Ship-Iceberg Discrimination with Convolutional Neural Networks in High Resolution SAR Images

Carlos Bentes, German Aerospace Center (DLR), carlos.bentes@dlr.de, Germany
Anja Frost, German Aerospace Center (DLR), anja.frost@dlr.de, Germany
Domenico Velotto, German Aerospace Center (DLR), domenico.velotto@dlr.de, Germany
Björn Tings, German Aerospace Center (DLR), bjoern.tings@dlr.de, Germany

Abstract

The application of Synthetic Aperture Radar (SAR) for ship and iceberg monitoring is important to promote maritime safety in Arctic waters. Although the detection of ships and icebergs in SAR images is well established using adaptive threshold techniques, the discrimination between the two target classes still represents a challenge for operational scenarios. This paper proposes the application of Convolutional Neural Networks (CNN) for ship-iceberg discrimination in high resolution TerraSAR-X StripMap images. The CNN model is compared with a Support Vector Machine (SVM), and the final results indicate a superior classification performance of the proposed method.

1 Introduction

Spaceborne Synthetic Aperture Radar (SAR) is an important instrument for oceanographic observations. Due to its active radar, it is able to monitor the oceans and floating structures in all weather conditions. In SAR images, ships and icebergs typically have a stronger backscatter response than the surrounding open water, and are therefore detectable using adaptive threshold techniques. The Constant False Alarm Rate (CFAR) algorithm is typically applied to detect ships and icebergs [5, 8, 13, 14]. Icebergs occur in a large variety of sizes and shapes, imposing additional challenge to their detection in SAR images. In [7], the detection of icebergs is performed in high resolution TerraSAR-X images using iterative censoring CFAR (IC-CFAR), resulting in a higher iceberg detectability at lower false alarm rates.

After detection, additional processing is needed to distinguish between ships and icebergs. The discrimination between the two classes is carried out through feature extraction and target classification steps. In [6], intensity and polarimetric parameters are used as features to a Support Vector Machine (SVM) classifier, in order to discriminate ships from icebergs in simulated, dual polarized, medium resolution SAR data. In [9], the differences in the dominant scattering mechanism between ships and icebergs are used to divide the two classes. These methods exploit polarimetric information from medium resolution SAR products, and therefore, not applicable for the discrimination of ships and icebergs in single-polarization data. However, with the availability of high resolution SAR image, the discrimination can be performed using intensity and shape features.

This abstract proposes the application of Convolutional Neural Networks (CNN) to ship-iceberg discrimination in high resolution TerraSAR-X data. Convolutional Neural Networks are able to learn complex representations from the input data, without the need of handcrafted features, and have been successfully used in many image classification tasks [2, 10, 11]. Only recently CNN have been successfully adopted for demanding SAR classification tasks, as in [3, 12]. This abstract is organized as follows: Section 2 describes the dataset and CNN architecture used in our experiments. Section 3 presents the experimental results using the proposed model trained with targets extracted from TerraSAR-X StripMap products. Finally, Section 4 presents our conclusions and future work.

2 Development

Figure 1-(a)-(b) illustrate the typical SAR signature of a ship and an iceberg in an X-Band high resolution image. The high resolution of TerraSAR-X StripMap mode allows the detection of structural components from both floating structures. Even though ships and icebergs can have similar intensity and size values, their structures and shapes typically follow different patterns. The CNN is designed to learn a set of features from the input image in a supervised training process, in a way to capture the differences between the two observed classes.

The classifier architecture is presented in Figure 2. It is composed of two convolutional layers (Feature Map layer followed by a Pooling layer) and a fully connected layer D. The output class is defined using a soft-max operation in the final layer S. The CNN model is implemented in Python using the optimized library Theano [1, 4].

The classification dataset is composed of ships and icebergs extracted from TerraSAR-X Multi Look Ground Range Detected (MGD) products, with ground resolution of 3 meters. A total of 277 ships and 68 icebergs are used in the classification dataset. The targets (ships and icebergs) are extracted from the SAR images and stored in a

sub-image set of Regions of Interest (ROI) R_{target}.

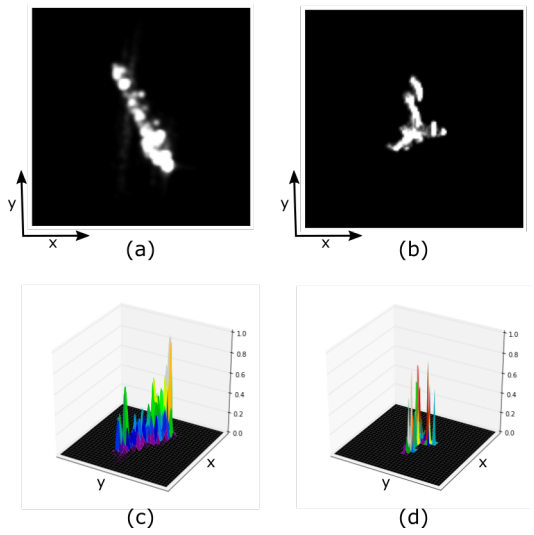

Figure 1: Example of a ship and an iceberg SAR signature: (a) Ship structure, (b) Iceberg structure, (c) 3D visualization of the ship structure, (d) 3D visualization of the iceberg structure. Note that both ship and iceberg are similar in size and intensity, but differ significantly in their spatial signature

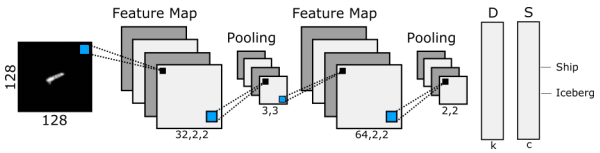

Figure 2: Convolutional Neural Network architecture used in our experiments is composed of two convolutional layers (Feature Map + Pooling operation), one full connected layer (D), and one soft-max layer to generate the classification probabilities.

To all elements in the set R_{target}, the following operations are performed: 1) Image sampling to a pixel spacing of 1.5 meters; 2) Image cropping to 128×128 pixels centered in the ROI element; 3) Image normalization.

The image sampling and cropping transform the image in a standard input vector of 128×128 elements. This is necessary because the neural network input filed has a fixed number of elements. The image normalization process is responsible to reduce strong backscattering signals from targets. The strong backscattering represented by high pixel values can introduce instabilities in the training process. In addition to that, target's peak values do not generalize well in the classification task, once the high value is highly dependent of target orientation towards the satellite. The following nonlinear normalization function is proposed:

$$N(x) = \frac{L(x)}{maxL(x)} \tag{1}$$

where:

$$L(x) = \begin{cases} 1 + \log x & \text{if } x > 1 \\ x & \text{if } x \leq 1 \end{cases} \tag{2}$$

The log function in equation 2 attenuates the high intensity values of the target signal. The normalization function changes the target pixel distribution from a highly skewed (right tail) distribution to a normal-like symmetric distribution.

The effects of the normalization processing step are illustrated in Fig. 3, where Fig. 3-(a) and Fig. 3-(c) are scaled from $min(ROI)$ to $max(ROI)$, in order to better depict the information content from the target's signal. The target profile is presented in Fig. 3-(b) and Fig. 3-(d), using $3D$ plots (ROI coordinates $x \times y$ vs ROI intensity).

The classification dataset is artificially enlarged, in order to balance the number of samples per class and to avoid model overfitting. This process is performed by sampling (with replacement) the ships and icebergs in the dataset, applying a set of label-preserving transformations. The following transformations were used: Horizontal and vertical reflections; Image rotation; Image translation. The augmented dataset contains 600 elements (300 ships, 300 icebergs), and is divided in two groups: training dataset, with 90% of data; test dataset, with 10% of data.

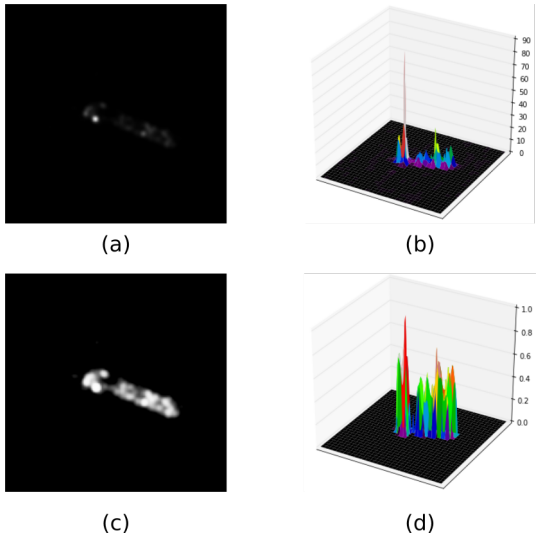

Figure 3: ROI normalization preprocessing. (a) The input target, with image scaled from $min(ROI)$ to $max(ROI)$ to better visualize the unbalanced intensity distribution of the target. (b) $3D$ representation of the input ROI. (c) Visualization of the target intensity distribution, from $min(ROI)$ to $max(ROI)$. (d) $3D$ representation of the normalized ROI.

3 Experimental Results

The CNN structure is optimized using training error curves, to determine the network performance during the training phase. Different convolutional layer sizes were tested, and the best model over the training dataset was selected. The final architecture was then validated in the

test dataset. Two Support Vector Machine (SVM) models were also trained using the same dataset, for comparison purposes. The first SVM model uses as feature vector parameters extracted from the segmented target contours. This feature vector is composed of: area, perimeter, seven central moments, and seven Hu invariant moments. The resultant vector has 16 elements per target sample. For the second SVM model, the target sub-image (128×128 pixels) is transformed in a vector of 16384 elements (stacking sub-image elements, row by row, from the 128×128 input image). After that, the input vector is compressed to a feature vector of 60 components using Principal Component Analysis (PCA). The SVM models are optimized using cross validation over the training dataset, and hyper-parameters are chosen using grid search methods. The best model over the training dataset is then validated using the same test dataset from the CNN model.

Table 1 presents the results obtained with all models in the test dataset. The final score is evaluated using the f1-score, defined as the harmonic mean between the precision and the recall percentages, as illustrated below in Equation 3:

$$\text{f1-score} = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \qquad (3)$$

The CNN model, with f1-score of 97%, performs better than the SVM models. This result indicates that the convolutional layer is able to extract useful features during the training process, resulting in a better generalization during the test phase.

Classifier	Target	Precision	Recall	F1-Score
CNN	Iceberg	95%	100%	98%
	Ship	100%	95%	97%
	Average	98%	97%	97%
SVM	Iceberg	88%	88%	88%
	Ship	88%	88%	88%
	Average	88%	88%	88%
PCA+SVM	Iceberg	100%	88%	94%
	Ship	90%	100%	94%
	Average	95%	94%	94%

Table 1: Classification results: All models are trained and have their hyper-parameters selected using the training dataset. The best final model is then validated using the test dataset. The results are presented in the form of the f1-score, to balance both precision and recall performances.

Figure 4 illustrates the CNN classification output. Figure 4-(a) shows a ship, correctly classified with high probability output. Figure 4-(b) shows an elongated iceberg structure, correctly classified as iceberg. It is possible to observe that the network also considered the elongated structure as a ship, but with lower probability. Figure 4-(c) illustrates a misclassified ship. The visible signature of the ship in the image is small, misguiding the network to select the iceberg class.

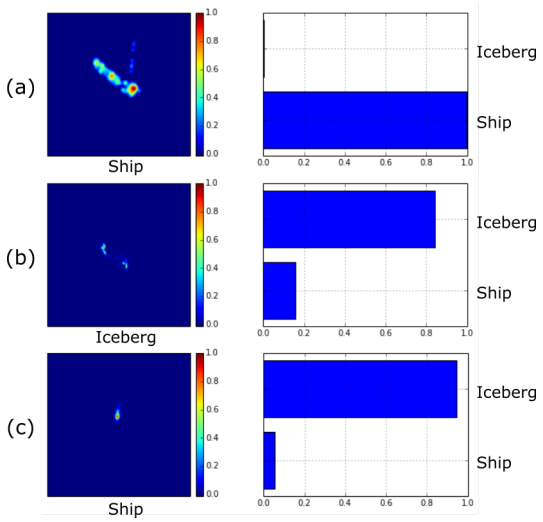

Figure 4: CNN classification results, illustrating the normalized target sub-image (on the left) and the classification probability plot (on the right). (a) example of a correctly classified ship; (b) example of a correctly classified iceberg; (c) example of a misclassified ship.

4 Conclusion

This paper presents the use of Convolutional Neural Networks for ship-iceberg discrimination in high resolution TerraSAR-X images. A total of 277 ships and 68 icebergs were extracted from StripMap MGD images, and used to build a balanced augmented dataset of 600 samples. The CNN architecture is trained using 90% of the data, and the final model is validated in the test dataset, composed of the remaining 10% of the data. Two SVM models were used to evaluate the CNN performance. The final results show that the CNN model outperform the SVM models with a total f1-score of 97%. The CNN has the disadvantage of a longer training time, but is able to learn relevant features from the input image, resulting in a better generalized model. Future work will focus on the utilization of CNN models in high resolution dual-pol SAR images, incorporating the polarimetric channels into the network convolutional layers.

Acknowledgment

TerraSAR-X data were provided by German Aerospace Center (DLR) via AO accounts OCE2917, OCE1045, OCE2082.

References

[1] Frédéric Bastien, Pascal Lamblin, Razvan Pascanu, James Bergstra, Ian J. Goodfellow, Arnaud Bergeron, Nicolas Bouchard, and Yoshua Bengio. Theano: new features and speed improvements. Deep Learning and Unsupervised Feature Learning NIPS 2012 Workshop, 2012.

[2] Yoshua Bengio. Learning deep architectures for AI. *Foundations and trends in Machine Learning*, 2(1):1–127, 2009.

[3] Carlos Bentes, Domenico Velotto, and Susanne Lehner. Target classification in oceanographic SAR images with deep neural networks: Architecture and initial results. In *IEEE International Geoscience and Remote Sensing Symposium (IGARSS) 2015*, page 4, 2015.

[4] James Bergstra, Olivier Breuleux, Frédéric Bastien, Pascal Lamblin, Razvan Pascanu, Guillaume Desjardins, Joseph Turian, David Warde-Farley, and Yoshua Bengio. Theano: a CPU and GPU math expression compiler. In *Proceedings of the Python for Scientific Computing Conference (SciPy)*, June 2010. Oral Presentation.

[5] Stephan Brusch, Susanne Lehner, Thomas Fritz, Matteo Soccorsi, Alexander Soloviev, and Bart van Schie. Ship surveillance with TerraSAR-X. *Geoscience and Remote Sensing, IEEE Transactions on*, 49(3):1092–1103, 2011.

[6] Michael Denbina, Michael J Collins, and Ghada Atteia. On the detection and discrimination of ships and icebergs using simulated dual-polarized radarsat constellation data. *Canadian Journal of Remote Sensing*, (just-accepted):00–00, 2015.

[7] Anja Frost, Rudolf Ressel, and Susanne Lehner. Iceberg detection over northern latitudes using high resolution TerraSAR-X images. In *36th Canadian Symposium of Remote Sensing-Abstracts*, pages 75–75, 2015.

[8] RS Gill. Operational detection of sea ice edges and icebergs using SAR. *Canadian journal of remote sensing*, 27(5):411–432, 2001.

[9] Carl Howell, James Youden, Kelley Lane, Desmond Power, Charles Randell, and Dean Flett. Iceberg and ship discrimination with ENVISAT multipolarization ASAR. In *Geoscience and Remote Sensing Symposium, 2004. IGARSS'04. Proceedings. 2004 IEEE International*, volume 1. IEEE, 2004.

[10] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.

[11] Yann LeCun and Yoshua Bengio. Convolutional networks for images, speech, and time series. *The handbook of brain theory and neural networks*, 3361:310, 1995.

[12] David AE Morgan. Deep convolutional neural networks for ATR from SAR imagery. In *SPIE Defense+ Security*, pages 94750F–94750F. International Society for Optics and Photonics, 2015.

[13] Desmond Power, James Youden, Kelley Lane, Charles Randell, and Dean Flett. Iceberg detection capabilities of RADARSAT synthetic aperture radar. *Canadian Journal of Remote Sensing*, 27(5):476–486, 2001.

[14] Rudolf Ressel, Anja Frost, and Susanne Lehner. Navigation assistance for ice-infested waters through automatic iceberg detection and ice classification based on TerraSAR-X imagery. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, 2015.