SIMULTANEOUS ESTIMATION OF MULTIPLE SHIP PARAMETERS FROM SAR IMAGES USING A FORKED CONVOLUTIONAL NEURAL NETWORK

J. Imber, B. Tings, D. Velotto

German Aerospace Centre (DLR)

ABSTRACT

The detection of ships and other objects at sea using space-bourne SAR is well established. The current focus is to extract as much additional information as possible from these detections. Convolutional neural networks (CNN) enable the empirical modelling of an arbitrary function relating an image patch to a data point in the desired parameter space. This paper describes the application of a forked CNN to the problem of ship parameter estimation using a compound loss function. In this way, the ship heading, dimensions and type are simultaneously estimated using a single network. The results suggest that this is an efficient method for generating an accurate model.

Index Terms— neural networks, synthetic aperture radar, object classification

1. INTRODUCTION

Space-bourne synthetic aperture radar (SAR) is capable of measuring large areas in all weather conditions, independent of natural illumination. As such, it has become indispensable for monitoring the world's oceans. The observation of metallic structures on the ocean surface using SAR is a common use of the technology and considered important for maintaining safety and security at sea. As large, angular, metallic objects, ships have a very large radar cross-section and are thus readily detected amongst the clutter of the ocean surface [1][2]. Oil platforms, offshore wind turbines and icebergs are similarly distinct, usually presenting a much larger radar cross-section than the surrounding environment [3][4]. Once detected, the next task is to determine the characteristics of the objects observed. Neural networks have been introduced to perform classification of maritime objects and shown excellent results [5][6]. In the study presented here, a neural network was trained to estimate the heading, length and width, and type, of a ship from small image patches extracted from the full SAR acquisition.

2. NETWORK IMPLEMENTATION

Large neural networks are currently amongst the most favoured tools for analysing complex datasets. Convolutional neural networks (CNN) are particularly suited to image analysis tasks. By restricting the spatial dimensions of the extracted features and constructing a sequence of layers, a feature hierarchy from simple edges to complex shapes can be produced. These features can then be used to discriminate between different input images. In this study the network was constructed in Python using the TensorFlow library with ancillary processing using the OpenCV and Numpy libraries [7][8][9]. Figure 1 shows a schematic of the network architecture. It was designed to be large enough to accomplish the tasks at hand, but small enough to remain trainable with the limited dataset available and within a reasonable time frame. Three convolutional layers are used to produce a set of features. These features are then fed into three separate fully-connected network paths, one for each output. Identification of the relevant image features is shared, and in this way the network can benefit from a simple form of curriculum learning [10]. The ship heading is the easiest problem for the network to solve and thus the initial training was performed using only the heading part of the loss function. The other components were then added and benefited from the approximate features already discovered.

The ship heading is a cyclical variable and thus difficult to extract via regression, as virtually identical observations can be paired with very different parameter values. Instead a classification approach has been shown to give better results [11]. A set of 8 bins are used to span the possible values [0,360]. The number of bins is kept small to reduce the fraction of edge cases and make the training easier. Additional sets of bins are then added, offset from the original by a fraction of the bin width. The network is trained against each set of bins independently, yet simultaneously, using a softmax function to give a probability for each bin at the output, and cross-entropy loss. In this study, nine sets of 8 bins were used, each offset by $5°$, as shown in the left-hand branch of the network architecture in Figure 1. Figure 2 shows two sets of bins as an example in red (solid) and blue (dashed). As discussed in [11] the heading can be extracted by constructing a probability density function and then finding the angle that maximises that function. Combining the results in this way, a high angular resolution is recovered.

The other outputs do not require additional processing. The ship dimension component was trained using simple re-
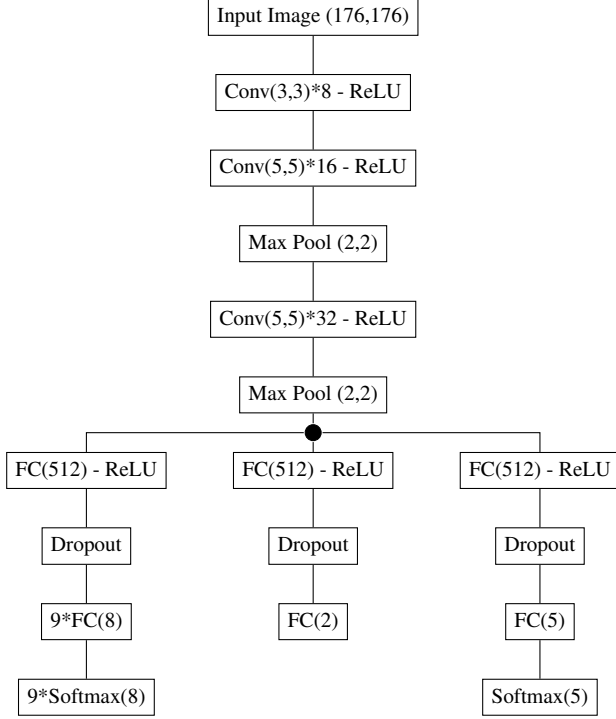
Fig. 1. A schematic of the CNN network architecture. The three branches return the heading (left), dimension (centre) and type (right) of the detected vessel.

gression with a Huber loss function to give the values in metres directly. The ship-type is another classification problem. The probability is again calculated using a softmax function at the output and trained using cross-entropy loss. For this work, ships were divided into broad type classes (Cargo, Tanker, Fishing, Passenger, Other). The number of classes was determined by the data available. Classes without sufficient representation in the data were grouped into "Other" rather than discarded. This gives the class set a degree of completeness required for operational deployment. The total loss function for the network is then the sum of the three individual loss functions.

$$L_{total} = L_{heading} + L_{dimension} + L_{type}$$

$$L_{heading} = \sum_{j=1}^{9} \sum_{i=1}^{8} p_{ij} \log q_{ij}$$

$$L_{dimension} = \begin{cases} 0.5 \times |x-y|^2, & |x-y| \leq 1 \\ |x-y| - 0.5, & |x-y| > 1 \end{cases}$$

$$L_{type} = \sum_{i=1}^{5} p_i \log q_i$$
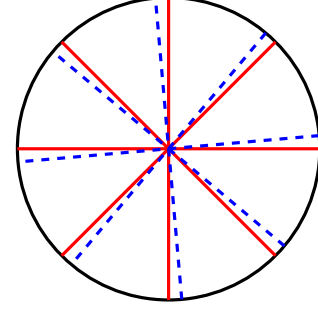
Fig. 2. An illustration of sets of 8 bins used for heading determination. Only two of the nine used are shown for clarity.

3. DATASET

The data used in this study is derived from TerraSAR-X (TS-X) stripmap multi-look ground range detected (MGD) images with VV and HH polarisation and pixel size between 2-4 m [12]. The acquisitions were made over the North Sea and the Mediterranean. Ships were detected and extracted from the SAR images using a constant false alarm rate (CFAR) algorithm [13][2]. The ground truth information used for training the network was derived from the automatic identification system (AIS). The AIS messages were automatically paired to the SAR ship detections by matching the latitude and longitude, and the time stamp, interpolating between subsequent AIS data points where necessary. Additional checks were performed by hand. Ships were extracted within small patches, resampled at 3 m pixel spacing and trimmed to squares of 176 by 176 pixels. The image patches were normalised via a logarithmic function to the range [-1,1] using the following procedure, where max and min values are found separately for each patch.

$$\sigma_0 \to a = \log(\sigma_0 + 1)$$
$$a \to b = 2 \left(\frac{a - \min(a)}{\max(a) - \min(a)} \right) - 1$$

In total, 7535 unique detections of ships were included in the data set. The data were augmented by a factor of four by performing horizontal and vertical mirroring operations on the original image patches. The data were divided into three parts; 70% for training the network, 15% to test the performance during training, and 15% for a final validation of the network performance. For the training data set, patches from the sub-dominant ship-type classes were duplicated such that each class was approximately equally represented. All of the results in §4 were produced using the validation data subset.

4. RESULTS AND DISCUSSION

From space, ships display a significant degree of fore-aft symmetry leading to 180° ambiguities in ship heading retrieval.

One of the key results of this paper is the ability of the CNN to break that degeneracy. Figure 3 shows the accuracy of the estimated heading. On a log-scale it is clear that the forward direction is favoured by approximately two orders of magnitude over the reverse. Preliminary investigation suggests that the network makes use of any wake pattern within the surrounding sea as well as features of the vessel itself.
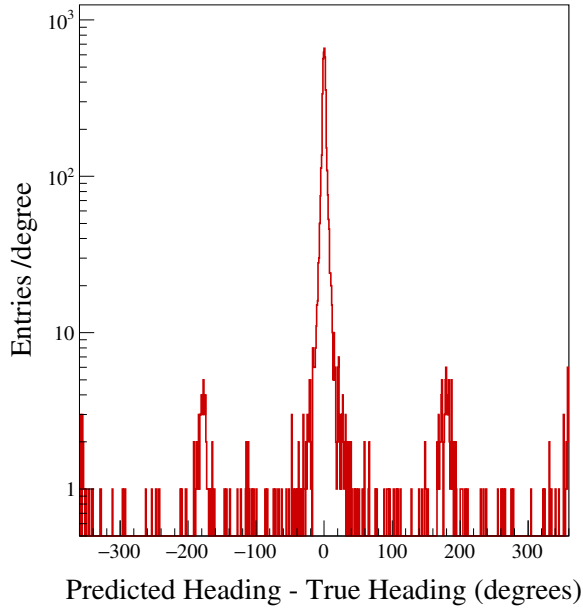

Fig. 3. A histogram showing the accuracy of the predicted ship heading.

Right-angles or corner geometries naturally produce extremely high radar return signals with large side lobes. Such imaging artefacts can substantially extend the apparent size of a ship [14]. The ship width, as the smaller dimension, is more susceptible to distortion in this manner. Figure 4 shows the fractional error in the estimated dimensions, $(x_{estimated} - x_{true})/x_{true}$. The plots show a slight bias towards underestimating the length and width, although in both cases it is well below the standard deviation. This may be a result of unevenness in the ship dimension distributions. Vessels over 200 m in length are comparatively rare in this dataset, and thus less well modelled by the network.

The results of the ship-type classification are shown in Figure 5 in terms of the F1-score, a metric combining precision and recall.

$$\text{F1-score} = 2 \left(\frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \right)$$

The diagonal elements are clearly dominant with the most significant confusion arising between the cargo and tanker classes. Whether it is a particular subset of these classes that
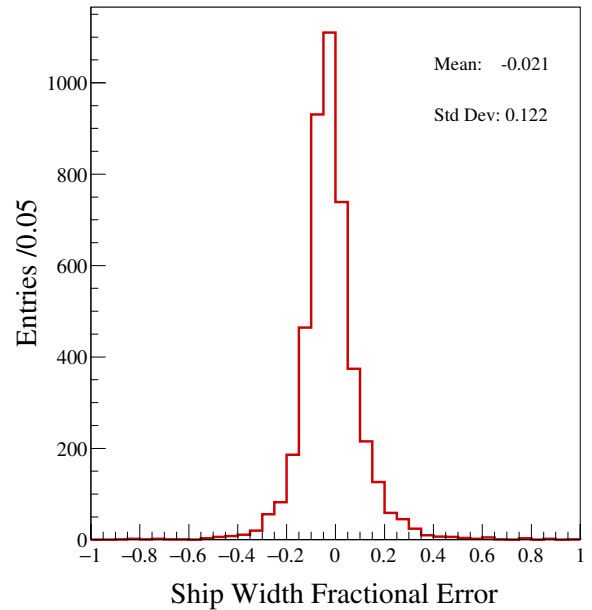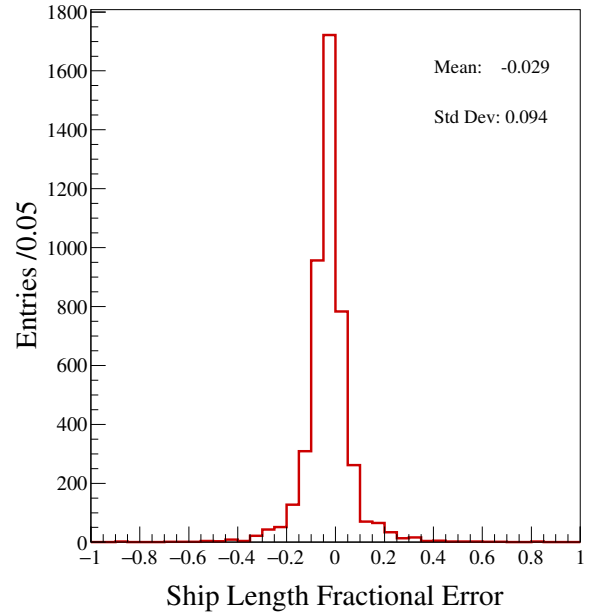
Fig. 4. Histograms of the estimated ship dimension fractional error for length (top) and width (bottom)

tend to be misidentified or if these observations were made under sub-optimal conditions is the subject of further study.
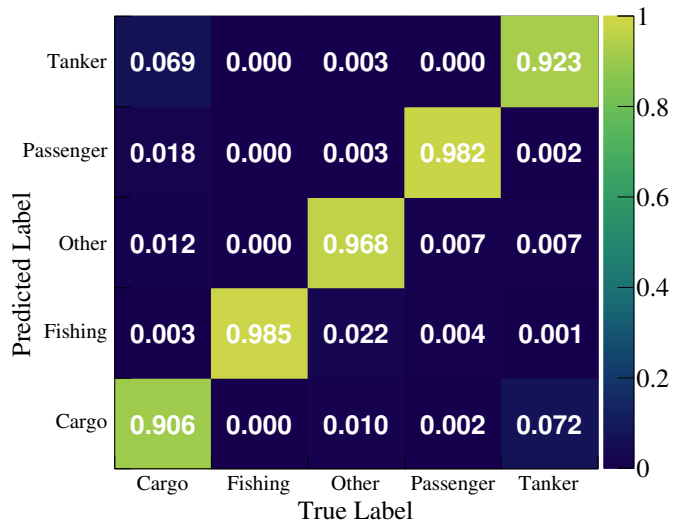
Fig. 5. The ship type estimation F1-score matrix. Note that bins with no significant figures are indeed empty.

This study suggests that a shallow CNN trained to model multiple, complementary parameters simultaneously, and with very limited preprocessing of the input images, can perform accurate ship parameter estimation.

5. REFERENCES

[1] S. Brusch, S. Lehner, T. Fritz, M. Soccorsi, A. Soloviev, and B. van Schie, "Ship surveillance with terrasar-x," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 3, pp. 1092–1103, March 2011.

[2] Björn Tings, Carlos Bentes, Domenico Velotto, and Sergey Voinov, "Modelling ship detectability depending on TerraSAR-X-derived metocean parameters," *CEAS Space Journal*, Oct 2018.

[3] H. Li and W. Perrie, "Detection of wind farm using the relative phase of compact polarimetry SAR," in *2012 IEEE International Geoscience and Remote Sensing Symposium*, July 2012, pp. 7636–7639.

[4] A. Marino, D. Velotto, and F. Nunziata, "Offshore Metallic Platforms Observation Using Dual-Polarimetric TS-X/TD-X Satellite Imagery: A Case Study in the Gulf of Mexico," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 10, no. 10, pp. 4376–4386, Oct 2017.

[5] C. Bentes, A. Frost, D. Velotto, and B. Tings, "Ship-Iceberg Discrimination with Convolutional Neural Net-

works in High Resolution SAR Images," in *Proceedings of EUSAR 2016: 11th European Conference on Synthetic Aperture Radar*, June 2016, pp. 1–4.

[6] Carlos Bentes, Domenico Velotto, and Bjorn Tings, "Ship Classification in TerraSAR-X Images With Convolutional Neural Networks," *IEEE Journal of Oceanic Engineering*, vol. PP, pp. 1–9, 11 2017.

[7] Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Yangqing Jia, Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dan Mané, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viégas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng, "TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems," 2015, Software available from tensorflow.org.

[8] G. Bradski, "The OpenCV Library," *Dr. Dobb's Journal of Software Tools*, 2000.

[9] Travis Oliphant, "NumPy: A guide to NumPy," USA: Trelgol Publishing, 2006–, (http://www.numpy.org/).

[10] Yoshua Bengio, Jérôme Louradour, Ronan Collobert, and Jason Weston, "Curriculum Learning," in *Proceedings of the 26th Annual International Conference on Machine Learning*, New York, NY, USA, 2009, ICML '09, pp. 41–48, ACM.

[11] K. Hara, R. Vemulapalli, and R. Chellappa, "Designing Deep Convolutional Neural Networks for Continuous Object Orientation Estimation.," arXiv preprint, 2017, arXiv:1702.01499.

[12] H. Breit, T. Fritz, U. Balss, M. Lachaise, A. Niedermeier, and M. Vonavka, "TerraSAR-X SAR Processing and Products," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 2, pp. 727–740, Feb 2010.

[13] David Crisp, "The state-of-the-art in ship detection in Synthetic Aperture Radar imagery," *DTIC Doc., Tech. Rep. DSTO-RR-0272*, 05 2004.

[14] Björn Tings and Carlos Augusto Bentes da Silva and Susanne Lehner, "Dynamically adapted ship parameter estimation using TerraSAR-X images," *International Journal of Remote Sensing*, vol. 37, no. 9, pp. 1990–2015, 2016.