Transferred Deep Learning for Sea Ice Change Detection From Synthetic Aperture Radar Images

Yunhao Gao, Feng Gao, Junyu Dong, and Shengke Wang

***Abstract*— High-quality sea ice monitoring is crucial to navigation safety and climate research in the polar regions. In this letter, a transferred multilevel fusion network (MLFN) is proposed for sea ice change detection from synthetic-aperture radar (SAR) images. Considering the fact that training data are limited in the task of sea ice change detection, a large data set was used to train the MLFN, and the deep knowledge can be transferred to sea ice analysis. In addition, cascade dense blocks are employed to optimize the convolutional layers. Multilayer feature fusion is introduced to exploit the complementary information among low-, mid-, and high-level feature representations. Therefore, more discriminative feature extraction can be achieved by the MLFN. Furthermore, the fine-tune strategy is utilized to optimize the network parameters. The experimental results on two real sea ice data sets demonstrated that the proposed method achieved better performance than other competitive methods.**

***Index Terms*— Change detection, deep learning, fine-tune, neural network, synthetic-aperture radar (SAR).**

I. Introduction

THERE have been growing interests in the Arctic and Antarctic observations for shipping and climate research. As a result, shipping activities begin to increase rapidly recently. However, many ships are not ice-strengthened, and then high-quality sea ice monitoring is very important for navigation safety. In order to provide valuable information for safe navigation, changed information of the ice coverage is of great significance.

In this letter, we focus on sea ice change detection which is defined as detecting changed areas between two images captured at different times over the same geographical region [1]. Due to the frequent cloud coverage and long periods of darkness in the polar regions, optical sensors can hardly generate continues observation. Instead, synthetic-aperture radar (SAR) sensors are used quite common to sea ice monitoring due to their strong ability to generate observations regardless of the weather conditions (cloud or rain). However, the intrinsic speckle noise makes the interpretation of SAR images challenging. Currently, due to the lack of suitable automatic sea ice change detection methods, SAR images are manually interpreted by human experts working in national ice agencies [2]. The manual interpretations are time-consuming and are highly dependent on the skills of the experts. Therefore, efficient automatic sea ice change detection techniques need to be developed.

There have been constant efforts in automating the sea ice change detection task. Clustering methods have been widely used for change detection. Celik [3] proposed a change detection method by partition pixel into two groups by k-means clustering. Gong *et al.* [4] presented a reformulated fuzzy local information c-means clustering method which could reduce the influence of speckle noises by adding a fuzzy factor to the objective function. Besides clustering methods, threshold methods are also very popular with the change detection task. Bruzzone and Prieto [5] modeled the difference image (DI) by Gaussian distribution where the expectation maximization (EM) algorithm was used to determine an optimal threshold. Xiong *et al.* [6] proposed a threshold method for SAR change detection which combined the features of two change detection measures by the Markov random field model. Clustering and threshold methods are simple to implement and easy to understand. However, these methods are prone to be affected by the speckle noise, and the contextual information around each pixel is not fully exploited.

Sea ice image interpretation is a complex recognition task that requires machine learning algorithm with strong capabilities to learn the nonlinear relationship between multitemporal images. Recently, with the great breakthroughs by deep learning methods in visual recognition tasks [7]–[9], remote sensing image interpretations also benefit from deep models. Hou *et al.* [10] proposed a change detection method based on multilevel convolutional neural networks (CNN) and low-rank decomposition. Wang *et al.* [11] presented a general end-to-end 2-D CNN framework for hyperspectral image change detection. For remote sensing images, CNN can effectively model local image structures at different scales, which is achieved by local connectivity between the adjacent layers and weight sharing within one layer. However, a large volume of data is desired for a robust CNN model training, which becomes intractable when only limited samples are available. For the sea ice change detection task, the observation data of ice shelf breakup are scarce. One way to solve the problem is transferring a model pretrained on related tasks from a large data set to sea ice change detection since it is widely acknowledged that features in lower layers of CNN are less specific to the classification task.

In this letter, a multilevel fusion network (MLFN) was established for sea ice change detection. We build a large data

Manuscript received November 1, 2018; revised February 20, 2019; accepted March 15, 2019. Date of publication April 9, 2019; date of current version September 25, 2019. This work was supported in part by the National Natural Science Foundation of China under Grant 41606198, Grant 41576011, and Grant U1706218. *(Corresponding author: Feng Gao.)*

The authors are with the Qingdao Key Laboratory of Mixed Reality and Virtual Ocean, School of Information Science and Engineering, Ocean University of China, Qingdao 266100, China (e-mail: gaofeng@ouc.edu.cn).

Color versions of one or more of the figures in this letter are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LGRS.2019.2906279

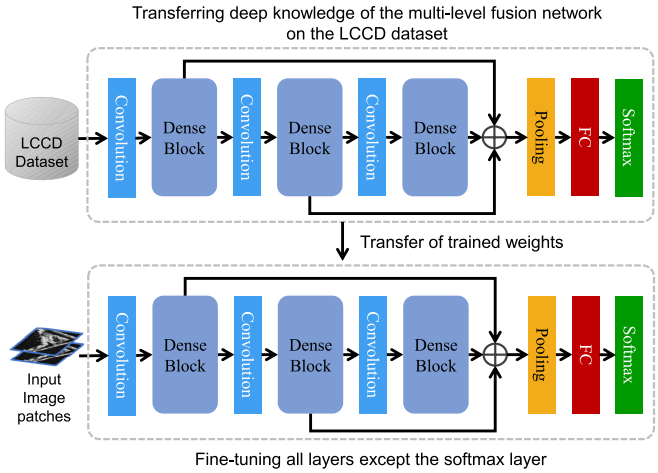

Fig. 1. Framework of the transferred MLFN. The LCCD data set is used to train the MLFN, and deep knowledge is transferred to sea ice change detection task. Dense blocks and multilayer feature fusion are employed to extract more discriminative features for the MLFN.

set to transfer deep knowledge from the data set to the limited training data in sea ice change detection. As far as we know, it is the first time that transferred deep learning is applied for sea ice change detection. In the proposed MLFN, cascade dense blocks are employed to optimize the convolutional layers. Multilayer feature fusion is introduced to exploit the complementary information among low-, mid-, and high-level feature representations. Therefore, more discriminative feature extraction can be achieved by the proposed MLFN. Experimental results on real SAR data sets demonstrated that the proposed MLFN has a better performance than several closely related methods.

II. Methodology

A. Problem Statements and Overview of the Proposed Method

Given two coregistered SAR images I_1 and I_2, which are captured over the same polar region at different times t_1 and t_2, respectively. Our goal is to highlight the changed regions occurring between t_1 and t_2. Finally, a change map will be generated that represents the change information.

The general framework of the proposed method is illustrated in Fig. 1. A great number of image patch pairs are randomly extracted from multitemporal images in the LCCD data set, and these image patch pairs are fed into the MLFN for training. After training, the trained weights are transferred to the input sea ice images. A preclassification process as mentioned in [13] is employed to select reliable training samples from the sea ice images. Then, image patches around each training sample are extracted, and these patches are used to fine-tune the MLFN. Features from all the layers of the MLFN are considered to be important, and all the layers are fine-tuned. After fine-tuning, all image patches from the input sea ice images are fed into the MLFN for classification, and then the final change map can be obtained.

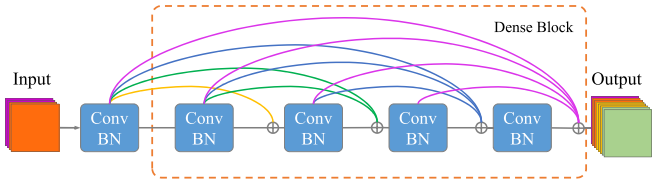

Fig. 2. Illustration of the dense block. ConvBN represents three operations: convolution, batch normalization and activation layer. The dense block has five layers, and each layer takes all preceding feature-maps as input.

B. Multilevel Fusion Networks

Densely connected convolution networks (DenseNets) [14] are known to be extreme cases of residual networks [15], where each convolution layer is connected by multiple short-cut connections. DenseNets strengthen feature propagation and encourage feature reuse.

Inspired by DenseNets, this letter proposes MLFN for sea ice change detection. The proposed MLFN uses three dense blocks for feature extraction. The three dense blocks can extract low-level, mid-level, and high-level features, respectively. The first dense block can extract minor details of the image such as lines or dots. The second block can extract mid-level features, which correspond to a combined output of low-level features. Finally, the third block can capture the structured information and semantic context of the input images.

As illustrated in Fig. 2, the dense block has four layers. The lth layer has l inputs, consisting of the feature maps of all preceding convolutional layers. Its own feature maps are passed to all $4 - l$ subsequent layers. This introduces $4 \times (4 + 1)/2 = 10$ connections in the dense block.

Feature fusion can promote feature reuse and reduce the number of parameters to some extent. In vision recognition tasks, feature fusion has been proven to be successful in many visual recognition tasks. Zhao *et al.* [16] presented a person reidentification methods which fused extracted features from different body regions with a competitive strategy. Chaib *et al.* [17] fused features from fully connected layers for remote sensing scene recognition. In this letter, we introduced the feature fusion strategy to exploit the complementary information among three dense blocks. Three dense blocks are used to extract the low-level, middle-level, and high-level features of the input image. The outputs of the three blocks are expressed as F_L, F_M, and F_H, respectively. Outputs of the three blocks are merged as follows:

$$\mathbf{F} = \text{pooling}(h(\mathrm{F}_L) + h(\mathrm{F}_M) + h(\mathrm{F}_H)) \tag{1}$$

where $\mathbf{F}$ represents the fused features, and pooling is a global average function. $h(\cdot)$ is a dimension matching function. Dimension matching is performed before feature fusion. 64 kernels with the size of 1×1 are employed to convolute F_L, F_M and F_H. With such convolution, the number of features maps of the three dense blocks all become 64. Therefore, the feature fusion can be achieved by elementwise summation.

The fused features are fed into a global pooling layer and then processed by one fully connected layer. Finally, the fused features are transformed into one high-dimensional vector.

The vector is used as the input of a softmax layer to calculate the possibility of being changed or unchanged. The possibility of being changed or unchanged can be represented as p_c and p_u, respectively. And if $p_c > p_u$, the input image patch pairs belong to the changed class, otherwise, they belong to the unchanged class.

C. Network Fine-Tuning

In this letter, a large data set named as *Land Cover Change Detection* (LCCD) data set was constructed. There are 10 multitemporal SAR image pairs (256×256 pixels) together with annotated ground truth change maps. There are two types of pixels in the ground truth change maps. Specifically, the changed class (the label is denoted as 1) and the unchanged class (the label is denoted as 0). The images in LCCD data set were captured by Radarsat, ENVISAT, and ERS satellites.

The goal of our work is to transfer deep knowledge from the LCCD data set to the limited training data in sea ice change detection. We randomly selected 30 000 image patch pairs for the changed class and 30 000 image patch pairs for the unchanged class from LCCD data set. These selected samples are fed into the MLFN to train a model. Next, the model is adapted to the input multitemporal sea ice images. There are two possible approaches for performing fine-tuning in the pretrained MLFN. The first approach is to fine-tune all the layers. The second is to keep some of the low-level layers fixed and then fine-tune the high-level layers. In the first approach, the softmax layer is removed from the pretrained MLFN. In the other approach, the low-level layers are frozen to keep the features already learned and the high-level layers are adjusted for input multitemporal SAR images. In this letter, we fine-tuned all layers, and therefore, features from all layers are assumed to be crucial for sea ice change detection.

D. Training Samples Selection and Change Map Generation

The input multitemporal SAR images are first handled by the log-ratio operator to generate the DI. The DI is computed as $I_D = |\log(I_2/I_1)|$. It is assumed that the multiplicative noise can be transformed to additive noise by such operation.

After DI generation, a hierarchical fuzzy c-means algorithm [13] is performed on I_D to classify pixels into three groups: the changed class G_c, the unchanged group G_u, and the intermediate group G_i. Pixels in G_c have high probability of being changed, while pixels in G_u have high probability of being unchanged. Pixels in G_i will be further classified by the MLFN.

Image patches pairs centered at pixels from G_c and G_u are extracted from the input multitemporal SAR images. Let R_p^1 represent the image patch centered at position p in image I_1, and R_p^2 represent the corresponding image patch in image I_2. The size of each patch is $r \times r$. We randomly select 10 000 samples from G_c and G_u. The selected samples are used to fine-tune the pretrained MLFN. Finally, all the pixels from the G_i are further classified by the MLFN. In the classification results, the changed class is labeled as 1, and the unchanged class is labeled as 0. The classification results together with G_c (labeled as 1) and G_u (labeled as 0) form the final change map.

TABLE I
IMPLEMENTATION DETAILS OF THE PROPOSED MLFN

Layers	Kernel size	Stride	Number Output	Output Size
Input				13×13
Convolution	1×1	1	16	13×13×16
Dense Block	3×3	1	16	13×13×16
Convolution	1×1	2	32	7×7× 32
Dense Block	3×3	1	32	7× 7× 32
Convolution	1× 1	2	64	4× 4× 64
Dense Block	3× 3	1	64	4× 4× 64
Global Pooling	-	-	-	1× 1× 64
Full Connection	-	-	2	-

III. EXPERIMENTAL RESULTS AND ANALYSIS

In this section, to evaluate the performance of the proposed MLFN, experiments were conducted on two real sea ice data sets. The implementation details of the proposed MLFN are listed in Table I. We first briefly describe the data set together with the evaluation criteria. Then, the important parameters that influence the change detection performance are discussed. Finally, comparisons with several excellent change detection methods are provided to demonstrate the effectiveness of the proposed MLFN.

A. Experiment Setup

In order to demonstrate the effectiveness of the proposed method, we apply the proposed method on two sea ice data sets. Both data sets were acquired at Sulzberger Ice shelf by the Advanced SAR (ASAR) on the ENVISAT satellite and were provided by the European Space Agency. The images were acquired on March 11 and 16, 2011, respectively. The data sets show the calving of large icebergs from the Sulzberger Ice Shelf as a result of the huge waves generated by the tsunami in Japan. The original size of the two SAR images is 2263×2264 pixels. Both images are too huge to display the detailed information. Therefore, we select two typical regions, denoted as Sulzberger I and Sulzberger II, respectively.

In this letter, the performance of the proposed MLFN will be measured by the number of false positives (FP), false negatives (FN), overall error (OE), and percentage of correct classification (PCC) and kappa coefficient (KC).

B. Analysis of Parameters

In order to obtain the contextual information, an image patch is extracted around each sample pixel. We choose a size of $r \times r$ for each patch. The parameter r is a crucial parameter that may affect the performance of sea ice change detection. Therefore, we set r as 5, 7, 9, 11, 13, and 15, respectively. The PCC values of different r are shown in Fig. 3.

As can be observed that, on both data sets, the PCC values achieve best when $r = 9$ or $r = 11$. There is a tendency

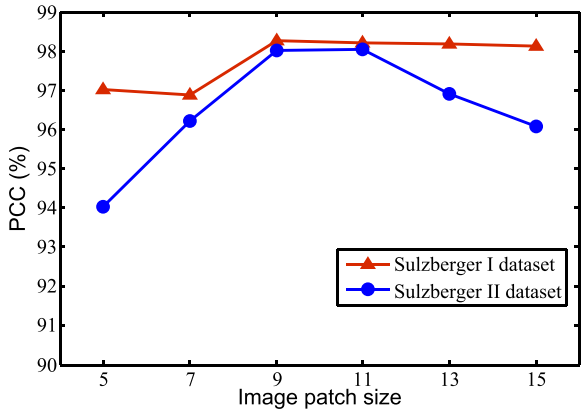

Fig. 3. Relationship between different image patch size and the PCC.

TABLE II
CHANGE DETECTION RESULTS OF MLFN WITH TRANSFER LEARNING AND MLFN WITHOUT TRANSFER LEARNING

Dataset	Methods	PCC (%)
Sulzberger I	MLFN without transfer learning	97.56
	Fine-tune parameters of the high-level layers	97.83
	Fine-tune parameters of all layers	98.28
Sulzberger II	MLFN without transfer learning	96.88
	Fine-tune parameters of the high-level layers	97.76
	Fine-tune parameters of all layers	98.05

of decline in PCC values when $r > 11$. The reason lies in the fact that larger image patch may not be representative of the central pixel, and the change detection is prone to be affected by the speckle noise. When $r \leq 7$, the PCC values are relatively low since more contextual information needs to be taken into account. In this letter, r is set to 9 in our following implementations.

Next, we discuss the effectiveness of transfer learning. As shown in Table II, the MLFN combining transfer learning obtains the better PCC. Specifically, there are 0.72% and 1.17% improvements in PCC by fine-tuning parameters of all layers on two data sets, respectively. Furthermore, we can observe that fine-tuning parameters of all layers have a better performance than fine-tuning parameters of the high-level layers. The reason for this phenomenon lies in the fact that the features from all layers are important for our sea ice change detection task. Therefore, in our implementations, we fine-tune parameters of all layers.

C. Experimental Results and Discussions

In this section, we compare our method with three closely related methods, including PCAKM [3], NBRELM [12], and GaborPCANet [13]. It is worth mentioning that NBRELM, GaborTLC, and GaborPCANet are implemented by the authors' open source code with the default parameters. The final change detection results of different methods are shown in figure form and the corresponding evaluation criteria are listed in tabular form.

TABLE III
CHANGE DETECTION RESULTS OF DIFFERENT METHODS ON THE SULZBERGER I AND II DATA SETS

Methods	Results on the Sulzberger I dataset				
	FP	FN	OE	PCC (%)	KC (%)
PCAKM [3]	711	479	1190	98.18	94.23
NBRELM [12]	817	673	1490	97.73	92.76
GaborPCANet [13]	473	739	1212	98.15	94.04
Proposed MLFN	950	179	1129	98.28	94.61
Methods	Results on the Sulzberger II dataset				
	FP	FN	OE	PCC (%)	KC (%)
PCAKM [3]	3073	362	3435	94.76	86.74
NBRELM [12]	539	2782	3321	94.93	85.82
GaborPCANet [13]	2475	500	2975	95.46	88.35
Proposed MLFN	616	664	1280	98.05	94.78

Fig. 4 shows the final change maps of different methods on two data sets, and Table III lists the corresponding evaluation criteria. On the Sulzberger I data set, it can be observed that many changed regions are missed by GaborPCANet, and therefore, the FN value of GaborPCANet is high and the overall performance is affected. PCAKM and NBRELM generate higher OE values than the proposed MLFN. The proposed MLFN exploits rich feature representations from transferred knowledge from LCCD data set, and therefore achieves the best performance. The KC value of the proposed MLFN is improved by 0.38%, 1.85%, and 0.57% over PCAKM, NBRELM, and GaborPCANet, respectively. Therefore, we can conclude that the proposed MLFN outperforms other methods on the Sulzberger I data set.

On the Sulzberger II data set, we can notice that many changed pixels are missed by NBRELM. Therefore, NBRELM generates very high FN value. The FP values of PCAKM and GaborPCANet are extremely high, and it indicates that many unchanged pixels are falsely detected as changed ones. The proposed MLFN can draw a balance between the FP and FN value and achieves the best PCC and KC values. It should be noted that the proposed MLFN has powerful feature representations by introducing residual learning and multilayer feature fusion. The PCC value of the proposed MLFN is improved by 3.29%, 3.12%, and 2.59% over PCAKM, NBRELM, and GaborPCANet, respectively. As a result, we can draw the conclusion that the proposed MLFN is superior to the other methods on the Sulzberger II data set. Visual and quantitative analysis on both data sets demonstrated that the proposed MLFN is effective in sea ice change detection from SAR images.

The computational complexity of the MLFN and other closely related methods is reported in Table IV. All experiments were implemented on the Intel Xeon E5-2620, NVIDIA GTX 1080 platform. The computational cost of the proposed MLFN is higher than PCAKM and NBRELM, due to the fact that deep learning-based MLFN needs to optimize

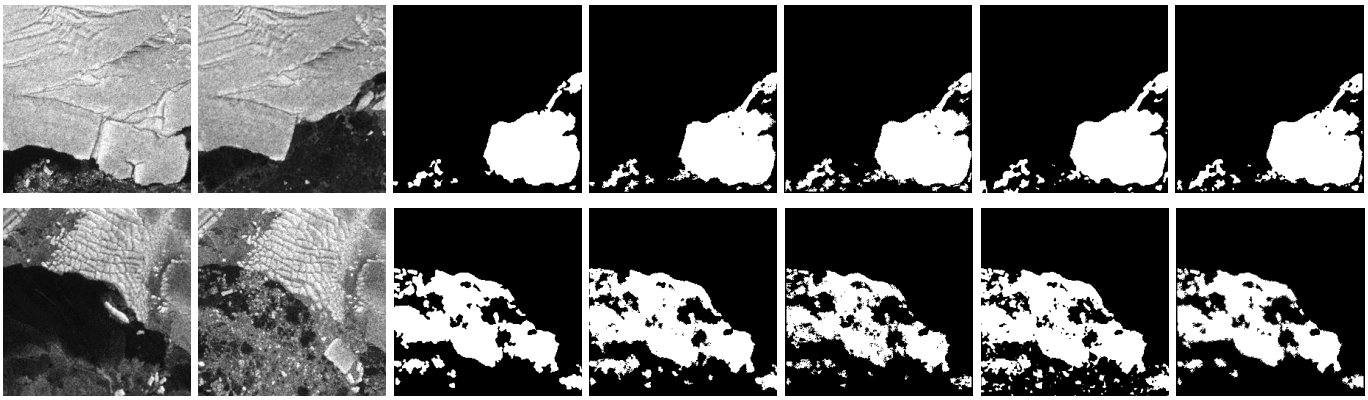

Fig. 4. Visualized results of different change detection methods on the Sulzberger I data set (first row) and Sulzberger II data set (second row). (a) Images captured on March 11, 2011. (b) Images captured on March 16, 2011. (c) Ground truth change maps. (d) Results by PCAKM [3]. (e) Results by NBRELM [12]. (f) Results by GaborPCANet [13]. (g) Results by the proposed MLFN.

TABLE IV
COMPUTING TIME OF DIFFERENT METHODS ON BOTH DATA SETS (IN SECONDS)

Dataset	PCAKM	NBRELM	GaborPCANet	MLFN
Sulzberger I dataset	2.15	22.47	435.78	111.57
Sulzberger II dataset	2.36	23.73	477.53	113.88

much more parameters. However, if more powerful graphics card is provided, the speed of the proposed MLFN can be further improved. Moreover, compared with GaborPCANet, the proposed MLFN is superior in computational time, which means that the speed of the proposed MLFN is computational efficient than typical deep learning-based model.

IV. CONCLUSION

In this letter, we proposed a transferred MLFN for sea ice change detection from SAR images. In the proposed method, a large data set (LCCD) is used for training. In the MLFN, cascade dense blocks and multilayer feature fusion are employed to improve the classification performance. Furthermore, the fine-tune strategy is utilized to optimize the network parameters and better change detection results are achieved with less computational cost. Experimental results on two real sea ice data sets demonstrated that the proposed MLFN outperformed several excellent change detection methods.

REFERENCES

[1] R. J. Radke, S. Andra, O. Al-Kofahi, and B. Roysam, "Image change detection algorithms: A systematic survey," *IEEE Trans. Image Process.*, vol. 14, no. 3, pp. 294–307, Mar. 2005.

[2] L. Wang, K. A. Scot, L. Xu, and D. A. Clausi, "Sea ice concentration estimation during melt from dual-pol SAR scenes using deep convolutional neural networks: A case study," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 8, pp. 4524–4533, Aug. 2016.

[3] T. Celik, "Unsupervised change detection in satellite images using principal component analysis and k-means clustering," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 4, pp. 772–776, Oct. 2009.

[4] M. Gong, Z. Zhou, and J. Ma, "Change detection in synthetic aperture radar images based on image fusion and fuzzy clustering," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 2141–2151, Apr. 2012.

[5] L. Bruzzone and D. F. Prieto, "Automatic analysis of the difference image for unsupervised change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 3, pp. 1171–1182, May 2000.

[6] B. Xiong, Q. Chen, Y. Jiang, and G. Kuang, "A threshold selection method using two SAR change detection measures based on the Markov random field model," *IEEE Geosci. Remote Sens. Lett.*, vol. 9, no. 2, pp. 287–291, Mar. 2011.

[7] Q. Wang, J. Wan, and Y. Yuan, "Deep metric learning for crowdedness regression," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 28, no. 10, pp. 2633–2643, Oct. 2018. doi: 10.1109/TCSVT.2017.2703920.

[8] Q. Wang, J. Gao, and Y. Yuan, "A joint convolutional neural networks and context transfer for street scenes labeling," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 5, pp. 1457–1470, May 2018.

[9] Q. Wang, J. Gao, and Y. Yuan, "Embedding structured contour and location prior in siamesed fully convolutional networks for road detection," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 1, pp. 230–241, Jan. 2018.

[10] B. Hou, Y. Wang, and Q. Liu, "Change detection based on deep features and low rank," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 12, pp. 2418–2422, Dec. 2017.

[11] Q. Wang, Z. Yuan, Q. Du, and X. Li, "GETNET: A general end-to-end 2-D CNN framework for hyperspectral image change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 1, pp. 3–13, Jan. 2019. doi: 10.1109/TGRS.2018.2849692.

[12] F. Gao, J. Dong, B. Li, Q. Xu, and C. Xie, "Change detection from synthetic aperture radar images based on neighborhood-based ratio and extreme learning machine," *Proc. SPIE*, vol. 10, no. 4, Dec. 2016, Art. no. 046019.

[13] F. Gao, J. Dong, B. Li, and Q. Xu, "Automatic change detection in synthetic aperture radar images based on PCANet," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 12, pp. 1792–1796, Dec. 2016.

[14] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 4700–4708.

[15] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.

[16] H. Zhao *et al.*, "Spindle Net: Person re-identification with human body region guided feature decomposition and fusion," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 1077–1085.

[17] S. Chaib, H. Liu, Y. Gu, and H. Yao, "Deep feature fusion for VHR remote sensing scene classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 8, pp. 4775–4784, Aug. 2017.