AN ADVANCED NEURAL NETWORK-BASED APPROACH FOR MILITARY GROUND VEHICLE RECOGNITION IN SAR AERIAL IMAGERY

Victor-Emil NEAGOE, Serban-Vasile CARATA, Adrian-Dumitru CIOTEC

”Politehnica” University of Bucharest, Bucharest, Romania (victor.neagoe@upb.ro)

DOI: 10.19062/2247-3173.2016.18.1.5

Abstract: *The paper presents a novel neural network approach for automatic target recognition (ATR) in the synthetic aperture radar (SAR) aerial imagery; this is applied to identify military ground vehicles. The proposed ATR algorithm consists of a processing cascade with the following stages: (a) object detection using a pulse-coupled neural network (PCNN) segmentation module; (b) a first feature selection module using Gabor filtering (GF); (c) a second feature selection module using principal component analysis (PCA); (d) a support vector machine (SVM) classifier improved by using virtual training data generation (VTDG) with concurrent self-organization maps (CSOM). The proposed model has been applied for the recognition of three classes of military ground vehicles of the former Soviet Union represented by the set of 2987 images of the MSTAR public release database. The experimental results have confirmed the method effectiveness, leading to a total success rate of 97.36%.*

Keywords: *Automatic target recognition (ATR), aerial SAR imagery; military ground vehicles; pulse-coupled neural network (PCNN) segmentation, support vector machine (SVM), concurrent self-organizing maps (CSOM).*

1. INTRODUCTION

In peace or war, the fate of millions can depend on the analysis of images gathered by radar or other imaging sensors. These images might reveal a significant truth, like the 1962 U-2 pictures of Soviet missiles in Cuba [19]. In military intelligence and military reconnaissance, knowledge is power and one way to get to know your opponent is to watch him from the air or from space. But it is not so simple. Enemy countermeasures, natural atmospheric disturbances and daunting technical challenges complicate the task. Even when you are watching your enemy under the best conditions, you may not understand what you are looking at. The defense image interpreters are usually concerned with images gathered from aircraft or satellites and have to be able to recognize objects and interpret their meaning [19]. The recent advances try to substitute the human image interpreters by *Automatic Target Recognition (ATR)* based on artificial intelligence. Target objects are often military vehicles as those shown in Fig. 1 considered for the MSTAR public release database used to experiment our ATR proposed method.

(a) (b) (c)

FIG. 1. Three military vehicles of the former Soviet Union: (a) BMP2 (infantry fighting vehicle); (b) BTR70 (armored personnel carrier); (c) T72 (tank).

Radar imagery brings the advantage of independence from a passive illumination source, such as sunlight or starlight, thus offers imaging capability at night and through clouds. Modern day radar imaging systems are capable of comparatively high resolution by using synthetic aperture radar (SAR) imagery [2]. The area of Automatic Target Recognition (ATR) for SAR imagery is an ongoing research in many branches of the military and large research institutions [6], [7], [16], [17], [20]. On the other side, there has been an increasing interest in using artificial neural networks (ANN) for image processing and pattern recognition [1], [2], [7], [11], [12], [13], [14], [15], [17]. A typical target recognition system consists of a detection module (filtering and segmentation) and a recognition module (feature selection and classification) [1]. Moreover, the speckle noise specific to SAR images makes segmentation and recognition difficult tasks [2], [6]. We further propose application and development of new and challenging neural network models both for target detection (filtering and segmentation) and also for classification tasks. *The segmentation uses the pulse-coupled neural network (PCNN)* model based on the implementation of the mechanisms underlying the visual cortex of cat [9], [10], [14], [15]. The visual cortex is the part of the brain that receives information from the eye. The waves generated by each iteration of the PCNN algorithm create specific signatures of the scene used for segmentation (target detection as a first processing step of the present approach), namely to decide for each pixel its potential belonging to a certain object. Second processing step means object feature selection performed firstly using Gabor filtering [4], [5], [8] and refined by the second step based on the application of principal component analysis (PCA). The fourth processing stage of the method corresponds to support vector machine (SVM) classification using an improved training based on *virtual training data generation (VTDG) by concurrent self-organizing maps (CSOM)* [12], [14]. The proposed method is applied for military ground vehicle recognition in aerial images, being experimented for Moving and Stationary Target Acquisition and Recognition (MSTAR) public release database.

2. ATR PROCESSING CASCADE

The flowchart of the proposed ATR model is shown in Fig. 2. It consists of the following processing stages : (a) object detection using a pulse-coupled neural network (PCNN) segmentation module; (b) a first feature selection module using Gabor filtering (GF); (c) a second feature selection module using principal component analysis (PCA); (d) a support vector machine (SVM) classifier improved by using virtual training data generation (VTDG) with concurrent self-organization maps (CSOM).

FIG. 2. Flowchart of the neural network-based ATR cascade.

2.1. PCNN segmentation. PCNN is a biological inspired type of neural network; its functions are found in the visual cortex of mammals [10],It is a single layered, two-dimensional, laterally connected network of pulse-coupled neurons. There exists a one-to-one correspondence between the image pixels and network neurons. The PCNN equations are given below [10]:

$$F_{ij}[n] = \exp(-\alpha_F) * F_{ij}[n-1] + V_F \sum_k \sum_l M_{ijkl} Y_{kl}[n-1] + S_{ij} \quad (1)$$

$$L_{ij}[n] = \exp(-\alpha_L) * L_{ij}[n-1] + V_L \sum_k \sum_l W_{ijkl} Y_{kl}[n-1] \quad (2)$$

$$U_{ij}[n] = F_{ij}[n](1 + \beta L_{ij}[n]) \quad (3)$$

$$E_{ij}[n] = E_{ij}[n-1] \exp(-\alpha_E) + V_E Y_{ij}[n-1] \quad (4)$$

$$Y_{ij}[n] = \begin{cases} 1, & if\ U_{ij}[n] > E_{ij}[n] \\ 0, & otherwise \end{cases} \quad (5)$$

where α_F, α_L, and α_E are the time constants; V_F, V_L, and V_E are the magnitude adjustments; β is the linking strength of the PCNN. Each neuron is denoted with indices (i, j), and one of its neighboring neurons is denoted with indices (k, l). Feeding component $F_{ij}[n]$is combined with linking component $L_{ij}[n]$into neuron's internal activity $U_{ij}[n]$. The neuron receives input signals via feeding synapse M_{ijkl}, and each neuron is connected to its neighbors such that the output signal of a neuron modulates the activity of its neighbors via linking synapse W_{ijkl}. The pulse is able to feed back to modulate the threshold $E_{ij}[n]$ via a leaky integrator, raising the threshold by magnitude V_E that decreases with time constant α_E. During iterations, when a neuron's internal activity $U_{ij}[n]$ exceeds its dynamic threshold $E_{ij}[n]$, a pulse is generated (firings).

The PCNN model has proved to be very suitable for image segmentation [6], [9], [10], and [15]. To obtain an improved segmentation performance, we have proposed to run in parallel two PCNN segmentation models, one for n iterations and second for m iterations (see Fig. 3). The results of these two segmentations are added, and as a result the objects have been accentuated and any remaining noise is faded in the background. A good relation between n and m has proved to be m = n + 3. The segmentation is finished by thresholding the PCNN output image.

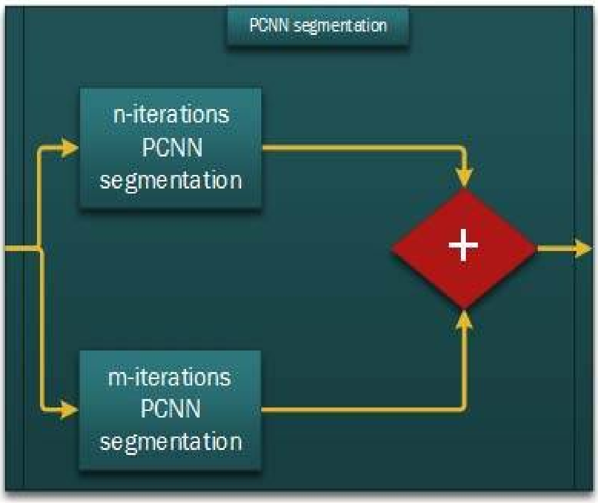


FIG. 3. Image segmentation using a combination of two parallel PCNN models.

2.2. Gabor Filtering (GF) as a first feature selection stage. First feature selection step has been performed by the standard 2-D Gabor filter [3], [4], [5], [8], [16]. It accepts the segmented images, from PCNN segmentation stage and provides an output column vector, consisting of the Gabor features. The feature vectors are normalized to zero mean and unit variance. The output GF vector has the length (m*n*u*v)/(d1*d2), where m and n represent the image with and height, u is the number of scales, v is the number of orientations, d1 is the factor of down sampling along rows and d2 is the factor of down sampling along columns.

2.3. Principal Component Analysis (PCA) as a second feature selection stage. We have applied PCA [1] to transform the GF output belonging to a (m*n*u*v)/(d1*d2)-dimensional space into a reduced p-dimensional space.

2.4. Support vector machine (SVM) classifier using an improved training based on virtual training data generation (VTDG) by concurrent self-organizing maps (CSOM). In order to improve training set quality, we use the innovative idea to build a training set composed by virtual samples only that completely substitute the input original samples [12]. The structure of the VTDG-CSOM is shown in Fig. 4. The steps of the proposed algorithm are the following:

a. Building M pattern subsets. One splits the original labeled sample set into M pattern subsets corresponding to each class, where M is the number of classes.
b. Training of each of the Self-Organizing Map modules SOM (k) (k=1,…, M). One uses the SOM unsupervised algorithm [11], [12], [13] to train each of the SOM modules. Namely, each SOM module has been trained only with the samples having the same label with the neural module label.
c. Virtual training data generation. After training, the weight vectors of the SOM modules have become virtual samples that substitute the input real samples, building an improved training set to obtain a more accurate classifier.

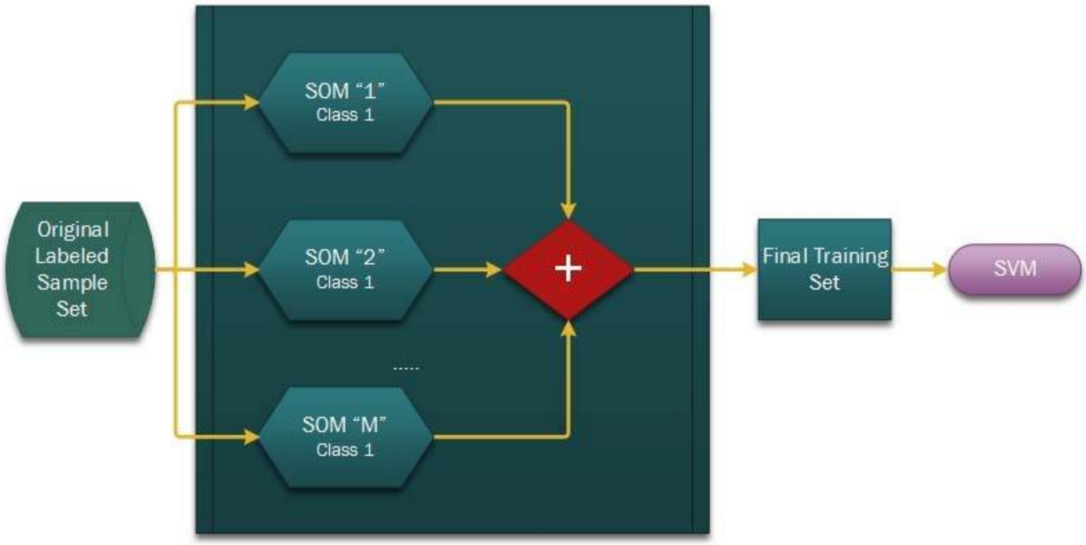
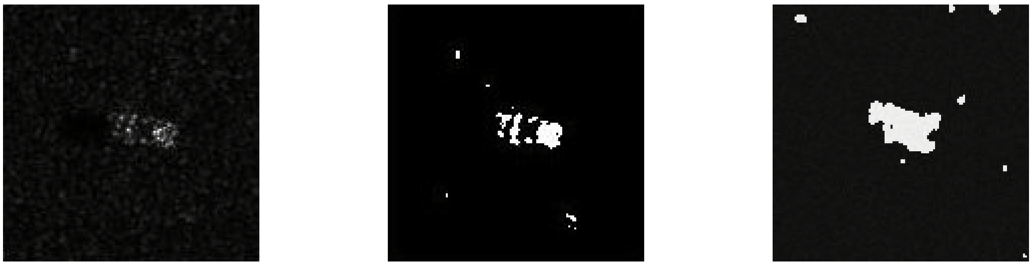


FIG. 4. Flowchart of the VTDG-CSOM.

3. EXPERIMENTAL RESULTS

3.1. MSTAR-database of military ground vehicles. We have intended to test the proposed model for the special defense application regarding identification of military ground vehicles. This why we have chosen MSTAR as a standard dataset for automatic target recognition (ATR) tasks. It is performed at the Redstone Arsenal, Huntsville, AL by the Sandia National Laboratory (SNL) using a Synthetic Aperture Radar (SAR) sensor platform. The collection was jointly sponsored by DARPA and Air Force Research Laboratory as part of the Moving and Stationary Target Acquisition and Recognition (MSTAR) program. SNL used an X-band SAR sensor in one foot resolution spotlight mode. Three classes of military ground vehicles of the former Soviet Union have been considered for our experiments from MSTAR database: BMP2 (infantry fighting vehicle), BTR70 (armored personnel carrier), and T72 (tank). We have chosen 2987 images of 128 x 128 pixels, using two depression angles: 15 degrees and 17 degrees. One chooses 1622 images corresponding to the 17-degree depression angle for training, while the other 1365 pictures corresponding to the 15-degree depression angle are considered for test.

3.2. Experimental performances. For our experiments, the image sizes are m=n=128. For PCNN, we have chosen the iteration parameters n = 21 and m = 24 to give the best results. We have denoted by RS (Reference Segmentation) the method consisting of median filtering, histogram equalization, and thresholding [18]. A comparison of PCNN versus RS performances may be subjectively evaluated in Fig. 5.



(a) (b) (c)

FIG. 5.(a) The original SAR image (BMP2); (b) Image obtained by reference segmentation (RS); (c) Image segmented with PCNN method (n = 21, m = 24).

The chosen GF parameters are u = 5, v = 21, d1 = d2 = 4, and consequently, the output of GF is a 107520-dimensional feature vector. After PCA, the GF output vector dimensionality is reduced to p=500. The experimental performances are shown in Table 1.

Table 1 – Maximum recognition score for MSTAR dataset.
R= (number of virtually generated training data using CSOM)/ (number of original training data).

Cascade	Maximum correct recognition score [%]	Parameters
PCNN-GF-PCA-{VTDG-CSOM}-SVM	97.36	R=3.25; rectangle sheet CSOM modules of sizes: (48x47), (28x27), (48x47); SVM of RBF type; gamma=0.0001
PCNN-GF-PCA- SVM	94.06	SVM of RBF type; gamma=0.0001
RS-GF-PCA-SVM	86.15	SVM of RBF type; gamma=0.01

4. CONCLUSIONS

The present paper has proposed a neural network-based ATR method using the synthetic aperture radar (SAR) imagery for the special application of military ground vehicle recognition. The proposed ATR has the following processing stages: PCNN segmentation for object detection; first feature selection stage using Gabor filtering (GF); second feature selection stage with PCA; SVM classification using an original neural-network procedure called VTDG-CSOM. Two of these stages are based on advanced neural networks: PCNN segmentation and VTDG-CSOM procedure for improvement of SVM classification performance. The experiments are performed using the dataset of military ground vehicles MSTAR (Moving and Stationary Target Acquisition and Recognition).

A comparison of PCNN versus RS performances subjectively evaluated in Fig. 3 can point out the objective advantage of PCNN segmentation over RS. For classical SVM classifier, Table 1 shows that PCNN segmentation leads to a 94.06% classification score versus 86.15% score obtained by the RS segmentation.

Table 1 also shows the advantage of the virtual data generation (VTDG) using CSOM in order to improve SVM classification. By substituting the original training samples with virtual samples generated by VTDG-CSOM model (the number of virtual samples being about R=3.25 times bigger than the number of original samples), one obtains an increasing of recognition score from 94.06% to 97.36%. Both neural network approaches (that of PCNN segmentation and also that of VTDG-CSOM classification) lead to the increasing pf classification performance from 86.15% to 97.36%, meaning a total score improvement of 11.21%.

REFERENCES

[1] M. Bishop, *Pattern Recognition and Machine Learning,* Springer, New York, 2006.
[2] C.H. Chen, *Signal and Image Processing for Remote Sensing,* 2^{nd} Edition, CRC. Press, New York, 2012.
[3] M. Haghighat, S. Zonouz, M. Abdel-Mottaleb, Identification Using Encrypted Biometrics, in *Computer Analysis of Images and Patterns, Lecture Notes in Computer Science,* Volume 8048, Springer, Berlin Heidelberg, pp. 440-448, 2013.
[4] J.K. Kamarainen, V. Kyrki, H. Kälviäinen, Fundamental Frequency Gabor Filters for Object Recognition, *Proc. 16th International Conference on Pattern Recognition (ICPR'02),* vol. 1, August 11-15, 2002, Quebec (Canada), pp. 628-631.
[5] J.K. Kamarainen, V. Kyrki, H. Kälviäinen, Robustness of Gabor Feature Parameter Selection, *Proc. IAPR Conference on Machine Vision Applications (IAPR MVA 2002),* Dec. 2002, pp. 132–135.

[6] J.A. Karvonen, Baltic Sea Ice SAR Segmentation and Classification Using Modified Pulse-Coupled Neural Networks, *IEEE Trans. on Geoscience and Remote Sensing,* Vol. 42, No. 7, pp. 1566-1574, July 2004.
[7] S. Kidera, T. Kirimoto, Accurate and Robust Automatic Target Recognition Method for SAR Imagery with SOM-based Classification, *IEICE Transactions on Communications*, Vol. E95-B No.11, pp. 3563-3571, Nov. 2012.
[8] V. Kyrki, J.K. Kamarainen, H. Kälviäinen, Simple Gabor Feature Space for Invariant Object Recognition, *Pattern Recognition Letters*, vol. 25, issue 3, pp. 311–318, Feb. 2004.
[9] G. Kuntimad and H.S. Ranganath, Perfect Image Segmentation Using Pulse Coupled Neural Networks, *IEEE Trans. on Neural Networks*, Vol. 10, No. 3, pp. 591-598, May 1999.
[10] Y. Ma, K. Zhan, Z.Wang, *Applications of Pulse-Coupled Neural Networks*, Springer, Heidelberg-New York, 2010.
[11] V.E. Neagoe and A. Ropot, Concurrent Self-Organizing Maps - A Powerful Artificial Neural Tool for Biometric Technology, pp. 291-298, in: *Harbour Protection Through Data Fusion Technologies*, Edited by: Shahbazian, E; Rogova, G; DeWeert, MJ, NATO Science for Peace and Security Series-C: Environmental Security, Springer, 2009.
[12] V.E. Neagoe, A.D. Ciotec, New Approach for Accurate Classification of Hyperspectral Images Using Virtual Sample Generation by Concurrent Self-Organizing Maps, *Proc. IEEE Internat. Geoscience and Remote Sensing Conf. (IGARSS 2013)*, July 21-26 2013, Melbourne (Australia), pp. 1031-1034.
[13] V.E. Neagoe, R.M. Stoica, A.I. Ciurea, L. Bruzzone, F. Bovolo, Concurrent Self-Organizing Maps for Supervised/Unsupervised Change Detection in Remote Sensing Images, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, Vol. 7, No. 8, pp. 3525-3533, August 2014.
[14] V.E. Neagoe, S. V. Carata, A. D. Ciotec, Automatic target recognition in SAR imagery using Pulse-Coupled Neural Network segmentation cascaded with virtual training data generation CSOM-based classifier, *Proc. 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2015)*, Milano, Italy, July 26-31, 2015, pp.3274-3277.
[15] H.S. Ranganath and G. Kuntimad, Object Detection Using Pulse Coupled Neural Networks, *IEEE Trans. on Neural Networks*, Vol. 10, No. 3, pp. 615-620, May 1999.
[16] H. Ruohong, Y. Ruliang, SAR Target Recognition Based on MRF and Gabor Wavelet Feature Extraction, *IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2008)*, July 2008, vol. 2, pp. II-907–II-910.
[17] N.M. Sandirasegaram, *Automatic Target Recognition in SAR Imagery using a MLP Neural Network*, Technical Memorandum, Defence Research and Development Canada (DRDC), Ottawa, TM 2002-120, Nov. 2002.
[18] M. Sezgin, B. Sankur, Survey over Image Thresholding Techniques and Quantitative Performance Evaluation, *Journal of Electronic Imaging*, Vol. 13, no 1, pp. 146–165, Jan. 2004.
[19] A. Streicher, E-Learning for radar image interpreters, in: *New Security Learning*, http://www.newsecuritylearning.com/index.php/feature/105-e-learning-for-radar-image-interpreters.
[20] Y. Sun, Z. Liu, S. Todorovic, J. Li, Adaptive boosting for SAR automatic target recognition, *IEEE Transactions on Aerospace and Electronic Systems*, vol. 43, issue 1, pp. 112–125, Jan. 2007.

AIR FORCE AND AEROSPACE ENGINEERING