A Deep Learning SAR Target Classification Experiment on MSTAR Dataset

Conference Paper · March 2018

DOI: 10.23919/IRS.2018.8448048

2 authors:

Rene Thaens
NATO - North Atlantic Treaty Organisation

Cristian I. Coman
NATO - North Atlantic Treaty Organisation

A Deep Learning SAR Target Classification Experiment on MSTAR Dataset

Cristian Coman, René Thaens

NATO Communications and Information Agency
The Hague, THE NETHERLANDS

email: cristian.coman@ncia.nato.int
email: rene.thaens@ncia.nato.int

Abstract: *Yet another deep learning method is proposed in this paper for the problem of automated target recognition (ATR) in Synthetic Aperture Radar (SAR) images. Deep convolutional neural networks (CNN) classifiers have been demonstrated on benchmark datasets like MSTAR to outperform classical machine learning algorithms based on formal feature extraction. Most of these deep learning solutions use amplitude information only and reverse the SAR classification problem to a simple image classification task. In this paper we analyze the potential of using additional radar information, such as phase information, in the deep learning process.*

1. Motivation

Images generated from Synthetic Aperture Radar (SAR) observations are commonly used in the Joint Intelligence, Surveillance and Reconnaissance (JISR) domain to develop the so called Recognized Ground Picture (RGP). Analysts are exploiting raw SAR data to respond to intelligence requirements, which often require identification of targets of interest or significant activities. The exploitation of SAR images is a complex process supported by trained image analysts, who manually search and classify targets extending just a couple of meters in large SAR images covering tens of kilometers. The time required to complete this manual analytic task is significant and impacts on the performance of the JISR community to respond to time sensitive requests. Developing and maintaining experiences SAR image analysts is a lengthly process, which can take years of training and significant resources.

Beside being a time consuming and expensive process, SAR image exploitation poses another challenge, which is the poor quality (low resolution) of images. Physical and technical limitations in radar technologies restrict to image resolutions to tens of centimeters or meters per pixel. When observing an armored vehicle, which extends a few meters the numbers of pixels is limited to a few tens as well. The image is more or less a collection of scatters and often even highly train analysts have difficulties with the detection and classification of targets in SAR images through visually guessing the shape of the target.

Researchers in the deep Convolution Neural Networks (CNN) domain have recently adopted SAR automated target classification (ATR) as one of the benchmark problems for highlighting

978-3-7369-9545-1

the potential of these new methods in the radar domain. In [1] and [2] a five layer CNN is reported to achieve a significant improvement in the classification accuracy when compared with popular machine learning methods such as Support Vector Machines (SVM) or Adaptive boosting (AdaBoost) used for SAR ATR. A regularization technique is proposed in [5] to reduce the free parameters of CNN model and consequently reduce the training time and improve the convergence properties. The problem of lack of SAR training data is addressed in [6] where a multilayer auto-encoder (AE) is used to prevent the over-fitting caused by training on restricted datasets. Most of the solutions proposed in the literature are exclusively using the magnitude information from SAR images and disregard other radar data, such as the phase of the signal.

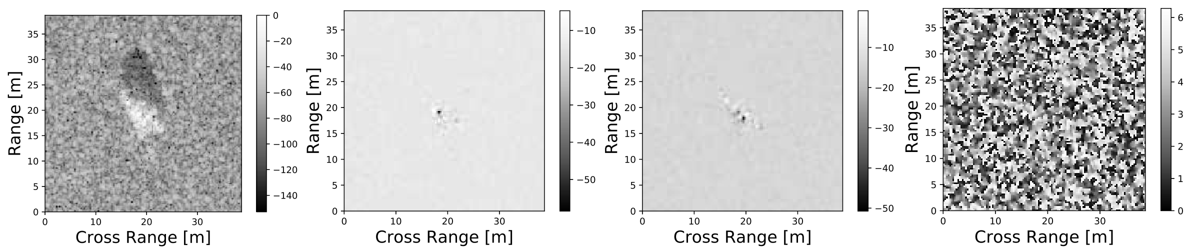
In this paper we analyze an simple solution for automated SAR image target recognition, which includes phase information into the deep learning training process. The phase provides additional discriminatory information, which increases the classification accuracy and resolve some of the challenges related to over-fitting and training on reduce datasets.

The remainder of this paper is organized as follows. Section 2 introduces the SAR data available for the experiment and explains how the phase information is exploited in the classification process. The architecture of the convolution neural network is presented in Section 3 where details of the training approach is discussed as well. Experimental results are presented in Section 4 and Section 5 contains concluding remarks.

2. Synthetic Aperture Radar Data

Airborne radars used in Earth observation missions employ range and azimuth compression techniques to generate high resolution images of observed scenes. The images constructed from microwave radar signals are expressed in the complex domain by using amplitude and phase information. The amplitude information gives and indication of the radar reflectivity and is commonly presented like a black and white picture. The phase information is more difficult to be interpreted by an analyst and often is not shown on the radar display. Still the phase contains information about the targets and through additional feature extraction techniques this information can be utilized for detecting and classifying target in SAR data. For example, interferometric processing of phase information is a popular technique for detecting changes in SAR images taken at different moments in time.

The Moving and Stationary Target Acquisition and Recognition (MSTAR) [7] collected by Sandia National Laboratory in a project jointly sponsored by the Defense Advanced Research Projects Agency and the Air Force Research Laboratory has been used in the experiment discussed in this paper. The data is available on the Internet and contains SAR images of ten targets including tanks, armored vehicles, weapons systems and military engineer vehicles (armored personnel carrier: BMP-2, BRDM-2, BTR-60, and BTR-70; tank: T-62, T-72; weapon system: 2S1; air defense unit: ZSU-234; truck: ZIL-131; bulldozer: D7). The data was collected with a Sandia X-band radar operating at 9.60GHz with a bandwidth of 0.591GHz. The range and cross range resolution are identical and equal to 0.3047m. Each target file contains a header



(a) (b) (c) (d)

Figure 1: MSTAR chip images for target ”t72 tank”: (a) normalized amplitude [dB], (b) normalized real [dB], (c) normalized imaginary [dB] and (d) phase information [rad].

where detailed sensor and sensor-target information is provided. Following the header the SAR data is provided, first the amplitude information then the phase information. The target data is presented as SAR chips containing between 128 and 179 rows and columns. For purpose of this paper all the data was normalized to block of 88 rows by 88 columns by taking the central part of the SAR chip provided in MSTAR data set. The amplitude and phase information for the target ”t72 tank” are depicted in Fig. 1, (a) and (d) respectively. The target can be easily observed in the amplitude image. The phase information does not present easily observable discriminating features but was used to generate the real (Fig. 1, (b)) and the imaginary (Fig. 1, (c)) parts of the original complex image. The real and imaginary parts were obtained by multiplying the normalized (0 to 1) amplitude data with the cosine and sine functions of the phase information. The data was further normalized by translation into the positive domain of values and re-scaling (0 to 1). This normalization process facilitates the representation of data in dB values in Fig. 1. The use of real and imaginary parts of the signal helps with the augmentation of the information contained in the training data set. CNNs used in visual image recognition often employ operations such as dimension reduction and translations and rotations to increase the training data set. For SAR classification problems the phase data can also be used as a simple technique to extend the information contain in limited training data sets. In the data analysis phase of the experiment reported in this paper it was possible to look into the Fourier transform of the complex SAR image as well. Classification based on images in the frequency domain is yet another way radar information can be employed in deep neural networks.

3. Network Model

The scope of the work reported in this paper is limited to exploration of potential benefits of phase information in SAR classification problems. One of the idea is to compare the performance of amplitude only and amplitude and phase SAR classification methods. For this purpose, two similar networks have been tested in the experiment. The networks are composed of two convolution layers, a flatten layer, two dense layers, and two dropout layers The parameters of the networks are presented in Tab. 1. The only difference between the two networks is in the shape of the input of the first convolution layer: the amplitude network receives a 2D image with the size of 88×88, whereas the amplitude and phase network takes images with the shape

Table 1: Parameters of the deep CNNs used in SAR classification.

layer (type)	input shape	kernel	features	output shape	param
input (Conv2D)	$88 \times 88 (88 \times 88 \times 3)$	3×3	32	$86 \times 86 \times 32$	320(896)
conv_1 (Conv2D)	$86 \times 86 \times 32$	3×3	32	$84 \times 84 \times 32$	9248
pool_1 (MaxPool)	$84 \times 84 \times 32$	2×2	32	$42 \times 42 \times 32$	0
drop_1 (Dropout)	$42 \times 42 \times 32$	0.25	32	$42 \times 42 \times 32$	0
flat_1 (Flatten)	$42 \times 42 \times 32$	1	56448	56448	0
dense_1 (Dense)	56448	1	128	128	7225472
drop_2 (Dropout)	128	0.25		128	0
dense_2 (Dense)	128	1	10	10	1290

$88 \times 88 \times 3$. This difference is also reflected in the number of parameters to be trained in the input layer: 320 for the amplitude only and 896 for amplitude and phase network. All layers use rectified linear units (ReLu) activation functions, with the exception of the last dense layer where a softmax activation function is used.

The input images for the amplitude only network simply represent the SAR absolute value information. For the amplitude and phase network the input images are constructed with three layers: the first layer is the amplitude data, the second layer is the real data and the third layer is the imaginary part (see Fig. 1 (a), (b) and (c)).

4. Experimental Results

The public MSTAR data set is already partitioned on training and test subsets for all ten classes of targets. Images for training are captured at a depression angle of $17°$, and images for testing are acquired at $15°$ depression angle. Test and training images are collected over the full aspect angle range ($0°$ to $360°$). The train/test partition based on depressions angle suggests concerns with the classifier when images are acquired by a platform flying at different altitudes. The range and cross-range resolution in SAR images do not depend on the target sensor geometry (they are related to the signal bandwidth and length of the synthetic aperture) but the shadow of objects is the main feature, which depend on the depression angle. The complete set of training images (3671 in total for all ten classes) and test images (3203 for all ten classes) is used in this paper. A mosaic of images (five samples per class) used for training of the amplitude only network is depicted in Fig. 2.

The networks were trained for 40 epochs with a batch size of 32 samples. The mean value of the correct classification accuracy on the test dataset, across all ten classes, was slightly better for the case when the phase was used, 91% compared to 90% in case of the amplitude only network. Such a small increase in the classification accuracy can be explained by relatively limited amount on discriminatory information contained in the phase part of the MSTAR dataset.

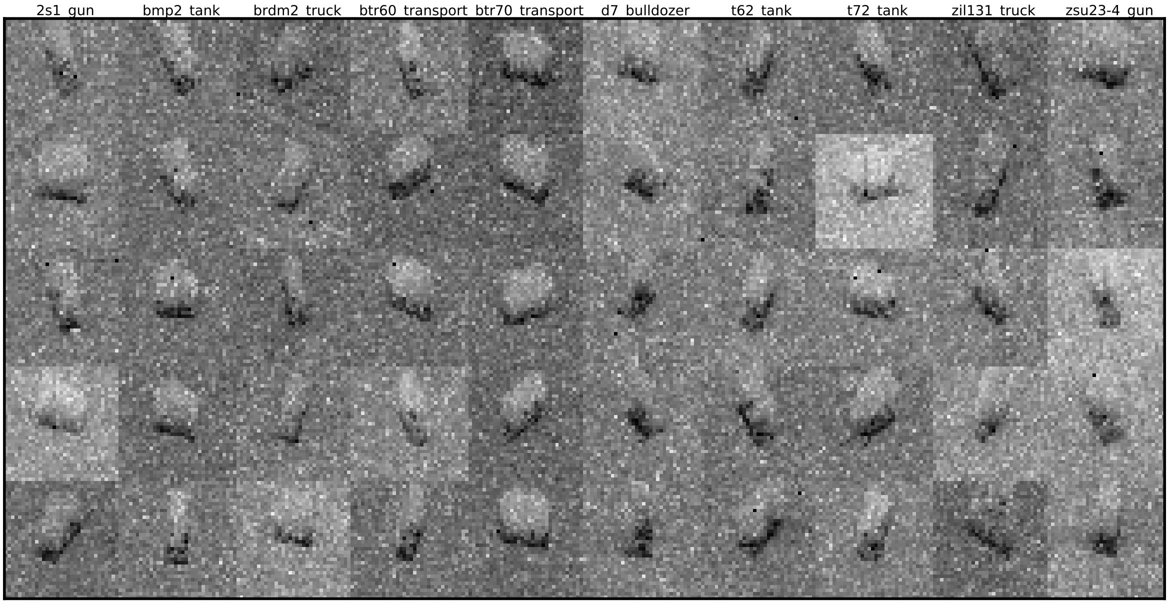
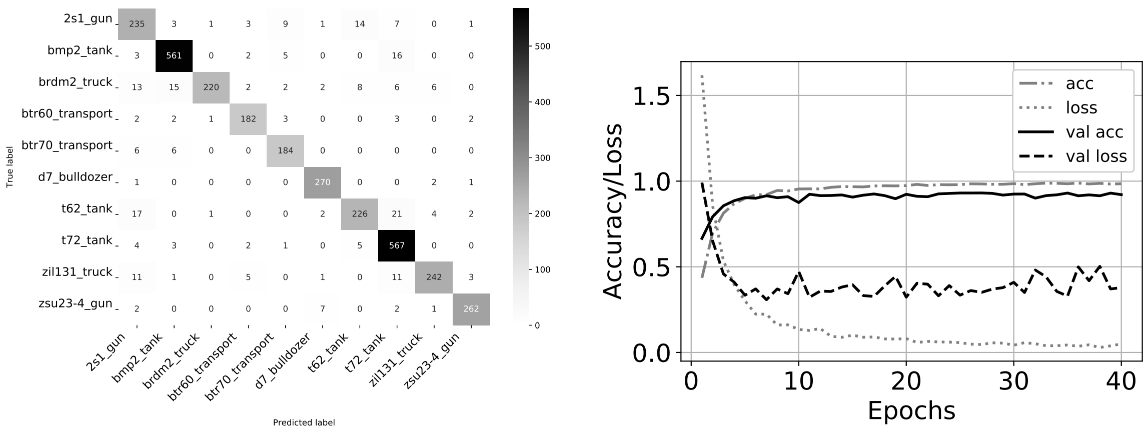


Figure 2: Mosaic of SAR image samples - amplitude information.

Common SAR focusing algorithms do preserve discriminant information in the phase component, however for the MSTAR dataset it is not evident how the focusing was performed and it is difficult to estimate why target information is not fully reflected in the phase data.

The confusion matrix of the amplitude and phase network is depicted in Fig. 3 (a), with the horizontal axis representing the predicted class labels and the vertical axis the true labels.



(a) (b)

Figure 3: Amplitude and phase network metrics: (a) confusion matrix, (b) accuracy and loss.

The training of the networks converged fast and after about 10 epochs the validation accuracy on the train data (acc) and on the test data (val acc) already reaches about 90% (see Fig. 3 (b)).

5. Conclusions

Deep convolutional neural networks show good results when used in automatic SAR target recognition problems. SAR phase information has the potential to improve the classification accuracy, especially when the focusing algorithm preserves target information in the phase data. The phase information can also be used to enhance the sample dataset when a limited number of samples is causing over-fitting during the training phase. In this paper a simple methods was introduced for employing phase information in the CNN process, namely the use of real and imaginary data to generate multichannel images.

Two lines of efforts could be identified to further investigate the potential use of deep CNN in SAR target recognition in production systems: developments in the CNN domain such as the use of ensemble models, capsule networks, or transfer learning techniques and the enhancement of SAR data set used in the CNN training (e.g. spectral images, polarimetry images etc.). Reliable CNN classification techniques are essential for the automation of the production of the Recognized Ground Picture (RGP), in particular in the exploitation SAR, electro-optic, infrared and full motion video data.

References

[1] H. Wang, S. Chen, F. Xu, and Ya-Qiu Jin, "Application of deep-learning algorithms to MSTAR data", *Proc. of 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*,2015, pp. 3743-3745.

[2] S. Chen, H. Wang, F. Xu, and Ya-Qiu Jin, "Target Classification Using the Deep Convolutional Networks for SAR Images", *IEEE Transactions on Geoscience and Remote Sensing*,vol. 54, no. 8, AUGUST 2016, pp. 4806-4817.

[3] Q. Zhao and J. C. Principe, "Support vector machines for SAR automatic target recognition", *IEEE Trans. Aerosp. Electron. Syst.*, vol. 37, no. 2, pp. 643654, Apr. 2001.

[4] Y. J. Sun, Z. P. Liu, S. Todorovic, and J. Li, "Adaptive boosting for SAR automatic target recognition", *IEEE Trans. Aerosp. Electron. Syst.*, vol. 43, no. 1, pp. 112125, Jan. 2007.

[5] S. Wagner, K. Barth, and S. Brggenwirth, "A deep learning SAR ATR system using regularization and prioritized classes", *Proc. of 2010 2017 IEEE Radar Conference (RadarConf)*, 2017, pp. 0772-0777

[6] S. Deng, L. Du, C. Li, J. Ding, and H. Liu, "SAR Automatic Target Recognition Based on Euclidean Distance Restricted Autoencoder", *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*,vol. 10, no. 7, 2017, pp. 3323-3333.

[7] E. R. Keydel, S. W. Lee, and J. T. Moore, "MSTAR extended operating conditions: A tutorial", *in Proc. of 3rd SPIE Conf. Algorithms SAR Imagerytext*, 1996, vol. 2757, pp. 228242.