SIAMESE NEURAL NETWORK FOR AUTOMATIC TARGET RECOGNITION USING SYNTHETIC APERTURE RADAR

Yasmine Khenchaf, Abdelmalek Toumi

SIAMESE NEURAL NETWORK FOR AUTOMATIC TARGET RECOGNITION USING SYNTHETIC APERTURE RADAR

Yasmine Khenchaf

Université de Bretagne-Occidentale,
Brest, France

Abdelmalek Toumi

Lab-STICC, UMR CNRS 685,
ENSTA Bretagne, Brest, France

ABSTRACT

Automatic Target Recognition (ATR) is an interest problem in various application fields (security, surveillance, automotive, environment, medicine, communications, remote sensing, ...). Thus, SAR (Synthetic Aperture Radar) and ISAR (Inverse Synthetic Aperture Radar) radar images provide rich visual information about the observed radar target. From these radar images, several methods have been proposed to meet the expected requirements in several application domains, including target recognition, which is one of the main issues addressed in the present work. Traditional standard image classification techniques are not suitable for efficient classification of SAR images due to the limited data available in some classes (unbalanced data). To solve these problems, we introduce a deep learning model, the Siamese network with multi-class classification, built from a pre-trained model to improve the model performances on unbalanced classes. To evaluate the proposed method, the MSTAR dataset is used. The proposed method improves the recognition rate from 95,18% to 97.16%.

***Index Terms*—** Automatic Target Recognition (ATR), Synthetic Aperture Radar (SAR), Convolutional Neural Network (CNN), Siamese Neural Network (SiNN).

1. INTRODUCTION

With the remarkable growth of various developments in electromagnetism and signal processing, the acquisition of images is carried out by radar (RAdio Detection And Ranging) which is an active system that can be used day and night and regardless of weather conditions. These two advantages have led to an expansion of research work providing this system with new and increasingly sophisticated functionalities and capabilities. Radar was designed primarily to detect and locate observed land, sea and air targets, but today it is used in both civilian and military applications. The most widespread radar images are the inverse synthetic aperture radar images (ISAR) and the direct aperture radar images (SAR)[1].

The recent advances in learning are giving rise to a new trend, the Deep learning, which will be done with the association of the features extraction layers and the classification tasks [2]. This methodology using deep learning will be applied later throughout the work done, it is based on advances the convolutional neural network (CNN). The final paper will be organized around three parts. First, section 2, The MSTAR database is presented and then we introduce the proposed CNN architecture and the correspondent results of some simulations. In section 3, we introduce the Siamese multi-class network (SiNN) and transfer learning. In this section, experiments and results obtained mainly on the MSTAR (Moving and Stationary Target Acquisition and Recognition) database are presented and analyzed. Conclusions and futurs works will be given in the last section of this work.

2. TARGETS RECOGNITION

2.1. Image database -MSTAR

We used the public MSTAR database, which is made up of SAR images. This database was developed as a collaborative effort between two agencies: Air Force Research Laboratory (AFRL) and Defense Advanced Research Projects Agency (DARPA). As shown in Figure 1 the MSTAR is a rich data

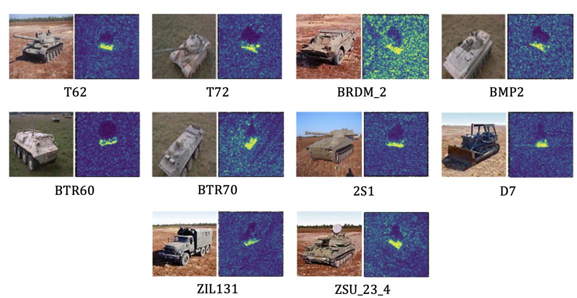

Fig. 1. Database MSTAR.

set of SAR images of 10 ground targets. It contains a total of 5165 images with 128x128 grayscale pixels. The SAR images are acquired in spotlight mode in the X-band (center frequency is 9.59 GHz) with a bandwidth of 0.591 GHz. The images were collected in HH (Horizontal/Horizontal) polarization in transmission and reception. The azimuth angle varies between 0° and 360°. The angular increment is approximately equal to 0.03°. The resolution in distance and azimuth is $0.3047m$ [3].

Targets	Train	Test
2S1	292	274
BMP2	233	195
BRDM2	298	274
BTR60	256	195
BTR70	233	196
D7	299	274
T62	298	273
T72	232	196
ZIL131	299	274
ZSU234	299	274

Table 1. Classes representations in MSTAR database

The database contains a total of 5165 images of size 128x128 in grayscale but throughout the work done, we reduced the size to 88x88. In the database, we have a training database containing 2740 images, and a test database with a total of 2425 images.

2.2. Proposed CNN Architecture

In the first step of this work, we design a CNN model adapted to SAR images for target recognition. Figure 2 shows the designed CNN architecture.

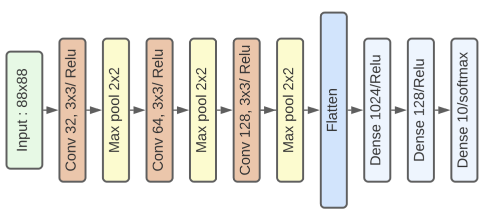

Fig. 2. A proposed CNN Architecture.

For the experimental part, we used the previously presented MSTAR database with an image size of 128x128x1 pixels. For computation time optimization and memory size constraints, all images have been resized to 88x88x1. The hyperparameters chosen for the training phase are presented in table 2. The choice of these hyperparameters values is motivated by the works and results obtained in the state of the art [4][5]. The Adam optimizer is chosen for its performance in optimization, which makes it the most often used in the literature. We have limited the number of epochs to 100 and a learning rate to 0.0001(1e-4). The algorithms are implemented in python 3.8 based on TensorFlow (2.5.0) and keras (2.5.0) with programming on GPU. The simulations are performed on a CentOS Linux 7 PC, Intel® Xeon(R) CPU E5-1660 v4 @ 3.20GHzx8, with NVIDIA TITAN V/PCle/SSE2 graphics card.

Batch size	32
Optimizer	Adam
number of epoch	100
learning rate	0.0001

Table 2. Training configuration in CNNs model.

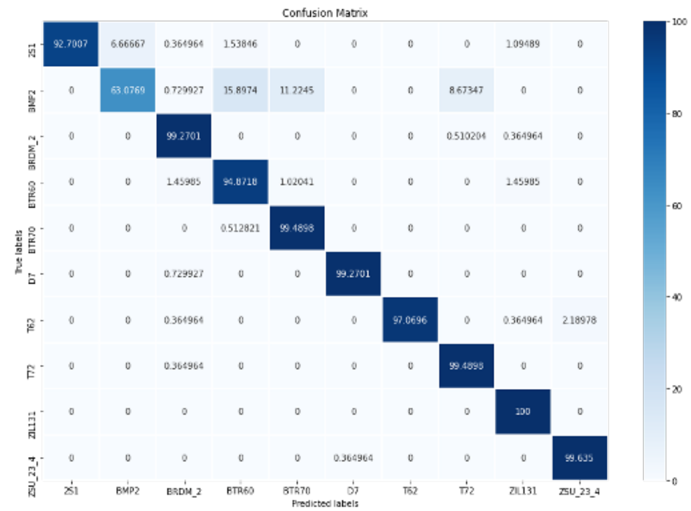

Fig. 3. Confusion matrix obtained by proposed CNN model.

The model gives us a recognition rate of 95.18%, with a learning time of 17s/epoch. The obtained confusion matrix is presented in figure 3. According to these results, we can note that some targets (BTR60, D7, T62, T72, ZIL131 and ZSU234) are recognized with the rates between 99.27% and 100%. We can see that for the proposed CNN model (see figure 3), 2S1 and BMP2 have the lowest recognition rate, unlike the other targets.

We have presented in this section a CNN model that we have implemented, the objective now is to improve the performance obtained, especially for targets that are poorly recognized, such as 2S1, BMP2 and BRDM2. For this purpose, we are interested in the combination of deep learning architectures, especially the Siamese network which will be introduced in the following section.

3. SIAMESE NEURAL NETWORK (SINN)

The Siamese network is a recent architecture, with little work on automatic target recognition. This architecture is applied in different application domains such as facial recognition, signature verification [6] and prescription pill identification. Siamese networks typically contain 2 (or more) identical subnetworks [7][8] that initially have the same architecture, parameters and weights. All parameter updates are reflected in both subnets, this means that updating the weights of one subnet will induce the automatic update of the weights of the other subnet. Generally, the Siamese network does not make multi-classes classification, but binary output which will allow to give us the similarity between the two inputs images (1 for similar inputs and 0 otherwise).

We have considered and implemented the following architecture (figure 4). In inputs for SiNN model, we insert

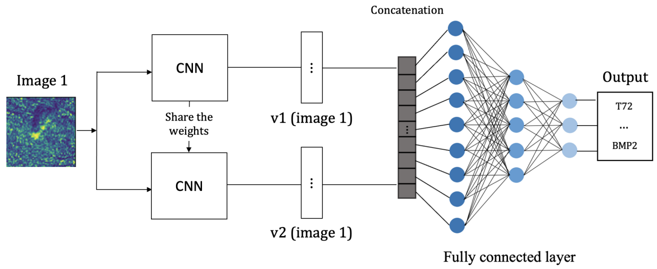

Fig. 4. Siamese network with multi-class classification.

identical images. Afterward, the feature vectors v1 and v2 will be concatenated instead of calculating the distance, or the inputs will pass through fully connected layers to then give us a probability of having fallen on the right target. Having two sub-networks, thus twice as much extraction, we need less data to adapt and train the network, which is the advantage of this network. The CNN model used in figure 2 is the one from section 2.2.

3.1. Siamese experimentation on 3 classes

The use of the Siamese network was motivated not for recognition but to improve the performance of the CNN against misclassified classes (targets). The selection of the images to make the pairs database concerns the misclassified targets. The selected images are obtained from class with important false alarm rate. Thus, three classes are selected BMP2, 2S1 and BRDM2 that given a low recognition rate. The selected images database contains 823 images for the learning dataset and 743 images for test dataset. In the learning process, we select in validation dataset 30% of the training data. In SiNN, the same CNN model is used on each branch (figure 4).

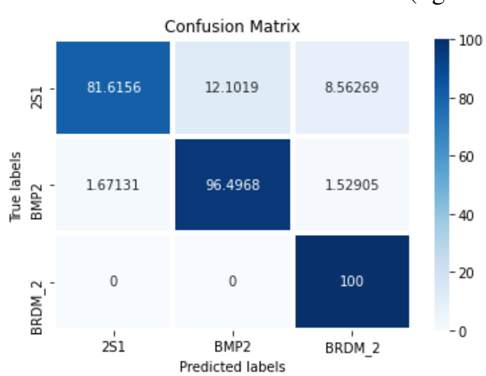

Fig. 5. Confusion matrix for SiNN (3 targets).

If we compare the recognition rates between the Siamese network and the initial proposed CNN. We can see a clear improvement with only the 2S1 target at a lower recognition rate. We note that the learning of the CNN model which composed SiNN model is performed from a random initialization (from scratch). To improve the performances, we used transfer learning which we present in the next part of this work.

3.2. Transfer learning

Transfer learning allows knowledge to be transferred from one model to another. In this case, transferring the weights of a pre-trained model to a specific task to associate them with another [9].

In order to improve the classification performances on the three problematic classes, we proposed the architecture detailed in figure 6. We freeze all the convolution layers in the pre-trained CNN (see figure 6) and then we add a convolution layer before the flatten, to extract a specific characteristics of the 3 considered targets. Afterward, the training process on new SiNN architecture is performed. In the last part of this

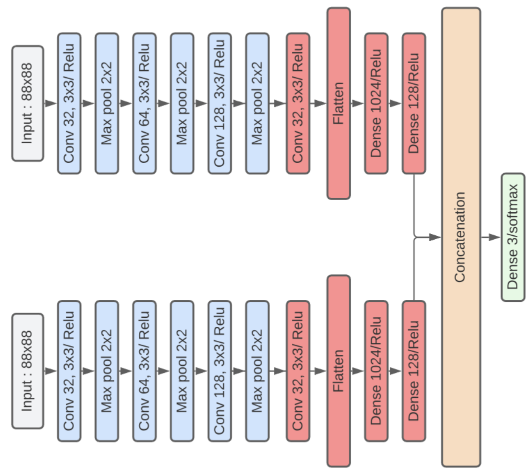

Fig. 6. Transfer learning.

paper, we have selected the CNN model (noted CNN2) which compose the new SiNN architecture to classify the 10 targets and try to obtain a better recognition rate than the initial proposed CNN (figure 2).

3.3. Siamese network for all targets

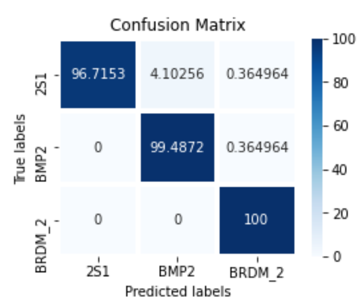

Fig. 7. Confusion matrix (SiNN) using transfer learning.

After using transfer learning, we go from 92% to 98.65% for a 3 classes classification, and an execution time that goes from 13s to 1s. After modifying the weights in the SiNN

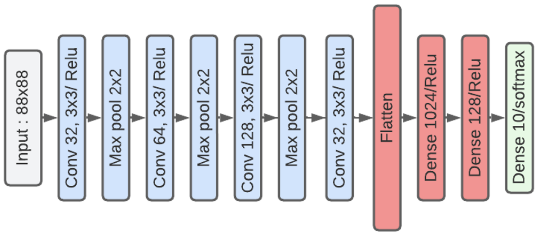

Fig. 8. 2nd configuration (fine tunning application)

architecture, The CNN2 is only selected to make all tagets classification. For this, we freeze the last convolution of the CNN2 model and then we train the entire model as shown by figure 8. The confusion matrices obtained, and which correspond to the two architectures are presented by figure 3 and 9, the architecture CNN2 provides a better results than the classical CNN. According to the results presented in the two confusion matrix, BMP2 remains the target with the lowest rate, but increases from 63% to 79.87%. However, we can notice that 2S1 and BRDM2 targets had a higher recognition rate.

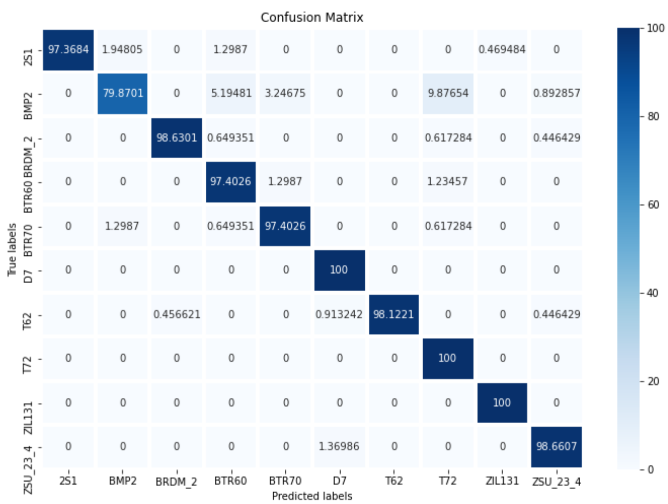

Fig. 9. Confusion matrix of the CNN2.

4. CONCLUSION

In this work, we propose to use the Siamese network to make in the first step a multi-class classification. So, The use of the Siamese network was motivated to improve the performance of the CNN against misclassified classes (targets). The use of Siamese networks can improve the shortcomings of some architectures (CNN). In this order, the transfer learning, like fine tuning can be interesting to exploit knowledge extracted from other databases. Experimental results on the MSTAR dataset confirm the effectiveness of the proposed methodology.

5. REFERENCES

[1] A. Toumi, A. Khenchaf, and B. Hoeltzener, “A retrieval system from inverse synthetic aperture radar images: Application to radar target recognition,” *Information Sciences*, vol. 196, pp. 73–96, 2012.

[2] A. Toumi, J-C. Cexus, and A. Khenchaf, “A proposal learning strategy on CNN architectures for targets classification,” in *2022 6th International Conference on Advanced Technologies for Signal and Image Processing (ATSIP)*, 2022, pp. 1–6.

[3] A. Karine, A. Toumi, A. Khenchaf, and M. El Hassouni, “Radar target recognition using salient keypoint descriptors and multitask sparse representation,” *Remote Sensing*, vol. 10, no. 6, 2018.

[4] J.A. Raj, S.M. Idicula, and B. Paul, “One-shot learning-based sar ship classification using new hybrid siamese network,” *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1–5, 2022.

[5] J. Tang, F. Zhang, Q. Yin, and W. Hu, “Synthetic aperture radar target recognition of incomplete training datasets via siamese network,” *The Journal of Engineering*, vol. 2019, 10 2019.

[6] S. Dey, A. Dutta, J.I. Toledo, S.K. Ghosh, J.Lladós, and U. Pal, “Signet: Convolutional siamese network for writer independent offline signature verification,” *ArXiv*, vol. abs/1707.02131, 2017.

[7] L.H. Hughes, N. Merkle, T. Bürgmann, S. Auer, and M. Schmitt, “Deep learning for sar-optical image matching,” in *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, 2019, pp. 4877–4880.

[8] W. Yang, J. Li, F. Fukumoto, and Y. Ye, “HSCNN: A hybrid-Siamese convolutional neural network for extremely imbalanced multi-label text classification,” in *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Online, Nov. 2020, pp. 6716–6722, Association for Computational Linguistics.

[9] R. Praneetha, T.M. Dhipu, and R. Rajesh, “Sar image classification using transfer learning,” in *2021 2nd International Conference on Range Technology (ICORT)*, 2021, pp. 1–6.