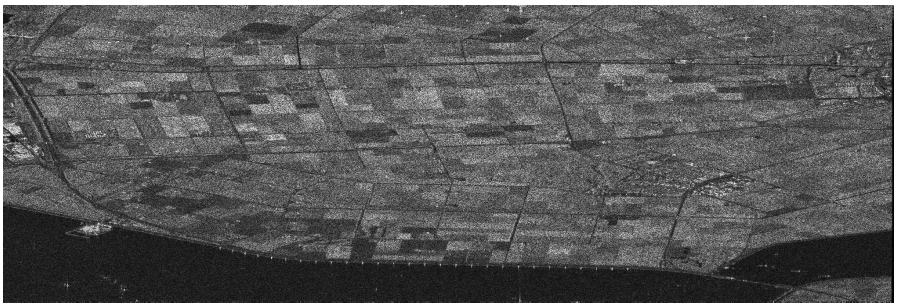
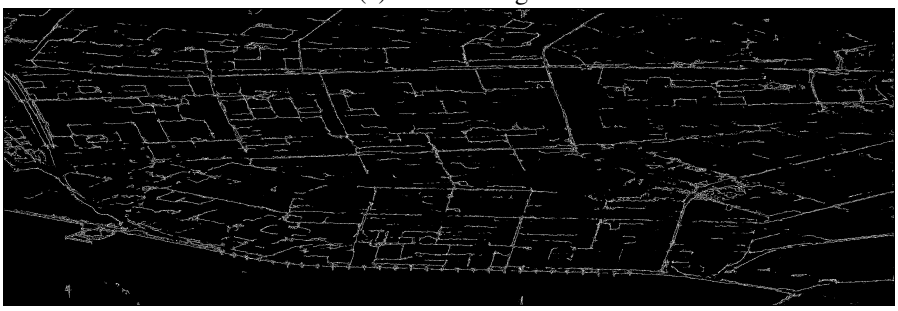


Fig. 12: F1-score curves computed for GR, HED, HED-log and GRHED in a 1-look synthetic SAR image of size 1024×3072 pixels. The 1-look synthetic SAR image is obtained by multiplying the clean SAR image in Fig. 11 with 1-look speckle noise.

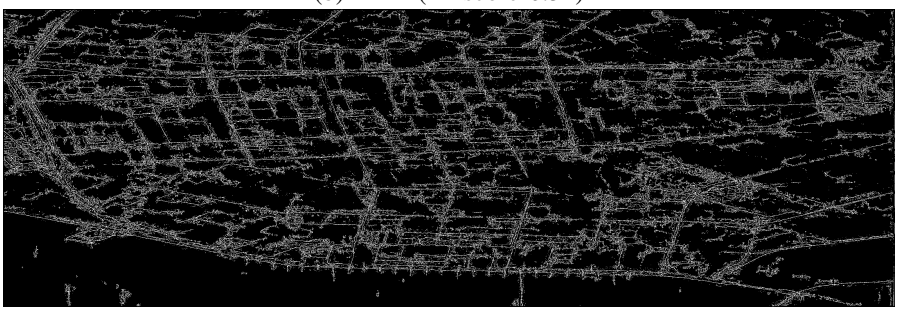
3) Comparison of different algorithms in two 1-look real SAR images using testing threshold corresponding to pfa (10^{-5}): In this part, we test the efficiency of GRHED in two 1-look real SAR images. We first test the performances of the state-of-art edge detector AMDR [16], but we find that it is not suitable for such complex and such noisy situations so we do not provide results with it. We compare HED and GRHED with GR in a 1-look real SAR image (Leystad, Sentinel 1) of size 1024×3072 pixels as shown in Fig. 14. From Fig. 14 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 using GR. In comparison, both HED and GRHED detect many true edge pixels while the number of false detections seems reasonable. However, HED produces spurious detections in very bright areas, this could be probably because the distributions of these areas are not included in those of the training dataset, HED does not learn the way to process them. Compared



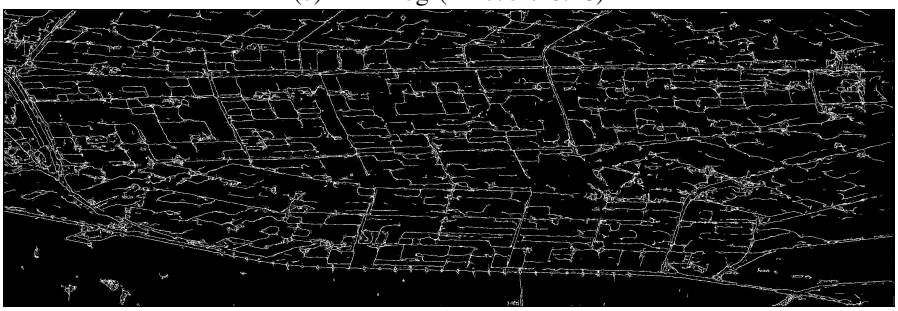
(a) 1-look image



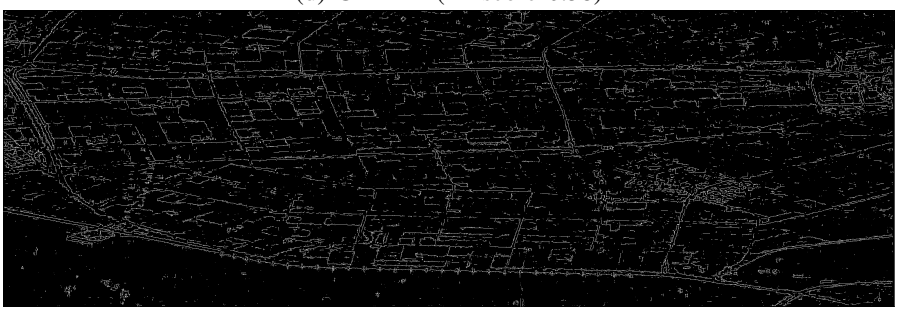
(b) HED (F1-score 0.34)



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



(d) GRHED (F1-score 0.38)



(e) GR (F1-score 0.28)

Fig. 13: edge maps computed with a threshold corresponding to pfa (10^{-5}) in a synthetic 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.

to HED, GRHED does not suffer from a similar problem. The detections of GRHED in all areas seem reasonable, the number of false detections is very small, and it detects many true edges. In addition, the connectivity of edges detected by GRHED is better than that of HED.

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. 15. 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's more important, GRHED is able to separate true edge pixels and false edge pixels efficiently.

V. Conclusion

In this paper we address the challenging task of edge detection in 1-look real SAR images. Leveraging the available optical dataset, we propose 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 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. We ignore this kind of situations because: first, in homogeneous areas with very high mean intensity values, the probability of false alarm rate is controlled; second, highly contrasted edges could also exist in speckled optical images because the ratio between two values drawn from $[0, 255]$ could also be very large, so the occurrences of extremely high-cotrasted edges (higher than those in speckled optical images) should be rare; third, even when this kind of edges occurs, the number of false detections is very small compared to the total number of detections and thus it is insignificant (the width of this kind of areas will not be larger than 25 pixels when the maximum value of α used in GR is not larger than 5, the length of the area depends on the length of the edges). From all the experiments we can see that GRHED is able to obtain stable and reliable results in all situations. In addition, introducing the hand-crafted layer GR enables GRHED trained on the speckled optical images to work well in real SAR images. Furthermore, to the best of our knowledge, this is the first edge detector that can achieve reasonable results in very noisy and very complex scenes, and it exceeds the existing edge detectors a lot.

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] Q. Wei and D. Feng, "Extracting Line Features in SAR Images Through Image Edge Fields," *IEEE Geoscience and Remote Sensing Letters*, vol. 13, pp. 540–544, 2016.

[3] Q. Wei, D. Feng, W. Zheng, and J. Zheng, "Rapid Line-Extraction Method for SAR Images Based on Edge-Field," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, pp. 1865–1869, 2017.

[4] 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.

[5] 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.

[6] JONG-SEN LEE and IGOR JURKEVICH, "Coastline Detection and Tracing in SAR images ," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 28, pp. 662–668, 1990.

[7] 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.

[8] 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.

[9] 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.

[10] 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.

[11] 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.

[12] 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.

[13] 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.

[14] 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.

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

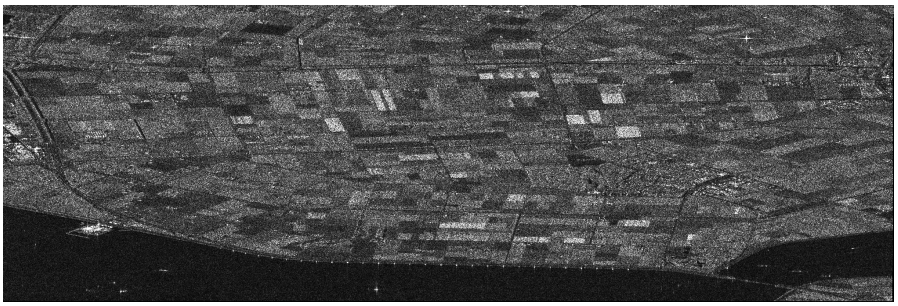
[16] PENGLANG SHUI and SHUCHEN FAN, "SAR Image Edge Detection Robust to Isolated Strong Scatterers Using Anisotropic Morphological Directional Ratio Test," *IEEE Access*, vol. 6, pp. 37 272–37 285, 2018.

[17] 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.

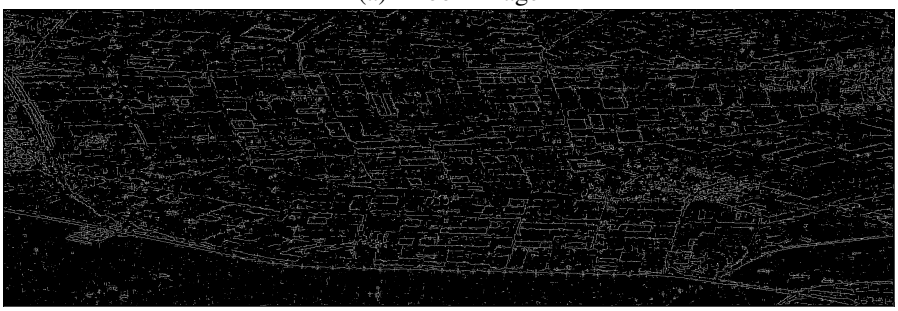
[18] Wei Shen, Xinggang Wang, Yan Wang, Xiang Bai and Zhijiang Zhang, "DeepContour: A deep convolutional feature learned by positive-sharing loss for contour detection," in *2015 IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 3982–3991.

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

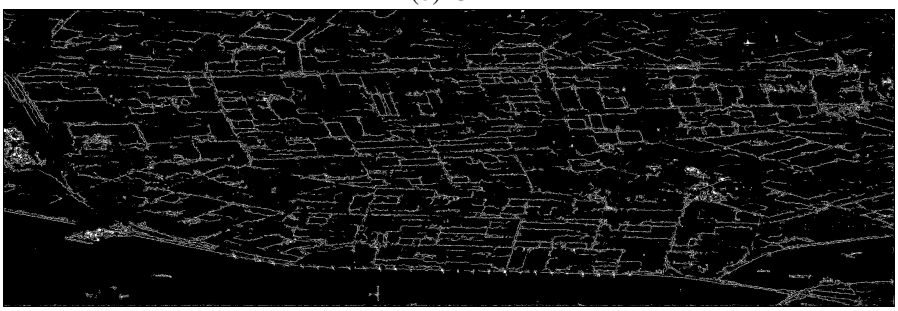
[20] ——, "Holistically-Nested Edge Detection," *International Journal of Computer Vision*, vol. 125, pp. 3–18, 2017.



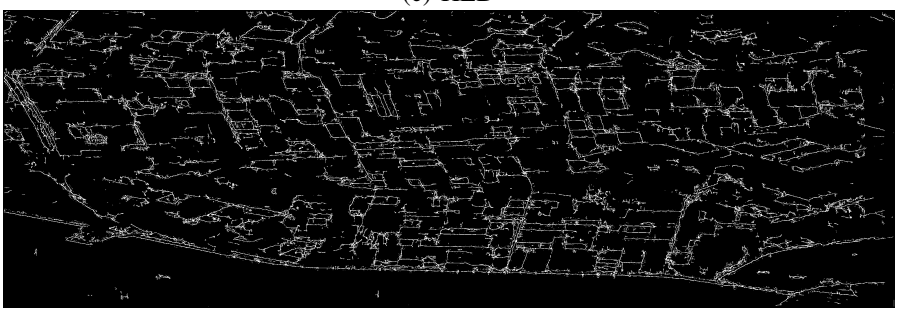
(a) 1-look image



(b) GR



(c) HED



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

Fig. 14: Comparison of different methods in a 1-look real SAR image (Leystad, Sentinel 1). The size of the image are 1024×3072 pixels.

[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.

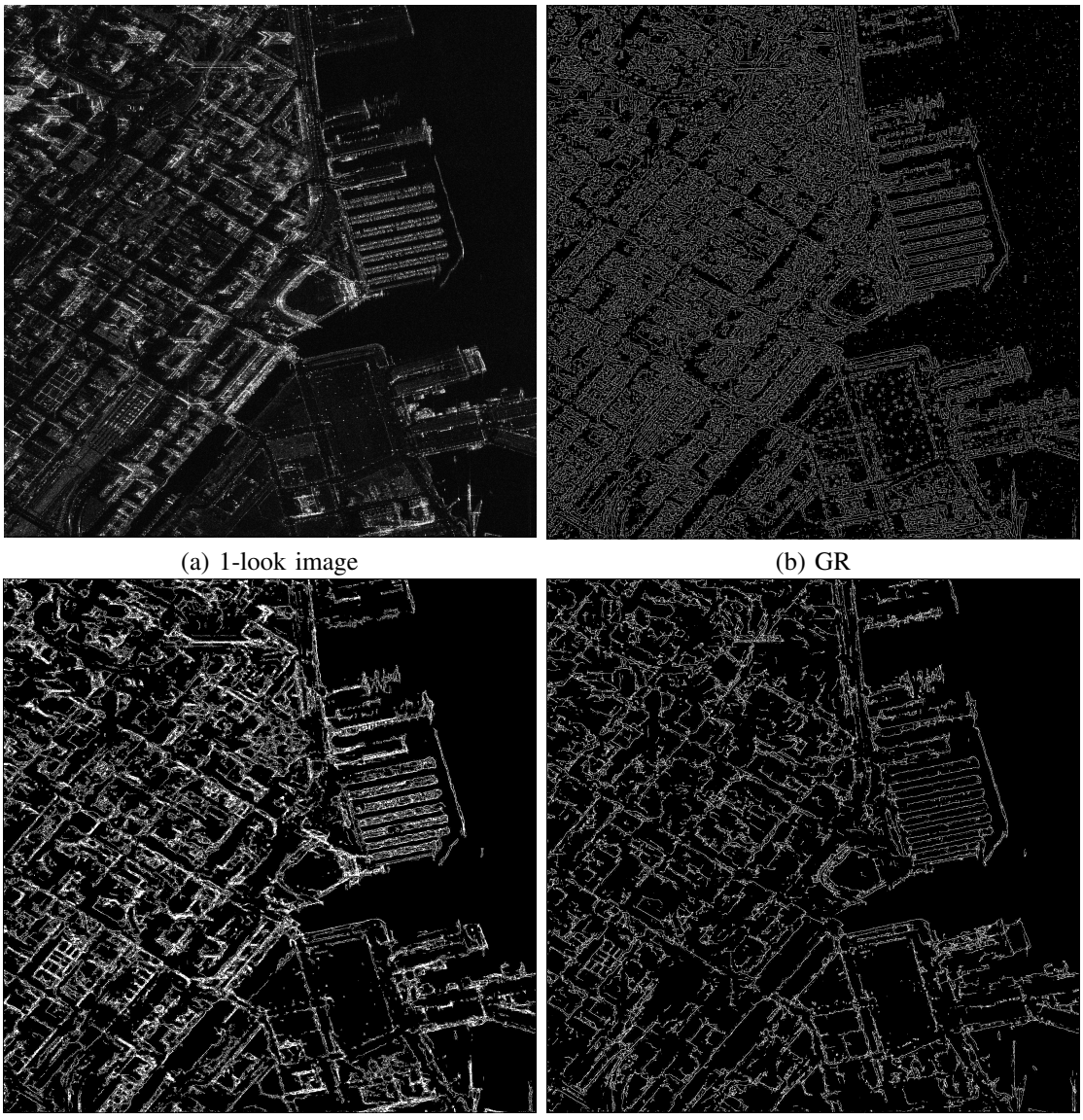
[24] Yun Liu, Ming-Ming Cheng, Xiaowei Hu, Jia-Wang Bien, Le Zhang, Xiang Bai, Jinhui Tang, "Richer Convolutional Features for Edge Detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 41, pp. 1939–1946, 2019.

[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.



(a) 1-look image (b) GR

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

Fig. 15: Comparison of different methods in a 1-look real SAR image (San Francisco, TerraSAR-X). The size of the image are 2048×2048 pixels.

[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] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun, "Deep Residual Learning for Image Recognition," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2016.

[36] 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.