Ship Classification from SAR Images based on Sequence Input of Deep Neural Network

Hongliang Zhu1,*, Nan Lin2 David Leung2
1School of Science and Engineering, The Chinese University of Hong Kong, Shenzhen, China
2Artificial Intelligence LAB, Washington University in St. Louis, St. Louis, USA

*Corresponding author e-mail: zhuhongliang@cuhk.edu.cn

Abstract. This article proposes an all brand new classification architecture for SAR images of the ship via deep learning. Compared with the most widely used conventional classification algorithm based on deep neural networks, we set a sequence of SAR images as input to the neural network, which is different from a single image as input in those kinds of conventional ones. Two central neural networks form the construction of this new classification architecture; the first is a convolutional neural network used to extract the features of each frame, and the second is the LSTM part, which is trained on the sequence to predict the labels. It is much more significant to enhance the connection among the SAR images primarily due to the limit of SAR data, which makes us combine all the SAR images to several sequences. The proposed new classification architecture is trained on the OpenSARShip[2] dataset captured by Sentinel-1 space-borne radar. Experiment results show that the classification architecture can get an accuracy of 99.24% in classifying the six kinds of targets, which is much more precise than any conventional methods before.

1. Introduction
Ship classification tasks in the ocean are of great significance for enhancing the maritime domain awareness (MDA), which aims to identify different categories of ship as accurately as possible. This topic is quite crucial in military and civilian fields.

To acquire such ship data, synthesis aperture radar performances a vital role in capturing the different types of ships due to its superiority of all-weather working conditions, day and night running time, and a large scale of surveillance surface in the ocean. It has become the leading equipment to offer data in the remote sensing field. With the rapid development of SAR, higher and higher radar imaging resolution has touched 0.3m like Pleiades Neo, etc. These kinds of radar provide powerful hardware support for ocean surveillance, which always decreases the influence by complex sea navigation conditions. Moreover, SAR data contains much more information than conventional optical images, such as the phase position and polarization information, which are quite useful to detect the target in the clutter. But some disadvantages like the speckle noise does exist in the SAR image, which is a challenging problem all the time.

For a ship classification problem, researchers have been exploring the most efficient method to tackle this kind of problem for several decades. The year 2012 was a watershed for image classification
problems, Geoffrey and Alex [3] won the image classification competition in ImageNet with super high accuracy rate among all the participants via a neural network construction named AlexNet. It makes deep learning, especially convolutional neural networks, become the principal methodology in the image classification field after 2012 for its excellent capability of feature learning and representation. Some researchers also use this kind of network to tackle the ship classification problem with SAR image. Most of them use a single SAR image as the input to train the whole convolutional neural network with the data augmentation technology for the limit of SAR data [2], [6]. The output does show not wicked accuracy in the test dataset, but it seemed the relationship between each SAR image had been omitted.

Due to this problem, this article proposes an all brand new classification architecture with a convolutional neural network and an LSTM module using the sequence of SAR images as the input. We use the convolutional layers to extract ship features by applying the convolutional operations to each frame of the sequence independently. After that, convert the output to vector sequences by using a sequence unfolding layer and a flatten layer. Finally, this vector sequence will be passed through the LSTM to output the final classification result. We organize the rest pages as follows.

Section II gives the structure of the proposed architecture and explains the whole training and classification process. Section III presents the classification results with discussions, and Section IV concludes this paper.

2. Proposed architecture for ship classification from SAR images

For a ship classification problem, two main stages are usually required during the whole process: feature extraction and classification. See Fig.1. There are many existed conventional algorithms applied in the feature extraction stage, such as Fourier Transform, Wavelet, EMD, Tight frame, etc. Also, such classifiers like Naive Bayes, SVM, HMM, Decision Trees are widely used in the second stage.

Nevertheless, the emergence of deep learning, especially convolutional neural networks, makes the era changed for its most potent capacity of feature extraction and classification in single pipeline. We found that most of such networks' input is one single image so far [5], [8], [9]. Hence, it makes us more interested to use a sequence of SAR images as the input instead of conventional ones. We use several convolutional layers to generate the feature vectors and an LSTM to classify those feature vectors.

2.1. Detailed description for proposed architecture.

As shown in Fig.2, the complete structure can be divided into four main parts, which are the sequence input layer, convolutional layers, LSTM and output layers.

![Figure 1. Ship classification problem with several steps](image1)

![Figure 2. Proposed architecture for ship classification with sequence input](image2)
SAR at different times with several degrees of rotation. Convolutional layer 1, 2, 3 contains 24, 48, 96 feature maps respectively with the same size convolutional kernel sharing of 4*4. The RELU layer is applied to enhance the nonlinearity of the network for excellent capability of generalization. Moreover, each of three convolutional layers is followed by a max-pooling layer, with a pooling size of 2 × 2 and a stride of 2 pixels. Convolutional layers convert the images sequences to sequences of feature vectors, where the feature vectors are the output of the activations function on the last pooling layer of the CNN network. Next, an LSTM network is created that can classify the sequences of feature vectors representing the image sequence. At last, output layer outputs the label of each input sequence.

**Figure 3.** Detailed description for proposed classification architecture

### 2.2. Training of the proposed architecture.

We train the whole classification architecture with back-propagation algorithm, which is a widely used method to update the network’s parameters during the training process. However, there are two main parts in our proposed architecture, which means the output of CNN is the input of the LSTM. We separate the whole back-propagation into two subparts connected with the feature vectors. In the output layer of LSTM, the error can be described as the difference between the real target and the prediction one.

\[ \delta_k(t) = y_k(t) - t_k(t) \]  \hfill (1)

As shown in Figure 4, this error goes back through the LSTM to the input of it and determines how accurate the feature vectors extracted by the CNN. Then, we got another error represented by the

\[ e_k(t) = \delta_k(t) \]  \hfill (2)
Training error propagation through the architecture

We can update the parameter of CNN with this error going back through the CNN from higher layer to lower layer. For the sequence of images \( x, (i) \), ship class \( y, (i) \), the cross-entropy loss function can be defined as:

\[
L(\omega) = -\frac{1}{m} \sum_{i=1}^{m} \log P(y^{(i)}|x^{(i)}; \omega)
\]  

(3)

3. Experiment result and analysis.

In this article, we conduct our experiment on the OpenSARShip dataset acquired by Sentinel-1. It contains 10,045 SAR ship chips with four categories: boat, cargo, container, and tanker. The optical image and SAR image of each category sample are shown in Fig 5.

\[\text{(a)} \quad \text{(b)} \quad \text{(c)} \quad \text{(d)}\]

Figure 5. Optical and SAR images of 4 ships (a) Boat (b) Cargo (c) Container (d) Oil Tanker

The first step in our proposed classification architecture is constructing the sequence of SAR ship images. As shown in Figure 6, we pick every 10 SAR images in each ship category to form a sequence as the input for the classification architecture. Moreover, to prevent the overfitting problem usually occurred during the deep neural network training period, we construct the different sequences with different order among these 10 SAR images. It will generate \(10^6\) sequences of SAR image, but we only select 500 of them randomly. Therefore, the training set contains 10,000 sequences. For the test-set, we
only use ten same SAR ship images to form each sequence because, in the practical classification problem, it usually compares the classification accuracy with a single image. Table 1 shows the concrete amount of each ship in OpenSARShip dataset and the quantity of training and test sequence we are to use.

Figure 6. Construction of SAR image sequence

| Ship Type       | Train | Test | Train Sequence | Test Sequence |
|-----------------|-------|------|----------------|---------------|
| Boats           | 2146  | 512  | 100,000        | 512           |
| Cargo Ship      | 1893  | 463  | 90,000         | 463           |
| Container Ship  | 2041  | 512  | 100,000        | 512           |
| Oil Tanker      | 1966  | 512  | 100,000        | 512           |

This proposed classification architecture is trained on the computation platform with 10 NVIDIA TitanV GPUs. Figure 7 shows the training progress and classification accuracy with our classification architecture. Training set curve and validation curve are pretty close, which means there is no overfitting occurred during the training progress. The accuracy of training progress is 100%. The classification accuracy on the test set shows that the average accuracy is 99.85%, which is much higher than any conventional algorithm before. We compare the classification accuracy of our architecture with other methods in Table 2.

Figure 7. Training progress
Figure 8. Classification accuracy heatmap

Table 2. Classification accuracy of our architecture compared with others

|                  | Our architecture | CNN     | CNN-MR  | CNN+SVM |
|------------------|------------------|---------|---------|---------|
| Boats            | 99.81%           | 94.12%  | 87.71%  | 98.82%  |
| Cargo Ship       | 100%             | 97.33%  | 93.40%  | 99.12%  |
| Container Ship   | 99.62%           | 95.29%  | 96.34%  | 97.40%  |
| Oil Tanker       | 100%             | 96.85%  | 97.52%  | 98.75%  |
| Average          | 99.85%           | 95.90%  | 93.74%  | 98.52%  |

4. Conclusion.
This article proposed a new classification architecture for SAR images of the ship, which get the highest accuracy among all the classification algorithms so far. We combine the convolution layers and LSTM in one integrated network with the input of a sequence image. Different from conventional CNN with a single image for the input, we strengthen the connection between each image and utilize the strong predict time series capability of LSTM. The experiment result shows that it works fantastic in the ship classification problem.

References
[1] J. I. Park and K. T. Kim, “Modified polar mapping classifier for SAR automatic target recognition,” IEEE Trans. Aerosp. Electron. Syst., vol. 50, no. 2, pp. 1092–1107, Apr. 2014
[2] Lanqing Huang, Bin Liu, “Opensarship: A Dataset Dedicated To Sentinel-1 Ship Interpretation”, Ieee Journal Of Selected Topics In Applied Earth Observations And Remote Sensing, Vol. 11, No. 1, January 2018
[3] Alex Krizhevsky, I Sutskever, G Hinton, “ImageNet Classification with Deep Convolutional Neural Networks”, Advances in Neural Information Processing Systems, 2012, 25 (2): 2012–2022
[4] D. Velotto, C. Bentes, B. Tings, and S. Lehner, “Comparison of Sentinel-1 and TerraSAR-X for ship detection,” in Proc. 2015 IEEE Int. Geosci. Remote Sens. Symp., 2015, pp. 3282–3285
[5] X. Xing, K. Ji, H. Zou, W. Chen, and J. Sun, “Ship classification in TerraSAR-X images with feature space based sparse representation,” IEEE Geosci. Remote Sens. Lett., vol. 10, no. 6, pp. 1562–1566, Nov. 2013.
[6] G. Margarit and A. Tabasco, “Ship classification in single-pol SAR images based on fuzzy logic,”
IEEE Trans. Geosci. Remote Sens., vol. 49, no. 8, pp. 3129–3138, Aug. 2011

[7] W. Chen, K. Ji, X. Xing, H. Zou, and H. Sun, “Ship recognition in high resolution SAR imagery based on feature selection,” in Proc. 2012 Int. Conf. Comput. Vision Remote Sens., 2012, pp. 301–305.

[8] C. Bentes, D. Velotto, and S. Lehner, “Analysis of ship size detectability over different TerraSAR-X modes,” in Proc. 2014 IEEE Geosci. Remote Sens. Symp., 2014, pp. 5137–5140.

[9] C. Wang, H. Zhang, F. Wu, S. Jiang, B. Zhang, and Y. Tang, “A novel hierarchical ship classifier for COSMO-SkyMed SAR data,” IEEE Geosci. Remote Sens. Lett., vol. 11, no. 2, pp. 484–488, Feb. 2014.

[10] P. W. Vachon, J. Wolfe, and H. Greidanus, “Analysis of Sentinel-1 marine applications potential,” in Proc. 2012 IEEE Int. Geosci. Remote Sens. Symp., 2012, pp. 1734–1737.