IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.Doi Number

A Novel Data Augmentation Method for Sea Ice Scene Classification of Arctic Aerial Images

Yiming Yan[1], Chunming Zhang[1], and Nan Su[1,*]

[1]Department of Information and Communication Engineering, Harbin Engineering University, Harbin 150001, China;

Corresponding author: Nan Su (e-mail: sunan08@hrbeu.edu.cn).

ABSTRACT In this paper, a novel data augmentation method was proposed for supervised sea ice scene classification with Arctic aerial images. Ice-type classification of sea ice scenes in a region is useful for instant navigation. However, some types of sea ice scenes are difficult to collect. The small number of available samples usually limit the performance in classification. Inspiring by transfer learning method, the training samples are augmented with simulated sea ice images. Considering the distribution characteristic of sea ice scene, simulation samples are synthesized by algebraic operations on the respective regions of interest from two true samples. One of the two true samples can be an additional image, which is more easily collected than others, such as full ice and snow. Simultaneously, it introduces new information. Generalization and error-correction capability of deep neural networks for training samples makes the proposed method feasible. Experiments on true sample sets, simulation sample sets, and mixed sample sets were implemented. Finally, the effectiveness of our data augmentation method was demonstrated, which improved the accuracy of sea ice classification.

INDEX TERMS Data augmentation; scene classification; sea ice; transfer learning

I. INTRODUCTION

With the accelerated melting of the Arctic sea ice, more and more ships try an Arctic waterway as a shorter commercial route in special time [1]. Sea ice is a critical factor of route planning for Arctic navigation. Even though choosing to marine in the warmer months with powerful ice-breakers, the complex and changeable sea-ice environment still poses a great threat. The distribution and type of sea ice is quite important information, which is conductive to short-term route planning of ships. Different types of sea ice scene can affect the ships to varying degrees. Some types can trap or damage ships while some types of sea ice allow ships to pass safely. Therefore, to obtain a reliable indication and avoid threats of sea ice, scene classification and analysis of sea ice is essential. Air-crafts are always employed to collect images of the sea ice scenes of Arctic Ocean for scouting, rescue or navigation in a short period of time. Since high cost Arctic flight experiments, images of some special types of sea ice scenes are not easily collected. Limited training data is a specific challenge for supervised learning based image classification method.

Insufficient data will cause a decrease in scene classification performance when using supervised learning methods. In order to obtain an ideal classification result, the strategy of transfer learning is commonly used by many researchers when training with deep neural networks. Transfer learning is a process of taking a pre-trained deep learning network and fine-tuning it to learn a new task [2], which can speed up the training progress and guarantee the performance of deep learning models with moderate amount of data. Although transfer learning can reduce the amount of data required to some extent, a certain number of samples are still needed. Therefore, data augmentation is adopted to achieve expansion of limited data. We proposed a novel data augmentation method to generate diverse sea ice scenes for limited raw data to guarantee an appropriate amount of input data for deep neural networks. For scene classification tasks, the commonly-used image augmentation methods are flipping [3], rotation [4], cropping [5], color jittering [6], noise [7] and so on. These data augmentation methods improve the diversity of data and the generalization ability of models. However, these augmented data are obtained by changing the direction, position, proportion and pixel color of objects in original data, they do not generate additional data. Generative Adversarial Networks (GANs) [8] have recently made great progress in generation of new images for training. Nevertheless, GANs also have their own limitations: there is no control over the generated data and the training is unstable [9]. Therefore, the improved methods of GANs have been proposed and widely used in diverse kinds of data generation

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2019.2929861, IEEE Access

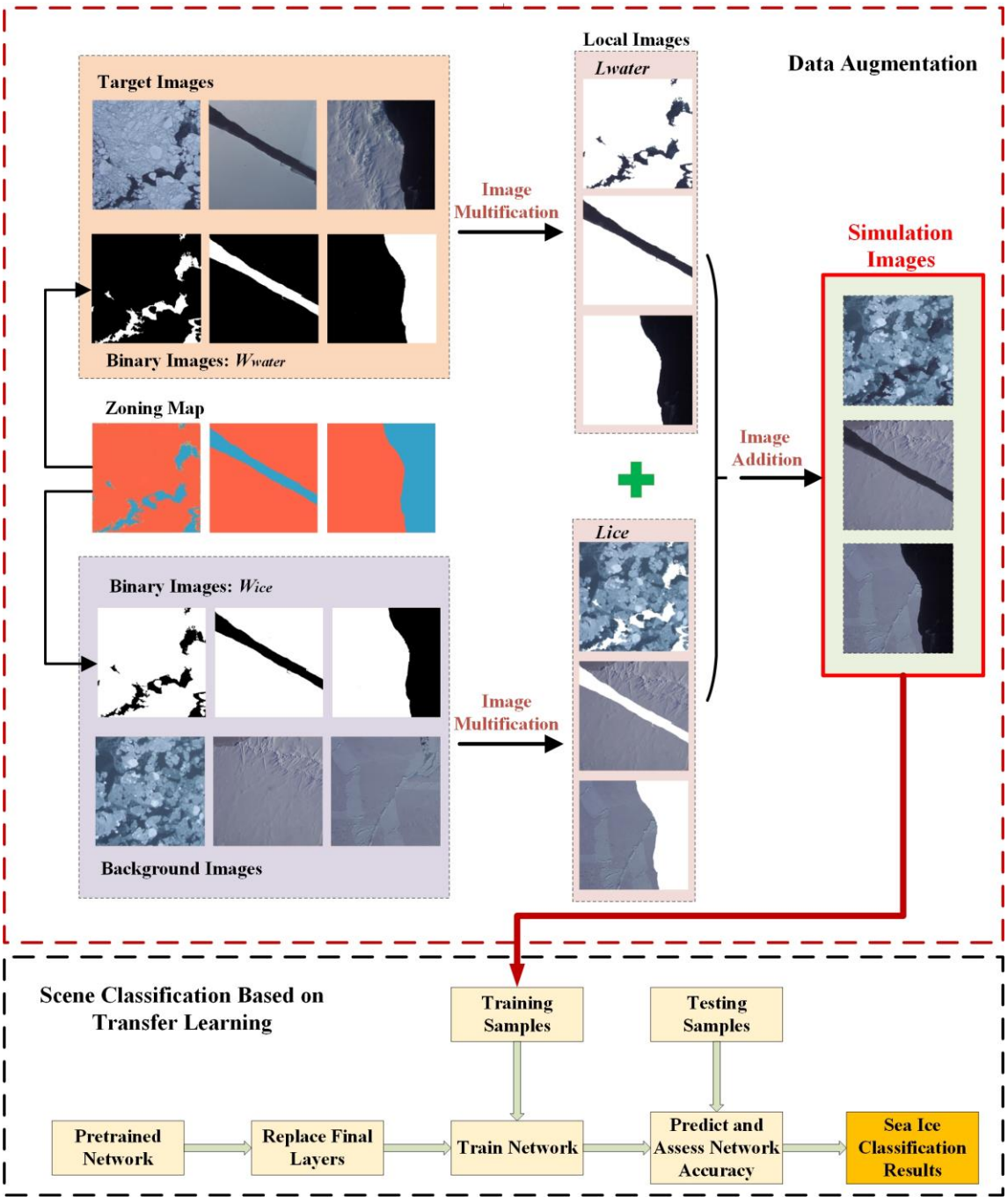

FIGURE 1. Flow chart of our proposed method.

tasks, such as conditional GANs (CGANs) [10] and Wasserstein GANs (WGANs) [11]. Nevertheless, these networks mentioned above usually require a large amount of data to train, which is not suitable for limited number of Arctic aerial sea ice images. To enlarge training data set and generate diverse data samples with new information, a novel data augmentation method for a limited data samples is proposed.

As we all know, when a baby comes into recognize new objects, 'simulation knowledge' is often enough for learning. For instance, we teach a baby what is a car with a cartoon picture for many times, and then he can recognize cars in the real world. Even when the 'simulation knowledge' is not very similar to the actual things, humans can still make an accurate judgment. Inspired by that, it can be inferred that the ‘simulation samples’ can be useful for supervised learning. Moreover, it should be especially helpful for deep neural networks due to the good generalization and fault tolerance [12]. Therefore, we proposed a new data augmentation method for aerial sea ice images by using simulation samples.

The main contributions of our method are as follows:

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2019.2929861, IEEE Access

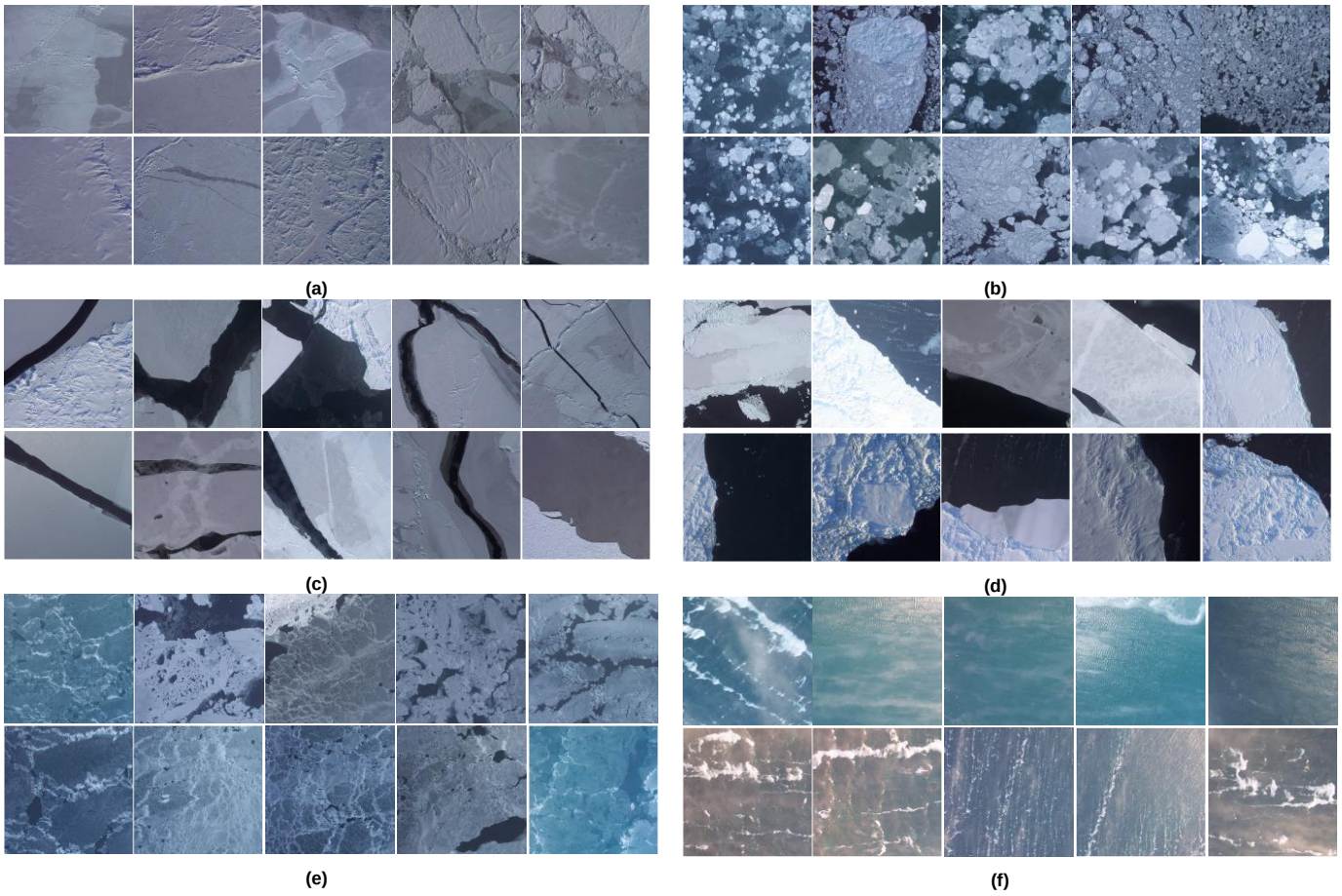

FIGURE 2. Typical aerial sea ice images picked from data set: (a) Gray Ice (GI); (b) Dense Grained Ice (DGI); (c) Lead-Ice (LI); (d)Transition Floes-Lead (TFL); (e) Thin Ice (TI); (f) Open Water (OW).

First, a novel data augmentation strategy was proposed, which provides a new solution to scarcity of aerial sea ice images. Second, a simulation method based on a linear superposition strategy was employed to generate diverse sea ice scenes. It makes the simulation sample contain additional information outside original training set. This paper is organized as follows. In Section II, details of the proposed data augmentation method are introduced for classification of aerial sea ice images. Section III presents experimental results of sea ice classification. Followed by a conclusion in Section IV.

II. METHODOLOGY

A. OVERALL FRAMEWORK

The flow chart of our proposed method is displayed in Figure 1. The task of sea ice scene classification from Arctic aerial images is completed by two steps: data augmentation and scene classification based on transfer learning. In the proposed data augmentation method, simulation samples are synthesized by image algebraic operation of pairing samples: target image and background image. The scene label of target image is the same with that of the simulation image and the scene label of background image is usually different from that of target image, i.e., the label of simulation image. We usually pick target image from special types of sea ice scenes in training data set and pick background image from scenes of common ice-type out of the training data set. Then, make zoning maps based on the distribution of sea ice and sea water for selected target images. As shown in Figure 1, the sea ice region is painted in orange and the sea water region is painted in blue. In this process, the transition zone between sea ice and sea water does not need to be divided in details, we just divide the transition zone into sea ice or sea water roughly. Next, according to the drawn color partition map, the sea area is set to white, the sea ice area is set to black, the binary image of the target image is obtained (W_{water}), and then the color setting method is reversed to obtain the binary image of the background image (W_{ice}). Finally, by means of image algebraic operation, the sea water region is extracted into local image (L_{water}) from target image and sea ice region is extracted into local image (L_{ice}) from background images in the same way. These two local images are combined together to form a simulation sample.

Experiments are performed in neural networks of different depths and structures. Modifications are made in pre-trained

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2019.2929861, IEEE Access

neural networks to implement the ice-type scene classification by fine-tuning the models.

The focus of this paper is the sea ice scene classification of Arctic aerial images, with the help of simulation samples. To clearly describe our method, the ice-type categorization and concrete implementation method of simulation samples are introduced in Sections II-B and II-C. Section II-D presents ways to train deep neural networks under the case for which the number of aerial sea ice samples available is limited or sufficient.

B. SEA ICE CLASSIFICATION SPECIFICATION

Sea ice images have certain particularity: there are different scene classification results under different classification criteria [13], [14]. Just considering instant navigation route for ships, the sea ice scenes in data set are divided into six main kinds of types in terms of the extent of threat to navigation:

- **Gray ice (GI)**: GI is named according to the color of sea ice. GI is a kind of scene that full of ice and snow, and its ice is generally thick, which is very dangerous for sailing.
- **Dense grained ice (DGI):** DGI is distributed with dense crushed ice. Although it is not as strong as gray ice, it is still composed of dense ice, so it is also relatively unsuitable for boating even if ships are under the leadership of icebreakers. Passage on it may cause damage to the hull.
- **Lead-ice (LI):** LI is named according to sea ice lead [15], [16], which forms when ice floes diverge or shear as they move parallel to each other. The width of lead varies from a couple of meters to over a kilometer. Even when it freezes, lead tends to contain thinner and weaker ice, which allows ships or icebreakers to traverse the ice more easily. Large leads in ice bodies are of great significance for the planning of safe shipping routes or for maritime search and rescue. In addition, even a small fraction of snow-free thin ice and open water within the sea-ice pack can significantly modify the total energy transfer between the ocean and atmosphere [17], [18]. An image traversed by a lead is classified as LI.
- **Transition floes-lead (TFL):** Transition of lead and ice floe is called TFL.
- **Thin Ice (TI):** TI is a very thin type of ice. When a ship passes through such ice, its obstacles to navigation are small, and the damage to the ship is minor. This kind of ice is the most suitable for navigation except open water.
- **Open water (OW):** OW is an ice-free zone where ships are safe enough to sail on it.

In this paper, the number of GI, TI and OW samples is limited while the samples of LI, TFL, and DGI are deficient. Therefore, our data augmentation method aims to synthesize augmented data for LI, TFL, and DGI, i.e., the LI, TFL, and DGI scenes are regarded as target images. Figure 2 shows examples of these typical aerial sea ice scenes picked from data set.

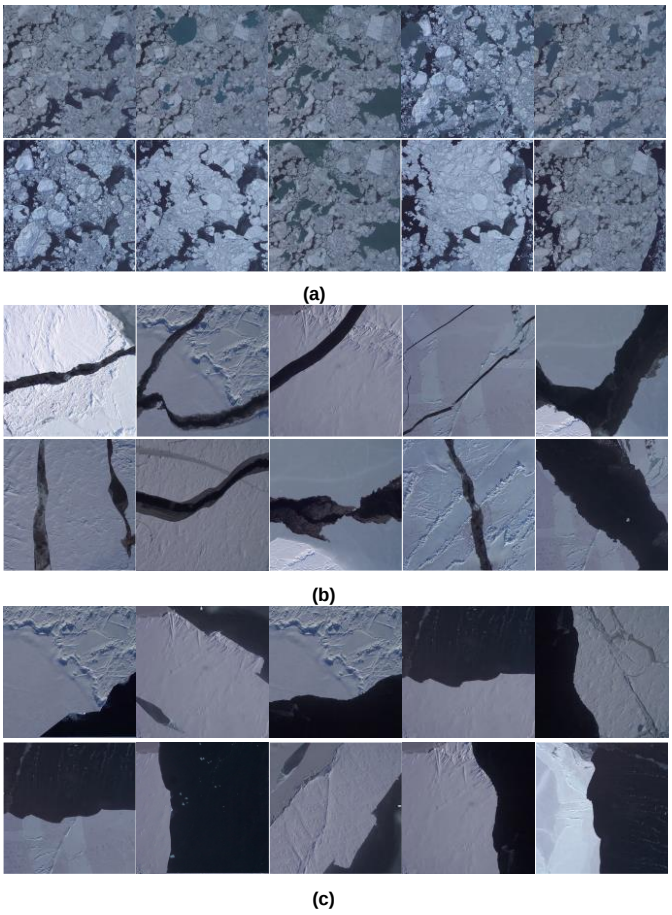

FIGURE 3. Sea ice simulation samples: (a) Dense Grained Ice (DGI); (b) Lead-Ice (LI); (c) Transition Floes-Lead (TFL).

C. DATA AUGMENTATION FOR SEA ICE IMAGES

The step of data augmentation aims to generate additional simulation samples using limited true samples for special type of sea ice scenes. As shown in Figure 1, three kinds of target images taken from LI, TFL, and DGI data sets are regarded as target images (I_{tar}) and their corresponding zoning maps are listed just above them. GI images picked outside training set are seen as background images (I_{bg}) for LI and TFI. It is worth pointing out that the background images for DGI are taken from the training set, due to the difference of ice shape between DGI images and GI images.

For a sea ice simulation sample, if the location of sea water is certain, the distribution of sea ice is also determined. In order to ensure the authenticity of the simulation samples, the sea water region of target image is extracted and overlaid into the background image, replacing the original pixels at the corresponding position in the background image. That means, features of sea water such as color, shape and textures are given more attention than sea ice region for target images. And new color and texture information are introduced from the sea ice region of background images.

To extract the region of sea water from the target images and sea ice from background images, the strategy we use for sea water and sea ice extraction is image algebraic operation,

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2019.2929861, IEEE Access

which has a wide range of applications in image processing [19], [20]. It is a process of obtaining an output image by performing addition, subtraction, multiplication or division point-to-point between pairs of images. Wherein, the image multiplication can be used to extract the region of interest (ROI). Specifically, when a RGB image is multiplied by a corresponding binary image where the pixel value of ROI is 1, the ROI in this RGB image is retained and rest region is black.

In this paper, the ROI of target image is sea water region, its corresponding binary image can be easily obtained through zoning maps. In such a binary image, white corresponds to sea water region and black corresponds to sea ice region. We recorded this kind of binary image as W_{water}. Similarly, the ROI of target image is sea ice region. We can also get its corresponding binary image where white corresponds to sea ice region and black corresponds to sea water region. And we called it W_{ice}.

When the W_{water} is multiplied with the target image, the output image retains the sea water region of target image, and this partial image, we called it L_{water}, defined as (1):

$$L_{\text{water}}(x,y) = W_{\text{water}}(x,y) * I_{tar}(x,y) \tag{1}$$

When W_{ice} is multiplied by the background image (GI image), the output image retains the sea ice region of GI image. This image with new information is called L_{ice}, as (2).

$$L_{\text{ice}}(x,y) = W_{\text{ice}}(x,y) * I_{bg}(x,y) \tag{2}$$

The simulation sample is the sum of the two local images, as (3).

$$S(x,y) = L_{\text{water}}(x,y) + L_{\text{ice}}(x,y) \tag{3}$$

Some simulation samples generated by our data augmentation method are shown as Figure 3.

D. TRANSFER LEARNING

This step aims to realize sea ice scene classification using limited data samples by the approach of transfer learning. As shown in (4), the input of pre-trained network model $M_{pre\text{-}trained}$ is a mixed training data set, which is composed of simulation samples S_i obtained from (3) and true samples T_i from original data set. The number of true sea ice samples for training is n_T and the number of simulation samples is n_S. The ice-type prediction result $P_{ice\text{-}type}$ is the corresponding output of $M_{pre\text{-}trained}$ after fine-tuning.

$$\left\{S_1,\cdots,S_{n_S},T_1,\cdots,T_{n_T}\right\} \rightarrow M_{pre-trained} \rightarrow P_{ice-type} \tag{4}$$

Transfer learning is performed by following these steps.

1. Choose a pre-trained network. The pre-trained network models involved in this paper are AlexNet [21], VGG-16, VGG-19, ResNet-50 [22], ResNet-101, Inception-v3 [23] and Inception-ResNet-v2 [24].

2. Replace the final three layers with new layers. Set learning rate to learn faster in the new layers. The learning rate of initial layers is 1×10^{-3}, the learning rate of three new layers is 1×10^{-4}.

3. Freeze initial layers. We freeze all layers except for the new layer that we add. The results of all experiments were performed with an iteration of 30 and a batch-size of 6.

When transferring different networks, all samples are resized to the image input size of various networks respectively. Table 1 lists the image input size of different networks.

TABLE I
THE IMAGE INPUT SIZE OF DIFFERENT NETWORKS

Network	Image input size
AlexNet	227-by-227
VGG-16	224-by-224
VGG-19	224-by-224
ResNet-50	224-by-224
ResNet-101	224-by-224
Inception-v3	299-by-299
Inception-ResNet-v2	299-by-299

E. Performance Metric

Overall Accuracy (OA) is an evaluation indicator that reflects the overall data set and is a determination of the entire samples, its calculation method is shown as (5),

$$OA = \frac{TP+TN}{TP+FN+FP+TN} \tag{5}$$

where:
TP (True Positives): predict the positive class as a positive class;
FN (False Positives): predict the positive class as a negative class;
FP (True Negatives): predict the negative class as a positive class;
TN (False Negatives): The negative class is predicted as a negative class.

III. EXPERIMENTS AND RESULTS

In this section, we designed a series of experiments to demonstrate the validity of simulation samples. Firstly, different numbers of true samples are used for training. Then, keep the number of training samples constant and replace the true samples with simulation samples until the true samples are completely replaced by simulation samples. Compare and analyze the feasibility of using simulation samples for training and the influence of the ratio of true samples and simulation samples on scene classification performance. In addition, comparison and compatibility of the proposed data augmentation method with basic data augmentation methods such as flip and rotation are also experimentally studied.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2019.2929861, IEEE Access

TABLE II
CLASS LABELS AND DISTRIBUTION OF SAMPLES

	Class	For training (each class)	For validation	For testing
1	Gray Ice (**GI**)		15	30
2	Dense Grained Ice (**DGI**)	**Group 1**: 10T, 10S; **Group 2**: 20T,10T+10S,20S;	15	30
3	Lead-Ice (**LI**)	**Group 3**: 30T,20T+10S, 10T+20S,30S;	15	30
4	Transition Floes-Lead (**TFL**)	**Group 4**: flip images of Group1 (Group2, Group3) left and right, up and down (retain original images);	15	30
5	Thin Ice (**TI**)	**Group 5**: rotate images of Group1 (Group2, Group3) 90 degrees, 180 degrees and 270 degrees (retain original images).	15	30
6	Open Water (**OW**)		15	30

TABLE III
OVERALL ACCURACY OF GROUP 1 TRAINING SAMPLES (IN %)

	AlexNet	VGG-16	VGG-19	ResNet-50	ResNet-101	Inception-v3	Inception-ResNet-v2
10 T	72.7778	**75.5556**	72.2222	68.3333	72.2222	68.8889	73.3333
10 S	69.4444	76.1111	**73.8889**	71.6667	65.0000	65.5556	62.7778

TABLE IV
OVERALL ACCURACY OF GROUP 2 TRAINING SAMPLES (IN %)

	AlexNet	VGG-16	VGG-19	ResNet-50	ResNet-101	Inception-v3	Inception-ResNet-v2
20 T	82.2222	**86.1111**	77.2222	77.7778	78.3333	72.2222	75.0000
10 T+10 S	78.8889	**81.1111**	78.8889	75.5556	77.7778	74.4444	71.1111
20 S	71.1111	**84.4444**	81.6667	74.4444	75.0000	76.6667	68.3333

TABLE V
OVERALL ACCURACY OF GROUP 3 TRAINING SAMPLES (IN%)

	AlexNet	VGG-16	VGG-19	ResNet-50	ResNet-101	Inception-v3	Inception-ResNet-v2
30 T	86.6667	**87.2222**	86.6667	86.6667	83.8889	76.1111	76.1111
20 T+10 S	82.2222	82.2222	**85.5556**	82.7778	83.3333	74.4444	75.5556
10 T+20 S	81.1111	**85.5556**	83.3333	81.1111	81.6667	78.8889	70.0000
30 S	76.6667	80.0000	78.8889	78.8889	**81.6667**	80.0000	71.1111

A. DATA SETS

This experimental data set contains photographs of sea ice in the Chukchi and Beaufort Seas of the Arctic Ocean, and of snow cover off the northern coast of Alaska, USA. Photographs were taken from a P3 aircraft during six days in March 2006 using two Kodak digital DC4800 cameras [25]. In order to observe the influence of the number of samples in training data set on the scene classification results, we design five different groups of training data set: Group 1, Group 2, Group 3, Group 4, and Group 5. The total number of training samples of each data set in Group 1, Group 2, and Group 3 is 10, 20, and 30 respectively.

In order to verify the validity of the augmented simulation samples, we designed data sets that are mixed of true samples and simulation samples and data sets that are consist of simulation samples only. Group 4 is obtained by flipping images of Group1 (Group2, Group3) left and right, up and down (retain original images), the quantity of samples in each data set is 3 times as much as before the flip. Group 5 is obtained by rotating images of Group1 (Group2, Group3) 90 degrees, 180 degrees and 270 degrees (retain original images). The quantity of samples in each data set is 4 times as much as before the rotation. The specific design is shown in Table II, it is worth pointing out that T means the six types of sea ice images in training set are composed of true samples, S denotes three types of sea ice images (DGI, LI, TFL) are made up of simulation samp-

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2019.2929861, IEEE Access

TABLE VI
OVERALL ACCURACY OF GROUP 4 TRAINING SAMPLES (IN %).

	AlexNet	VGG-16	VGG-19	ResNet-50	ResNet-101	Inception-v3	Inception-ResNet-v2
(10 T)$_{F}$	**77.7778**	76.1111	75.0000	68.8889	73.3333	73.8889	65.0000
(10 S)$_{F}$	75.5556	**81.6667**	81.1111	66.1111	71.1111	68.8889	63.8889
(20 T)$_{F}$	83.8889	78.8889	**85.5556**	78.8889	81.6667	81.6667	71.1111
(10T+10S)$_{F}$	78.3333	83.8889	**85.0000**	79.4444	76.1111	76.6667	74.4444
(20 S)$_{F}$	75.5556	81.1111	77.7778	**83.3333**	79.4444	78.3333	67.7778
(30 T)$_{F}$	83.3333	**88.8889**	**88.8889**	87.2222	87.2222	83.3333	75.0000
(20T+10S)$_{F}$	87.2222	86.1111	**90.0000**	86.6667	83.8889	85.0000	80.5556
(10T+20S)$_{F}$	81.1111	86.6667	**87.2222**	85.0000	82.7778	82.7778	77.2222
(30 S)$_{F}$	76.6667	82.2222	82.2222	81.1111	80.0000	**83.3333**	77.7778

TABLE VII
OVERALL ACCURACY OF GROUP 5 TRAINING SAMPLES (IN %).

	AlexNet	VGG-16	VGG-19	ResNet-50	ResNet-101	Inception-v3	Inception-ResNet-v2
(10 T)$_{R}$	77.7778	**78.3333**	75.5556	72.7778	74.4444	73.3333	67.7778
(10 S)$_{R}$	77.7778	78.3333	**80.0000**	70.5556	70.5556	70.5556	66.1111
(20 T)$_{R}$	82.7778	82.7778	**87.7778**	83.8889	80.5556	81.6667	79.4444
(10T+10S)$_{R}$	79.4444	**87.7778**	85.0000	80.0000	77.2222	80.0000	73.8889
(20 S)$_{R}$	79.4444	**86.6667**	82.7778	83.3333	77.7778	83.8889	73.3333
(30 T)$_{R}$	87.7778	**91.6667**	90.5556	87.7778	85.5556	86.1111	85.5556
(20T+10S)$_{R}$	83.3333	**89.4444**	**89.4444**	87.2222	85.0000	85.5556	81.1111
(10T+20S)$_{R}$	82.2224	**88.8889**	85.0000	87.2222	81.6667	82.2222	81.1111
(30 S)$_{R}$	77.7778	**85.5556**	84.4444	81.1111	81.6667	81.6667	73.8889

les while three other categories of sea ice images (GI, TI, OW) are still made of true samples.

B. VALIDITY OF SIMULATION SAMPLES

From Table III, considering the overall accuracy of different networks, it can be concluded that true samples perform better than simulation samples, but there is not much difference in some networks such as AlexNet, VGG-16, and VGG-19. Due to the small number of training samples, the best scene classification accuracy of Group 1 is still difficult to be satisfactory, only 76.1111%. Compared with Table III, it can be seen that Group 2 outperforms Group 1 under the same training network from the overall view, showing the power of the quantity of training data. From Table IV, the OA of three data sets in Group 2 almost reaches 70% among various networks and even above 80% with VGG-16. Data set with 20 true samples performs better than that with 20 simulation samples or mixed samples. It is not difficult to see that when the number of training samples is the same, true samples perform better than simulation samples or mixed samples. However, when the total number of samples is not constant, the addition of simulation samples can greatly improve the classification accuracy, comparing with the result of '10T' in Table III and '10T+10S' in Table IV. This can also be verified in Table III and Table IV. From Table V, it can also be seen that when using simulation samples to replace true samples in the case of keeping the total number unchanged, the rate of accuracy reduction in Group 3 is slower than that in Group 2. For AlexNet, when 20 true training samples are replaced with 20 simulation samples, the OA reduced by 11.1111%, while 30 true training samples are replaced to 20 simulation samples and 10 true samples, the OA reduced by 5.5556%. Meanwhile, the conclusion mentioned above that simulation sample matters to improve classification performance, which is especial evident in complex network training results. Take ResNet-101 as an example, samples mixed by 10 true and 20 simulation samples are 3.8889% better than samples mixed by 10 true and 10 simulation samples, and are 9.4445% better than 10 true samples.

According to Table IV and Table V, it can also be found that when the number of training samples increases to 30, the OA of scene classification result increases, too. In Table IV, only some well-trained networks such as VGG-16 and VGG-19 achieves 80%, while in Table V, most training results can reach 80%. A main reason for this may be that the number of training samples is required for some complex networks such as Inception-v3 and Inception-ResNet-v2, if the training data set is too small, deep networks will not perform as well as networks with fewer layers.

C. COMPARISON AND COMPATIBILITY WITH OTHER DATA AUGMENTATION METHODS

Compared with Table VI and Table VII, it can be concluded that increasing the number of training samples by flipping and rotating can indeed improve the training performance of different data sets, especially for networks with deeper layers and different structures. For example, the training performance of Inception-ResNet-v2 is almost the worst in Table III and Table IV, after increasing the training samples by flipping and rotating, the scene classification accuracy is significantly improved. In Table VII, the OA of $(20T)\_{R}$, $(10T+10S)\_{R}$, and $(20S)\_{R}$ are almost around 80% under different networks, for networks such as VGG-16, VGG-19, and ResNet-50, the OA of them are all above 80% and even 85%. The classification accuracy of $(30T)\_{R}$, $(20T+10S)\_{R}$, $(10T+20S)\_{R}$ are all exceed 80% based on all kinds of networks, for $(30T)\_{R}$ under VGG-16 and VGG-19, the OA of this data set achieves the highest, both above 90%. Because of the amount of augmented data caused by rotation is bigger than by flip, it can be seen that the improvement of the OA after rotation is more. According to the above comparison and analysis, it can be demonstrated that basic data augmentation strategy (image flip and image rotation) of true samples can improve training performance. And under the same data augmentation strategy, the more true samples, the better the classification performance we can obtain. For the samples mixed with true and simulation, this conclusion is still suitable, and the classification accuracy of the mixed samples after flip and rotation is not much different from that of the true samples. Although the overall performance of one simulation sample is inferior to one true sample, when the quantity of simulation samples gradually increases, the scene classification accuracy can be stabilized above 85%. Such experimental results are satisfactory. The above experiments demonstrate the validity of the simulation samples, i.e., the validity of our proposed data augmentation method is effective.

IV. CONCLUSIONS

In this paper, a novel and effective data augmentation method has been proposed and implemented to solve the lack of special types of aerial images for sea ice scene classification. In the experiment, to test the effectiveness of the simulation samples, the proposed method is evaluated on five groups of training data sets. The influence of different quantity of samples, different mixing ratios of simulation samples and comparison and compatibility with different basic data augmentation methods (flip and rotation) were studied. Experimental results demonstrated that when the number of simulation samples is guaranteed to a certain extent, the data set composed of simulation samples only can also obtain good classification results; when the simulation samples and the true samples are mixed according to some ratios, a better classification result can be obtained. The data augmentation method proposed in this paper is compatible with some basic data augmentation methods and the combination of them can produce powerful effects. The results of experiments consistently showed that the effectiveness of the proposed method.

ACKNOWLEDGMENT

The authors would like to thank to be supported by the Fund of the National Natural Science Foundation of China under Grant No. 61801142, No. 61601135, No. 61675051, Natural Science Foundation of Heilongjiang Province of China under Grant No. QC201706802. China Postdoctoral Science Foundation under Grant No. 2018M631912, Postdoctoral Science Foundation of Heilongjiang Province under Grant No. LBH-Z18060.

REFERENCES

[1] Miaojia L, Kronbak J, “The potential economic viability of using the Northern Sea Route (NSR) as an alternative route between Asia and Europe,” Journal of Transport Geography. 2010. 18 (3): 434—444.

[2] Cai Z, Ma H, Zhang L, “Model transfer-based filtering for airborne LiDAR data with emphasis on active learning optimization,” Remote Sensing Letters, 2018, 9(2):111-120.

[3] Simonyan, Karen, and A. Zisserman, “Very Deep Convolutional Networks for Large-Scale Image Recognition,” Computer Science .2014.

[4] Vasconcelos, Cristina N., A. Paes, and A. Montenegro, “Towards Deep Learning Invariant Pedestrian Detection by Data Enrichment,” IEEE International Conference on Machine Learning & Applications IEEE, 2017.

[5] Asperti, Andrea, and C. Mastronardo, “The Effectiveness of Data Augmentation for Detection of Gastrointestinal Diseases from Endoscopical Images,” 2017.

[6] Vitas, Diiana, M. Tomic, and M. Burul, “Zooming Innovation in Consumer Technologies Conference (ZINC) - Image Augmentation Techniques for Cascade Model Training." 2018:78-83.

[7] Bae H, Kim C, Kim N, Park B, et al., “A Perlin Noise-Based Augmentation Strategy for Deep Learning with Small Data Samples of HRCT Images,” Scientific reports, 2018.

[8] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, “Generative adversarial nets,” In Advances in neural information processing systems, pages 2672–2680, 2014.

[9] Shi, Yilei, Q. Li, and X. X. Zhu, "Building Footprint Generation Using Improved Generative Adversarial Networks," IEEE Geoscience and Remote Sensing Letters, 2018.

[10] Goodfellow I J, Pouget-Abadie J, Mirza M, et al., “Generative Adversarial Nets,” International Conference on Neural Information Processing Systems, 2014.

[11] Arjovsky, Martin, S. Chintala, and Bottou, Léon, “Wasserstein GAN,” 2017.

[12] Patan, Krzysztof, "Neural Network-Based Model Predictive Control: Fault Tolerance and Stability," IEEE Transactions on Control Systems Technology, pp.1147-1155, 2015.

[13] Ressel, Rudolf, et al., "Investigation into different polarimetric features for sea ice classification using X-band Synthetic Aperture Radar," IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing pp.:3131-3143, 2016.

[14] Onana, Vincent-De-Paul, Kurtz, Nathan T., Farrell, Sinead Louise, “A Sea-Ice Lead Detection Algorithm for Use with High-Resolution Airborne Visible Imagery,” IEEE Transactions on Geoscience and Remote Sensing 51.1(2013):38-56.

[15] Johansson, A. Malin, et al., “Newly Formed Sea Ice in Arctic Leads Monitored by C- and L-Band SAR,” Living Planet Symposium Living Planet Symposium, 2016.

[16] Bäck, Daniel, B. Holt, and R. Kwok, "Analysis of C-band Polarimetric Signatures of Arctic Lead Ice using Data from AIRSAR and RADARSAT-1," IEEE International Geoscience & Remote Sensing Symposium IEEE, 2009.

[17] Maykut, Gary A., "Energy exchange over young sea ice in the central Arctic," Journal of Geophysical Research Oceans 83. C7 (1978):3646-3658.

[18] Worby, A. P., and A. Ian, "Ocean-atmosphere energy exchange over thin, variable concentration Antarctic pack ice," Annals of Glaciology 15, 1991.

[19] An, Zhenyu, and Z. Shi, "Hyperspectral image fusion by multiplication of spectral constraint and NMF," Optik- International Journal for Light and Electron Optics 125.13(2014):3150-3158.

[20] Webber R L, Ruttimann U E, Grondahl H G, "X-ray image subtraction as a basis for assessment of periodontal changes," Journal of Periodontal Research, 2010, 17(5):509-511.

[21] Krizhevsky, A., Sutskever, I., and Hinton, G. E, "ImageNet classification with deep convolutional neural networks," In NIPS, pp. 1106–1114, 2012.

[22] He, Kaiming, et al., "Deep Residual Learning for Image Recognition," 2015.

[23] Szegedy, Christian, et al., "Rethinking the Inception Architecture for Computer Vision," 2016 IEEE Conference on Computer Vision and Pattern Recognition, 2016:2818-2826.

[24] Szegedy, Christian, et al., "Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning," 2016.

[25] AlvaroIvanoff.AMSRIce06AerialPhotographs.http://polynya.gsfc.nasa.gov/seaicearctic2006.html.2014-7-25.