COMPARATIVE STUDY OF FEATURE EXTRACTION APPROACHES FOR SHIP CLASSIFICATION IN MODERATE-RESOLUTION SAR IMAGERY

Shreya Sharma, Kenta Senzaki and Hirofumi Aoki

Data Science Research Laboratories, NEC Corporation

ABSTRACT

This paper presents a comparative study of existing feature extraction approaches for ship classification in moderate-resolution synthetic aperture radar (SAR) images. Ship classification is a key functionality in many maritime surveillance applications. For efficient ship classification, appropriate feature extraction is crucial. Most of existing studies have used high-resolution images. For maritime surveillance, however, wide-area coverage is essential whereas it inevitably reduces the spatial resolution. In this paper, we evaluate the applicability of representative methods to moderate-resolution images. The evaluated methods are hand-crafted feature extraction (HCF), principal component analysis (PCA) and autoencoder (AE) based on neural-network. The evaluation is done on the basis of accuracy for two-class ship classification into tanker and cargo. The experiments demonstrate that AE outperforms HCF and PCA in classification accuracy by 7.4% and 2.6%, respectively. Furthermore, AE performs best even in classification of challenging cases such as small ships.

Index Terms— ship classification, feature extraction, deep learning, maritime, SAR, moderate resolution

1. INTRODUCTION

In recent years, maritime traffic has increased tremendously due to high demand of global trade and sea-food products. As shipping traffic grows, not only problems of accidents and environmental damages but also illegal maritime activities, such as illegal fishing, oil spills and human trafficking are increasing. With respect to maritime security and environmental protection, ship surveillance by remote sensing has proven to be a key technology in many maritime surveillance applications. Out of all remote sensing technologies, synthetic aperture radar (SAR) is an important tool which provides wide spatial coverage with all-weather and day-and-night acquisition capability.

For reliable and quick responses to maritime security issues, we need ship classification as well as ship detection to identify the ships involved in illegal activities. For efficient ship classification, appropriate feature extraction of ships is crucial.

Numerous features have been proposed so far for effectively discriminating the appearance of different ships such as cargos, tankers and containers [1]. Some features describe the discriminative geometry of different ships such as length, area while others distinguish their unique backscatter patterns such as width-ratios, auto-correlation and ratio of bright and dark pixels along the ship's main axis [2]. These features are generally hand-crafted and therefore they might be less effective under non-tested conditions such as lower spatial resolution.

Apart from the above hand-crafted features, automatic feature extraction approaches have been proposed. They are roughly classified into classical dimensionality reduction and deep learning (DL). A representative method for the former approach is principal component analysis (PCA) and it has been applied to ship classification by Gouaillier et al. [3]. Recently, as in the general image recognition field, DL-based approach has been introduced for SAR-based object recognition and Bentes et al. [4] applied autoencoder (AE), which is one of DL-based methods, to the ship classification.

Most of such existing feature extraction methods have been tested with high-resolution SAR imagery. It is because objects in high-resolution images tend to show discriminative patterns similar to that in optical images, whereas generally SAR images and optical images are totally different in terms of reflection properties, geometric distortion and so on. In contrast, for applications in maritime surveillance where the objective is to monitor large marine areas, moderate-resolution images are required as they provide wider swath [5]. Therefore, considering real-world application of ship classification technology, it is generally required for ship classification methods to be adaptable to lower resolution SAR imagery. This paper presents a comparative study on the effectiveness of the above mentioned feature extraction approaches to the ship classification in moderate-resolution SAR imagery.

2. FEATURE EXTRACTION METHODS

The process of ship classification using SAR image is shown in Fig. 1 wherein the input is a SAR image and the output is ship classes for ships captured in the image. The three major steps in the process are ship detection, feature extraction and classification. We focus on feature extraction in this paper. The input for feature extraction process is a set of SAR sub-

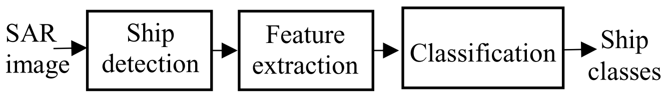
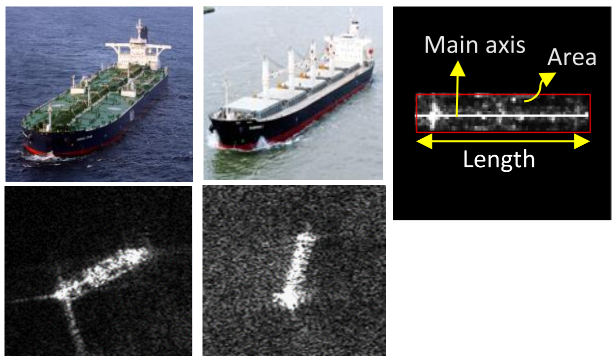


Fig. 1. Process of ship classification using SAR.



(a) (b) (c)

Fig. 2. Optical and SAR ship-chips (a) tanker, (b) cargo. (c) MBR extraction from a ship-chip.

images each containing a ship, which are extracted by ship detection. Such sub-images are hereafter referred to as ship-chips. Figure 2 (a)-(b) depicts exemplary optical and SAR ship-chips for the two classes: tanker and cargo. Features are extracted from each SAR ship-chip. The three conventional feature extraction methods are explained in detail below.

2.1. Hand-crafted feature extraction method (HCF)

In this paper, seven conventional hand-crafted features for ship classification are selected out of twenty-one conventional features based on preliminary experiments evaluating the factors such as stability, discriminability and multi-collinearity. The details of these factors for feature selection can be found in [6]. Out of the seven selected features, five are geometric features while two are backscattering features. Table 1 depicts the selected features and their corresponding formulas.

To compute the above features, we first rotate the ship-chips along their main axis as shown in Fig. 2 (c), and then extract minimum bounding rectangle (MBR) around each ship. Figure 2 (c) depicts a sample ship-chip with extracted MBR and two features L and A. The features L, A, SC and C are obtained from the dimensions of the extracted MBR while the features M and V are obtained by computing mean and variance of backscattering values of all the pixels within MBR. To calculate the feature R10, the ship-chips are binarized with an appropriate threshold. Then, R10 is computed as the ratio of number of one-valued and zero-valued pixels along the ship's main axis. Since tanker has an oil pipeline passing through its main axis, it has many one-valued pixels along the main axis. In contrast, in cargo, due to the alternate hatches, the number of one-valued pixels along the main axis is less. As a result, R10 is higher for a

Table 1. Extracted hand-crafted features of ship (W: width, P: Perimeter).

Feature	Formula
Length (L)	L
Area (A)	LW
Shape Complexity(SC)	$4\pi A/P^2$
Compactness (C)	$P/2\pi L$
R10	Ratio of no. of one and zero valued pixels along main axis
Mean backscatter (M)	Mean backscattering value
Variance (V)	Variance of backscattering value

tanker compared to that of a cargo and can be used as a discriminative feature [1]. Finally, a feature vector (FV) containing the seven features is constructed for each ship-chip and is denoted as

$$FV = \{L, A, SC, C, R10, M, V\}. \quad (1)$$

2.2. Principal Component Analysis method (PCA)

PCA is a dimensionality reduction method which has proved efficient in target recognition using SAR under different operational conditions. To extract features using PCA, each ship-chip of size $N \times N$ is flattened to a one-dimensional vector of size N^2 by concatenating all rows together. Taking each pixel as a feature, each ship-chip has N^2 features. PCA is applied on the ship-chips to reduce the number of features from N^2 to m.

Let $X \in \mathbb{R}^{N^2}$, $P \in \mathbb{R}^{N^2 \times m}$ and $Y \in \mathbb{R}^m$ be a concatenated input ship-chip, a projection matrix and a reduced feature vector, respectively. The reduced feature vector Y for each ship-chip can be computed by

$$Y = P^{\mathrm{T}} X. \quad (2)$$

In this paper, m is set to 177 which is equal to the number of ship-chips and captures 100% variance of features.

2.3. Autoencoder method (AE)

AE is a self-supervised neural-network, where the output is forced to be same as the input. The objective of the AE is to learn the best latent representation, i.e. implicit features, of the input data which can be used to reconstruct it. In order to prevent the AE from learning an identity mapping function between the input and the output, the size of the latent representation is kept lower than the input size. We use a three-hidden-layer AE for feature extraction of ships.

The architecture of our AE is depicted in Fig. 3 where the input ship-chip is denoted by X, the output reconstructed ship-chip is denoted by X' and the latent representation is denoted by Z. Each input ship-chip is flattened by concatenating all rows together and then input in the AE. The implicit features extracted in the latent representation layer Z are used as a feature vector for each ship-chip. The loss function which is used to compare the input ship-chip (X) and

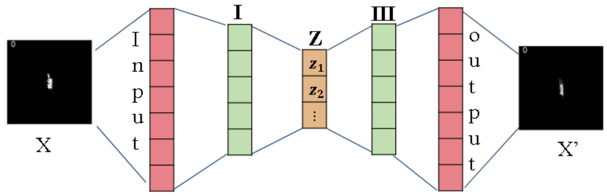


Fig. 3. Architecture of 3-hidden layer AE.

the output reconstructed ship-chip (X') is binary cross-entropy which is given by

$$L(X, X') = \frac{1}{N^2}\sum_{i=1}^{N^2} X_i \log(X_i') - (1 - X_i)\log(1 - X_i'), \qquad (3)$$

where i denotes the i^{th} pixel and N^2 is the total number of ship pixels in one ship-chip.

The input data is divided into training and validation data for training and hyper-parameter tuning respectively. The most important hyper-parameter to tune the AE is the size of the latent representation Z. For this purpose, six different size combinations of the first and the third hidden layers and the Z layer are used to train the network and then compared based on the reconstruction loss for the validation data. As shown in Table 2, the Z layer size of 128 resulted in the minimum reconstruction loss. The epochs used to train the network are 100 and the network is optimized using mini-batch gradient descent with the batch size of 4. Finally, the feature vector of each ship-chip obtained from Z is denoted as

$$Z = \{z_1, z_2, z_3, z_4, \ldots, z_{128}\}. \qquad (4)$$

3. EXPERIMENTS

The experimental data consists of 11 SAR images (over Tokyo and Osaka bay area) acquired by ALOS-2 PALSAR with HH polarization in Stripmap mode. It should be noted that Stripmap mode is the most common mode used for acquiring moderate-resolution and moderate-swath images instead of acquiring high-resolution images at the expense of narrow swath. The azimuth and range resolution of the images is 6m x 6m. The image acquisition dates range from August 9, 2016 to September 26, 2016. We used Constant False Alarm Rate (CFAR) algorithm to detect ships in the image which is an adaptive algorithm used in radar systems to detect targets from a background clutter [7]. Using CFAR ship detection on the 11 scenes, 177 ship-chips are extracted each of size 128x128, out of which 67 are tankers and 110 are cargos. All ship-chips were labeled using ground truth information provided by Automatic Identification System.
For each ship-chip, we extracted three feature vectors using (1), (2) and (4). We performed classification using SVM on the three feature vectors separately to get the class of each ship-chip. To remove any bias due to random splitting of the dataset into train and test sets, we performed five-fold nested cross-validation with grid search for parameter selection of SVM and training of AE. We evaluated methods on the basis of accuracy, precision, recall and F-measure.

Table 2. Reconstruction loss versus hidden layer sizes.

I/III	64	128	128	**256**	512	1024	1024
Z	20	32	64	**128**	256	256	512
Loss	243	241	240.7	**240**	240.9	240.6	241

Furthermore, ship properties such as length and speed affect the extraction of discriminative features. It is difficult to obtain discriminative features from small ships (relative to SAR image resolution) as the number of ship pixels becomes less and then distinguishing patterns are lost. Similarly, in case of high speed ships (relative to SAR platform velocity), the distinguishing patterns get blurred. Therefore, we evaluate the effectiveness of each feature extraction method for different ship lengths and speeds. For this purpose, the total 177 ship samples are divided into three bins based on length: small (<101m), medium (101m-150m) and large (>150m) and two bins based on speed: slow (<=10 knots) and fast (>10 knots). We compared percentage of correctly classified samples in each bin to evaluate the methods.

3.1. Results and Discussion

Table. 3 (a)-(c) depicts the confusion matrices of ship classification for HCF, PCA and AE, respectively. It can be concluded that AE outperforms the HCF and PCA method by 7.4% and 2.6%, respectively, in overall classification accuracy. To understand the effectiveness of feature extraction methods for each ship class, we performed class-wise analysis using precision, recall and F-measure as shown in Fig. 4. It shows that AE improves the recall for the small proportion class tanker by 44% compared to HCF. As a result, the number of false positives in class cargo is reduced and the precision is improved by 7.5% compared to HCF. Overall, AE achieves the best F-measure for both the classes followed by PCA and then HCF. The improvement in classification by AE can be explained as follows. AE is a data-driven approach and has a capability to extract features that are not limited by the domain knowledge which HCF method requires. Thus, AE is able to learn more information-rich features. Moreover, AE learns both non-linear and linear relationships among ship pixels, unlike PCA which only learns linear relationships, and thus understands complex parts of a ship better than PCA. These observations are consistent with previous research [8].

Table 3 (d) shows the percentage of the correctly classified samples by each method for different ship lengths and speeds. It can be observed that, for small ships, all the three methods give low percentage of correctly classified samples compared to medium and large ships. This shows that it is quite challenging to accurately classify small ships in moderate-resolution SAR images. For medium ships, PCA and AE give higher number of correctly classified samples compared to HCF, while AE gives the highest performance in large ships. Overall, AE gives the most correctly classified samples for each bin. When considering ship speeds, slow ships can be more accurately classified compared to fast ships

Table 3. Confusion matrices (a) HCF, (b) PCA, and (c) AE. (d) Percentage correctly classified samples on the basis of ship lengths and speeds.

Predicted Class / **Actual Class**

(a)

	Tanker	Cargo	Recall
Tanker	21.9	45.10	0.32
Cargo	17.03	93.97	0.84
Precision	0.56	0.67	**65.4% (+/-2)**

(b)

	Tanker	Cargo	Recall
Tanker	28.97	38.03	0.43
Cargo	17.83	92.17	0.83
Precision	0.62	0.71	**68.4% (+/-1)**

(c)

	Tanker	Cargo	Recall
Tanker	30.90	36.10	0.46
Cargo	16.77	93.23	0.85
Precision	0.64	0.72	**70.2% (+/-1)**

Method	Length			Speed	
	small	**med**	**large**	**slow**	**fast**
HCF	62%	61%	79%	69%	61%
PCA	60%	69%	81%	68%	**70%**
AE	**63%**	**70%**	**83%**	**76%**	65%

(d)

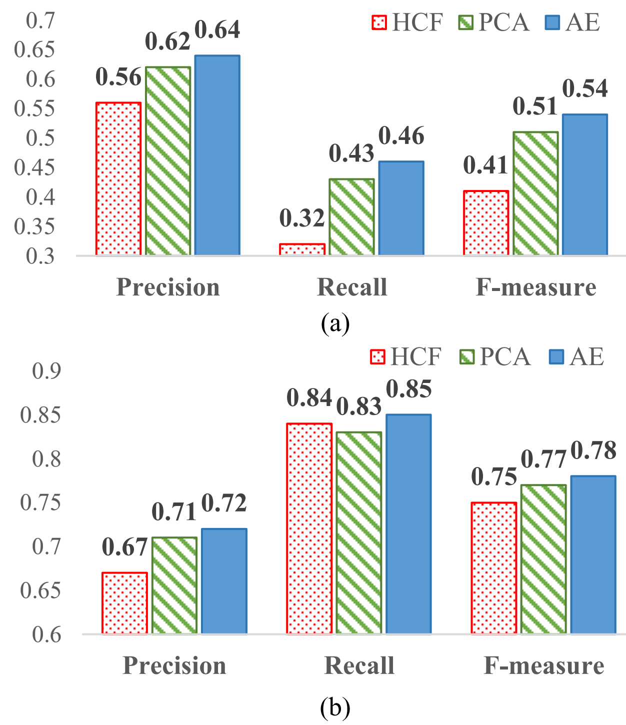


Fig. 4. Precision, recall, and F-measure (a) tanker, (b) cargo.

due the less distortion in ship SAR backscatter. AE gives the highest number of correctly classified samples in slow ships. However, in fast ships, PCA performs better than AE. Evaluation of the methods for fast ships is difficult as we are still unaware which ship features are getting distorted due to high speed. This can be taken as a future work of our research.

4. CONCLUSION

This paper presented a comparative study of conventional feature extraction methods for ship classification in moderate-resolution SAR imagery. As representative methods, HCF, PCA and AE are evaluated. The methods were compared quantitatively on the basis of overall classification accuracy and class-wise analysis. In addition to that, the methods were also evaluated for different ship lengths and speeds. The experiments verified that AE-based feature extraction outperforms the other conventional methods and can extract discriminative features even for challenging cases such as small ships. However, in cases such as fast ships, AE gives sub-optimal performance, the reasoning for which is still unknown. Further analysis using larger data set having a variety of ships in terms of class, size, speed and a variety of resolutions, is left as a future work.

5. REFERENCES

[1] H. Zhang, X. Tian, C. Wang, F. Wu, and B. Zhang, "Merchant vessel classification based on scattering component analysis for COSMO-SkyMed SAR images," *IEEE Geoscience and Remote Sensing Letters*, vol. 10, no. 6, pp. 1275-1279, 2013.

[2] M. Ziang, X. Yang, Z. Dong, S. Fang, and J. Meng, "Ship classification based on superstructure scattering features in SAR images," *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 5, pp. 616-620, 2016.

[3] Gouaillier, and L. Gagnon, "Ship silhouette recognition using principal components analysis," *SPIE Proc,* vol. 3164, pp. 59-69, 1997.

[4] C. Bentes, D. Velotto, and S. Lehner, "Target classification in oceanographic SAR images with deep neural networks: Architecture and initial results," In *Geoscience and Remote Sensing Symposium (IGARSS), 2015 IEEE International,* pp. 3703-3706, 2015.

[5] H. Lang, and S. Wu, "Ship Classification in Moderate-Resolution SAR Image by Naive Geometric Features-Combined Multiple Kernel Learning," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 10, pp. 1765-1769, 2017.

[6] W. T. Chen, K. F. Ji, X. W. Xing, H. X. Zou, and H. Sun, "Ship recognition in high resolution SAR imagery based on feature selection," *Computer Vision in Remote Sensing (CVRS), 2012 Int. Conf. on*, pp. 301-305, 2012.

[7] G. Gao, L. Liu, L. Zhao, G. Shi, and G. Kuang, "An adaptive and fast CFAR algorithm based on automatic censoring for target detection in high-resolution SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 6, pp. 1685–1697, 2009.

[8] G.E. Hinton, and R.R. Salakhutdinov. "Reducing the dimensionality of data with neural networks," *science*, 313 (5786), pp. 504-507, 2006.