Exploring Deep Transfer Learning Interference classification on Neural Style Transfer Generated Synthetic SAR Datasets

Richard Capraru[1,2], Matthew Ritchie[3], Jian-Gang Wang[1], and Boon Hee Soong[2]

[1]Institute for Infocomm Research, Agency for Science, Technology and Research, Singapore, Singapore
[2]Department of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, Singapore
[3]Department of Electrical and Electronic Engineering, University College London, London, United Kingdom
Email: {richard_capraru, jgwang}@i2r.a-star.edu.sg, ebhsoong@ntu.edu.sg, m.ritchie@ucl.ac.uk

***Abstract*—Synthetic Aperture Radar (SAR) imagery receives great attention in recent years due to the development of radar systems based on the use of active array antennas, with transmit/receive modules at each antenna element, enabling ultra-high resolution and uninterrupted surveillance over areas of interest. However, severe image distortion is a critical problem, and this is often a result of Radio Frequency Interference (RFI) and noise. Issues that arise from distortion include missing detection and inaccurate height maps.**

Furthermore, SAR images are particularly relevant in computer vision tasks such as classification and automatic target recognition (ATR). For such applications, access to comprehensive databases of SAR images as well as SAR images contaminated with RFI and noise is critical to enable the effective training and optimisation of classification algorithms and to provide a common baseline for benchmarking purposes. Given these challenges, the purpose of this paper is to show that Neural Style Transfer can be used to induce RFI and noise into SAR images (creating valuable datasets) and we can further classify the type of contamination using image classification techniques. The experimental results have verified the efficiency of our approach.

***Index Terms*—Synthetic Aperture Radar, Radio Frequency Interference, Transfer Learning, Neural Style Transfer**

I. Introduction

Synthetic Aperture Radar (SAR) has a very important role in remote sensing and monitoring applications due to its' capability of detecting, tracking, and imaging targets with high accuracy at a long range. Therefore, it is widely used for military and civilian purposes, such as wide-area surveillance, air defence and weapon control, remote sensing of the environment, and industrial automation.

A current disadvantage, in comparison to optical imagery, is that synthetic aperture radar data is much less common and more challenging to generate or access. As a result, the amount of SAR imagery training data containing RFI and noise available at the moment is limited, and gathering useful datasets for research and learning purposes is a relevant topic for different sectors such as industry or academia.

Neural style transfer has been mainly used in image stylization [1] and texture synthesis [2]. More recently, some work has been done in creating synthetic micro-doppler data for human activity recognition showing that neural style transfer has the ability to extract environmental factors (such as noise) more than any other synthetic dataset [3].

In this paper, we investigate the possibility of using neural style transfer ([4]) to extract SAR interference features and generate very realistic synthetic data. This proposed method adapted for our application extracts the effects directly from any interference data and transfers the specific features to the clean SAR data. This generates easily large quantities of realistic RFI datasets with high quality. These datasets can be used to augment measurement data in real-life scenarios.

Furthermore, we also propose to measure the classification accuracy of existing pre-trained CNNs in the context of interference classification using our dataset. Contrary to our expectations, even for visually-similar interference-contaminated SAR images, the classification accuracy is very high. However, we consider that when the data is more diverse, with more images available, the accuracy levels might drop slightly.

II. Style transfer and Interference Classification

A. Transfer Learning for Image Classification

Transfer learning is a deep learning method. Its' goal is to improve learning and classification by using a pre-trained Convolutional Neural Networks (CNNs), which has been used initially for a different classification task and so pre-trained with a different set of data. As presented in [5], transfer learning can be roughly divided into the following categories: transductive transfer learning, inductive transfer learning, and unsupervised transfer learning.

In our approach of classifying different types of SAR interference, we used deep networks previously trained with a large source dataset of optical images, fine-tuned the values of their internal parameters using a part of the available experimental radar data (the so-called 'target dataset'), and finally used the remaining radar data for validation.

We adopt inductive transfer learning in this paper. We provide a clearly labeled dataset of SAR interference-contaminated maps. As we will see in the results, the features from the different CNNs proved to be very effective for this application.

B. Neural Style Transfer for realistic synthetic database generation

Neural Style Transfer is a machine learning technique widely used in image processing. The working process of this technique defines 2 images: one style image and one content image. The 2 images are blended together in such a manner that the final output image looks like the content image but with a style painted in it. Further work is being performed until the output image takes the desired content statistics and the desired style statistics of the original images. The neural style transfer method used to create the dataset has been adapted to our task from the technique presented in [4]. Normally this technique takes 3 inputs: a style image, a content image and an input image, but for the purpose of our application, we used the same content and input image to obtain the desired output RFI SAR contaminated image.

This technique is based on a feature space originally designed to capture texture information [6]. Being placed at the top of the filter responses of neural networks, it uses these responses over the spatial extent of the feature maps, and the final product is a stationary, multi-scale representation of the input image, which captures its texture information but not the global arrangement. [4]

III. METHODOLOGY

A. Neural Transfer Learning Overview

Figure: 1 outlines the process of creating the SAR data set using the Neural Style Transfer method. The same overall process has been used for both the X-Band ICEYE SAR data and for the L-Band DLR SAR data

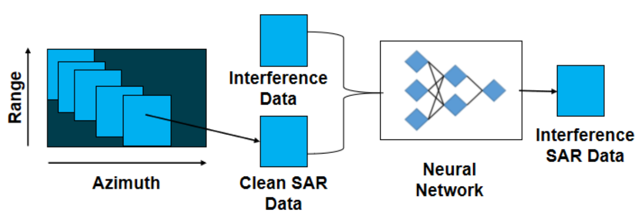

Fig. 1. Neural Style Transfer Process Diagram

Our original data is an SAR map of Range vs Azimuth. Because of the limited amount of data that this project had available, in order to produce more data, a code to automatically crop smaller maps out of a large map has been developed in order to save a good amount of images that could be used to create RFI data. This dataset of cropped SAR image will represent the Clean SAR Data in the Fig: 1.

The Interference data has been composed of 4 pure interference images taken from various research papers. The different types of interference used are: Narrow-Band Interference (NBI) [7], Chirp Modulated Wide-Band Interference (CMWBI) [7], Sinusoidal Modulated Wide-Band Interfernce (SMWBI) [7] and Time-Varying Wide-Band Interfernce (TVWBI) [8]. The ability of the Neural Style Transfer to use images from papers and create interference images provides the flexibility to be generate any kind of RFI SAR data using any type of interference and SAR data. This is particularly relevant when the amount of SAR RFI training data is limited, which is often the case.

After having both the style image (Interference Data) and the content image (Clean SAR Data), they must be resized to match the same size and small enough to be fed into the CNN. For our application, both the style image and the content image have been resized to 120x120 pixels

After defining the Content Loss function (which measures the dissimilarity between the content image and the output image) and the Style Loss function (which measures the dissimilarity between the style image and the output image) we will import the 19-layer VGG network, which has 16 convolutional and 5 pooling layers. (The reason for choosing this particular CNN is because it proved to have very good results in [4], and after some work on our application, it proved to display the desired output). Furthermore, we will normalize our images with mean=[0.485, 0.456, 0.406] and std=[0.229, 0.224, 0.225] before feeding them into the CNN, where std stands for standard deviation. The reason for this normalisation is that this is the kind of image VGG-19 has been previously trained with.

The final dataset of SAR RFI images has:

- 121 NBI L-band SAR images and 121 NBI X-band SAR images
- 121 SMWBI L-band SAR images and 121 SMWBI X-band SAR images
- 121 CMWBI L-band SAR images and 121 CMWBI X-band SAR images
- 121 TVWBI L-band SAR images and 121 TVWBI X-band SAR images

A good visualisation of the content image + style image and the resulting image can be seen in Figure: 2. The type of interference that was added is Time-Varying Radio Frequency Interference, and the original RFI image is from [8]. The output image looks very similar to real-life interference scenarios and can be used as a part of a dataset for different classification purposes.

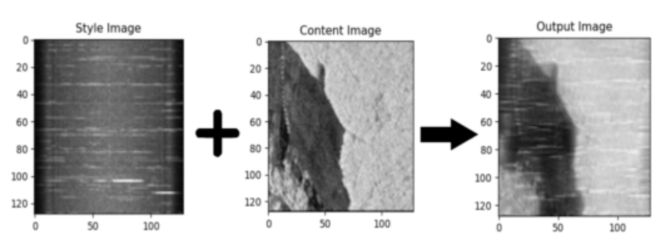

Fig. 2. Neural Style Transfer Process Example on DLR data

Figure: 3 provides a better visualisation to distinguish between the different types of RFI added. The first one is the CMWBI RFI, the second one is the NBI RFI which has a lower intensity, the third is the SMWBI RFI which has some similarities with the CMWBI due to the nature of the interference and the last one is the TVWBI RFI which is a more complex case.

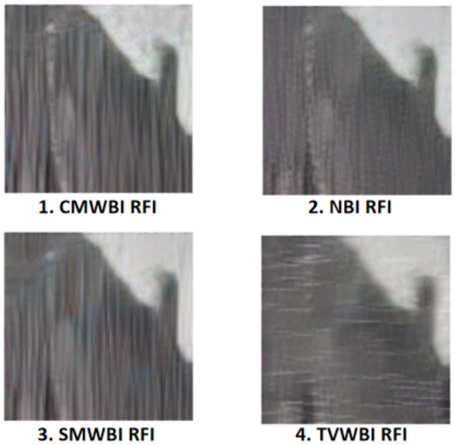

Fig. 3. RFI Types generated on the DLR F-SAR radar: 1. CMWBI 2. NBI 3. SMWBI and 4. TVWBI

B. Transfer Learning Overview

Several deep convolutional neural network architectures are used to classify SAR maps contaminated with RFI. The networks are all pre-trained with a large number of optical images and then fine-tuned using the training radar data. Although radar images are different from conventional optical images, this transfer learning approach enables to leverage of the very large training set of optical images (e.g. ImageNet [9]), as opposed to the amount of available experimental radar data which is inherently limited. The approach followed in this work included the following stages:

- We first process the data created using the Neural Style Transfer learning.
- The data is then augmented to avoid overfitting in the CNN, by randomly flipping the training images along the vertical axis and randomly translating them up to 30 pixels as well as scaling them up to 10% horizontally and vertically
- The images are fed into the CNNs for feature extraction. The trained CNNs are AlexNet [10], VGG16 [11], GoogLeNet [12] and ResNet50 [13]. All of these networks are pre-trained on the ImageNet dataset [9], which enabled the rapid progression of image classification processing and is particularly useful in this radar application.
- Lastly, the scores obtained from the CNNs are fed into a Softmax classifier which gives us the results available in Section: IV and Table: III, the classification accuracy is validated using a testing dataset.

C. AlexNet

AlexNet [10] was the first CNN that was trained on the dataset. The process of this CNN works the following way: the images are modified to the size of 227x227 RGB, and the images are fed to the neural network which is composed of 8 layers, 3 of which are fully connected layers and 5 are convolutional (each being followed by Relu functions). There are also present Max Pooling layers. The deeper the convolutional layers, the more advanced will be the features that AlexNet will use for the classification. To avoid overfitting, there are also present two 50% Drop Out layers. Lastly, the raw scores obtained are fed into a 1000-way Softmax classifier, which uses the cross-entropy loss function.

D. VGG16

VGG-16 is a very well-known CNN developed at the Visual Geometry Group (VGG) [11]. The required images for VGG-16 are 224x224 RGB format, and the network consists of a total of 41 layers. The number of convolutional layers is 16. All the convolutional layers are again followed by Relu functions, and some are followed by Max Pooling layers (2x2). To avoid the overfitting problem, at the end there are two 50% Drop Out layers. Lastly, the raw scores are again fed into a Softmax classifier, which gives us the percent of accuracy.

E. ResNet-50

ResNet-50 is a down-scaled version of the network VGG16 which utilises the same block structure. [13] It also requires 224x224 RGB format images, just like VGG16, but the advantage of using ResNet50 is that it allows an increased classification of data with very similar features, which is the case for radar images. This is due to the fact that, while in VGG subsequent blocks learn the features from the images anew, from the output of the preceding block, in ResNet50 those subsequent blocks only learn the residual of the output from the previous block.

It has overall 177 layers of which 50 are convolutional. Many convolutional layers are followed by Relu functions and Batch Normalization layers. There are also Max Pooling layers (3x3). The final 2 layers include an Average Pooling layer as well as 1000-way Softmax classifier.

F. GoogleNet

Since this CNN was designed considering mainly the computing resources for embedded systems, it makes it the perfect pre-trained network considering future mobile applications for RFI classification. [12] The required images for GoogLeNet are 256x256 RGB format. What makes this CNN powerful is the fact that it uses inception modules in 22 layers while requesting a good computing budget. These inception modules can enable the extraction of more features from the radar maps, increasing the accuracy, without making the network vulnerable to overfitting issues. There are 7 inception modules, each consisting of 3 convolutional layers, and a Max pooling Layer. It has overall 144 layers, of which 22 are convolutional.

IV. Data and Results

A. ICEYE data

The ICEYE's satellites' orbit height at the equator is 570 km, the inclination is 97.69 degrees and the number of orbits per day is 15. The SAR module used for recording the data is ICEYE-X2, an X-band Radar. The relevant parameters of this module are presented in Table I. [14]

TABLE I
ICEYE SATELLITE RADAR PARAMETERS. [14]

Feature	ICEYE Satellite
Center Frequency[GHz]	9.6
Imaging and Polarization	X-band SAR, VV polarization
Resolution (m)	1x1 / 3x3 / 20x20
Dynamic Range	16 bit
Imaging Mode	Stripmap, Spotlight
Communications (downlink)	X-band radio, 100+ Mbits/sChirp BW
Chirp Bandwidth	Up to 300 MHz
Georeferencing	Under 10 m (both azimuth and range)

The SAR ICEYE satellite has good radiometric accuracy, smaller than 1 dB for Stripmap data. The orbit accuracy levels are around 500 meters precision for image planning, and emergency services and around 3 meters for standard delivery along with the product. [14]

B. DLR data

Another dataset used as a part of this paper comes from an L-band radar developed by DLR Microwaves and Radar Institute in Germany. The L-band radar used on this project comes from the F-SAR. It is designed to cover an off-nadir angle range of 25 to 55 degrees at altitudes of up to 6000m above sea level, which is the maximum operating altitude with the DO228. The relevant parameters of this module are presented in Table II. [15]

TABLE II
DLR RADAR PARAMETERS. [15]

Feature	DLR Plane
Center Frequency[GHz]	1.325
Bandwidth [MHz]	150
Range resolution[m]	1.5
Azimuth resolution [m]	0.4
Range covered[km]	12.5 (at max. bandwidth)
Sampling	8 Bit real
Data rate	247 MByte/s (max. per rec channel)

C. Classification Results

Multiple CNNs were trained on the L-band dataset using 80% of the data for training and 20% data selected randomly for validation. The learning rate was set to 3e-4. The comparison was made between the 4 types of generated RFI Data: NBI, CMWBI, SMWBI as well as TVWBI. Table III contains the accuracy results, as well as other parameters for each CNN trained on the DLR F-SAR radar, generated data. The total number of samples used for this classification is 484.

In Figure: 4 we have a plot of validation accuracy of GoogLeNet. The blue solid line represents the training accuracy graph, while the black dotted line represents the validation graph. In Figure: 5 we have a plot of the loss graph of GoogLeNet. The orange solid line represents the training loss graph, while the black dotted line represents the validation loss

TABLE III
CNNS RESULTS AND PERFORMANCE METRICS

CNN	Acc.	Param(Mil)	Size
AlexNet	100%	61.0	227 MB
GoogLeNet	100%	7.0	27 MB
ResNet-50	100%	25.6	96 MB
VGG-16	100%	138	515 MB

graph. The number of iterations per epoch is 38 (the same number as it is for the other CNNs as well). The number of epochs used is 6, making it a total of 228 iterations. By looking at the loss graph it can be noticed that the validation loss decreases gradually with each iteration and the accuracy improves. It is also very important to notice that the validation accuracy and the training accuracy are very close to each other, successfully avoiding the overfitting problem.

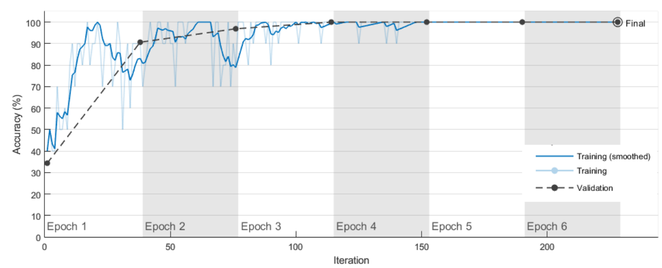

Fig. 4. Validation accuracy of GoogLeNet

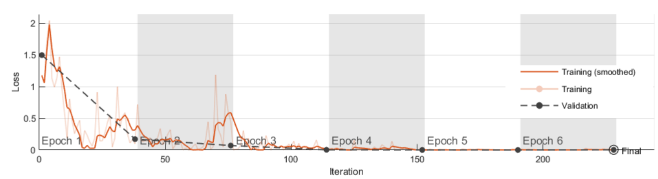

Fig. 5. Loss graph of GoogLeNet

Comparing the hyperparameters as well as the size of each CNN, it can be seen that GoogLeNet has the smallest number of parameters, 7.0 million, and a size of 27 MB making it easier to implement in constrained computational environments for future applications of SAR classification. In future applications, the generated RFI SAR data can be added to real-life data to create powerful datasets for classification applications as well as automatic target recognition.

To provide another interesting example, Figure: 6 shows the accuracy graph of the AlexNet CNN trained on NBI vs CMWBI vs SMWBI with a total of 363 samples. The classification accuracy this time is 98.61%, which is still very high. It is important to highlight that, for the classification, the data that creates false positives or false negatives are the CMWBI vs SMWBI due to their similar nature. However, the amount of false positives and false negatives are relatively low and can be ignored.

The same CNNs have also been trained on the X-band dataset using 80% of the data for training and 20% data

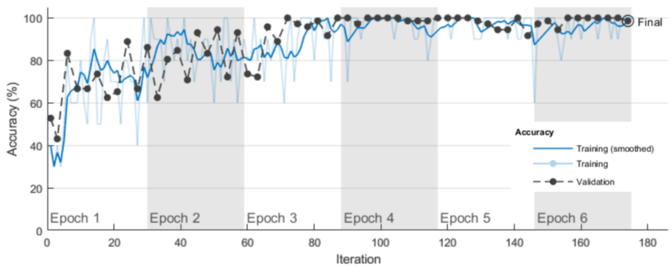

Fig. 6. Accuracy graph of AlexNet

selected randomly for validation. The learning rate was once again set to 3e-4. The comparison was made between the 4 types of generated RFI Data: NBI, CMWBI, SMWBI as well as TVWBI added to the X-band data this time. The accuracy results of each CNN trained on the ICEYE SAR radar generated data are the following: AlexNet (100%), GoogLeNet (100%), VGG-16 (100%), ResNet-50 (100%). The total number of samples used for this classification is 484.

In this section we proposed using pre-trained CNN architectures such as AlexNet [10], VGG16 [11], GoogLeNet [12] and ResNet [13] on the application of SAR RFI classification using a dataset created using Neural Style Transfer. This is an example of transfer learning from the source domain of optical images to the target domain of radar SAR images, which can be treated as 2D matrices of pixels values. The results showed very high classification accuracy. Of course, by adding more real life-examples of RFI into the data, the classification algorithms may require further optimisation.

Furthermore, this comparison provides useful insights on how RFI could be classified using state-of-the-art pre-trained CNNs. The results show that CNNs such as GoogleNet [12], which are designed considering the limited computing resources, still maintain very high accuracy. This can provide possibilities for high-accuracy, relatively low memory, classifiers which can be implemented for choosing the right RFI filtering method in real-life scenarios.

V. Conclusion

This work explores the possibility of using the style transfer method to synthesize realistic RFI contaminated SAR images which can be used for different computer vision tasks such as classification and automatic target recognition (ATR). The resulted synthetic RFI SAR images demonstrate that Neural Style Transfer has the ability to extract well RFI from images and induce it into measurement data. We further apply image classification techniques to demonstrate that pre-trained CNN architectures can be used to efficiently classify different types of interference presented in SAR images. The experimental results on the two generated datasets (based on DLR and ICEYE databases) have shown that our approach can achieve high image classification accuracy.

For future work, the authors may consider generating more such data and sharing the dataset as a benchmark for research purposes as well as possibly assessing the performance of CNNs trained only with measurement RFI samples versus CNNs trained with a combination of measurement RFI samples and Neural Style Transfer generated samples. Since the classification performance might differ in such experiments, the CNNs' comparison would become more meaningful.

Acknowledgment

The authors would like to acknowledge the support of the UCL Radar group for making this project possible and the support of ICEYE for facilitating access to their SAR X-band data. The first author would also like to gratefully acknowledge the research centre CISS, School of EEE, NTU for their partial fund support.

Furthermore, the authors would like to thank the anonymous reviewers for his/her valuable comments.

References

[1] A. Sanakoyeu, D. Kotovenko, S. Lang, and B. Ommer, "A style-aware content loss for real-time hd style transfer," 2018. [Online]. Available: https://arxiv.org/abs/1807.10201

[2] M. Elad and P. Milanfar, "Style transfer via texture synthesis," *IEEE Transactions on Image Processing*, vol. 26, no. 5, pp. 2338–2351, 2017.

[3] S. Vishwakarma, W. Li, C. Tang, K. Woodbridge, R. Adve, and K. Chetty, "Neural style transfer enhanced training support for human activity recognition," 2021. [Online]. Available: https://arxiv.org/abs/2107.12821

[4] L. A. Gatys, A. S. Ecker, and M. Bethge, "A neural algorithm of artistic style," 2015. [Online]. Available: https://arxiv.org/abs/1508.06576

[5] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on Knowledge and Data Engineering, vol. 22, no. 10, pp. 1345-1359,*, Oct 2010.

[6] L. A. Gatys, A. S. Ecker, and M. Bethge, "Texture synthesis using convolutional neural networks," 2015.

[7] Y. Huang, L. Zhang, J. Li, W. Hong, and A. Nehorai, "A novel tensor technique for simultaneous narrowband and wideband interference suppression on single-channel sar system," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 12, pp. 9575–9588, 2019.

[8] F. J. Meyer, J. B. Nicoll, and A. P. Doulgeris, "Correction and characterization of radio frequency interference signatures in l-band synthetic aperture radar data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 10, pp. 4961–4972, 2013.

[9] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A Large-Scale Hierarchical Image Database," in *CVPR09*, 2009.

[10] G. E. H. A. Krizhevsky, I. Sutskever, "Imagenet classification with deep convolutional neural network," *NIPS'12 Proceedings of the 25th International Conference on Neural Information Processing Systems, vol.1, pp. 1097-1105,*, December 2012.

[11] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, cite arxiv:1409.1556. [Online]. Available: http://arxiv.org/abs/1409.1556

[12] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 1–9.

[13] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778.

[14] V. Ignatenko, P. Laurila, A. Radius, L. Lamentowski, O. Antropov, and D. Muff, "Iceye microsatellite sar constellation status update: Evaluation of first commercial imaging modes," in *IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium*, 2020, pp. 3581–3584.

[15] "Dlr website," accessed: 2022-05-24. [Online]. Available: https://www.dlr.de/hr/en/desktopdefault.aspx/tabid-2326/3776_read-5691/