Volume 66, Issue 2, 2022

Journal of Scientific Research

of

The Banaras Hindu University

SAR Object Detection Using a Novel Convolutional Neural Network Architecture

Bibek Kumar[1[0000-0002-2081-2327]], Ranjeet Kumar Ranjan[1[0000-0002-8796-4579]] and Arshad Husain[1[0000-0003-0982-4789]]

[1]School of Computing, DIT University, Dehradun, India
bkknith@gmail.com

Abstract. In recent years, the importance of SAR (synthetic aperture radar) image analysis is growing day by day due to the vast applications in the field of oceanography, military war, land observation, agriculture, disaster management and geography. These applications required accurate classification of high-resolution SAR images. In the present study, two different machine learning techniques are applied on the SAR dataset to analyze the classification accuracy. The first classification technique is SVM (support vector machine) along with principal component analysis (PCA) for feature reduction. While, in the second technique, a novel CNN (convolutional neural network) has been proposed to perform the classification on SAR images. Essential experimental analysis has been carried out based on the MSTAR dataset. Experimental results of the current study showed that the proposed CNN model outperformed the SVM classifier with an accuracy of 98.69%.

Keywords: Convolutional Neural Networks, SAR Images, SVM, Automatic Target Recognition.

1 Introduction

The Synthetic aperture radar (SAR) is the field of remote sensing that has the power to capture images in the dark as well as cloudy weather, thus influencing a large dynamic area than optical images. The importance of SAR Imagery in the field of geology, oceanography, disaster management, land observation, agriculture and military operations makes them so popular. SAR is a radar, and it is working on a microwave band by using reflected signals from any targeted area(object) under any weather situation. SAR uses active sensors, and it provides its own illumination with a flexible look point and can select broad territory inclusion. The SAR images used range and azimuth compression method to produce a very high-resolution images of any targeted area. These images formed from microwave radar signals and represented in a complex domain with the help of phase information and amplitude. Due to the coherent nature of SAR images, these are affected by multiplicative noise also known as 'speckle' noise [11]. The speckle noise is responsible to affect the overall performance of any classification methodology. The amplitude data is the indication of the reflectivity of radar and it is generally represented in black and white images. The phase data is very complex to interpret but very helpful in classifying targets in SAR images with the help of feature extraction methods. ATR (Automatic target recognition) in the SAR imagery becomes important when we talk about military operations, disaster management kind of applications.

In this paper, the problem of ATR is SAR images is analyzed. The ATR is quite old as per the available literature [3]. There are so many feature types, classifiers and image processing techniques that have been developed so far [2, 17]. Different feature extraction techniques such ICA (Independent Component Analysis), PCA (Principal Component Analysis), and Hu moment invariants were developed with a variety of classifiers such as KNN (K-Nearest Neighbors), SVM (Support Vector Machine), QDC (Quadratic Discriminant Classifier) [15, 16, 19]. The appearance-based methodology was developed [20]. A good accuracy can also be achieved with limited dimensional features.

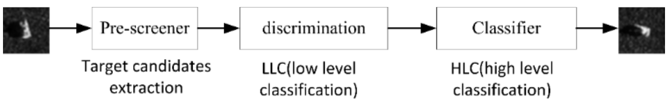

Fig. 1. Different steps in SAR image classification.

SAR-ATR is broadly categorized in three phases, as shown in Fig. 1), which are: pre-screener, discrimination, and classification. Pre-screener (also known as detection) detects an ROI (Region of Interest), discrimination (also known as a low-level classifier), discriminate ROI as a target or non-target region with output as a target chip. The last phase is classification (also known as a high-level classifier) and classifies target classes from a target chip.

Unlike in optical imagery, the issue with radar imagery classification is the presence of speckle noise. The presence of speckle noise in the radar imagery makes the data extraction process complicated. Due to the coherent nature of radar imaging, SAR images are affected by multiplicative noise also known as speckle noise [9, 12]. The presence of multiplicative noise in radar imagery can pose challenges in classification. It can be challenging to analyze the SAR images for classification and may lead to low accuracy. The high resolution of radar images produces a high number of features which leads to occurring overfitting issues with the classification of SAR images.

In this work, the authors have applied feature reduction over the available SAR dataset and then analyzed and compare the performance machine learning approach. The authors examined the performance of SVM (support vector machine) and CNN (convolutional neural network) to analyze the recognition accuracy on both techniques. The authors also

DOI: 10.37398/JSR.2022.660219

developed an advanced neural architecture to improve recognition accuracy.

In the section II, the authors explained SVM and CNN techniques in the form of feature set and network layer information. Experiment results have been discussed in the section III.

2 Machine learning approach and feature extraction

In this section, the authors explain the feature engineering and network layer approaches of the machine learning methods: SVM and CNN.

2.1 Support vector machine and feature engineering

SVM is the most popular supervised learning approach that can be useful in regression as well as in classification experiments. The SVM has been widely used in text recognition [7] and face recognition [13]. But SVM was mainly developed for binary classification problems by creating a hyperplane or line which separate the information into classes (see Fig. 2).

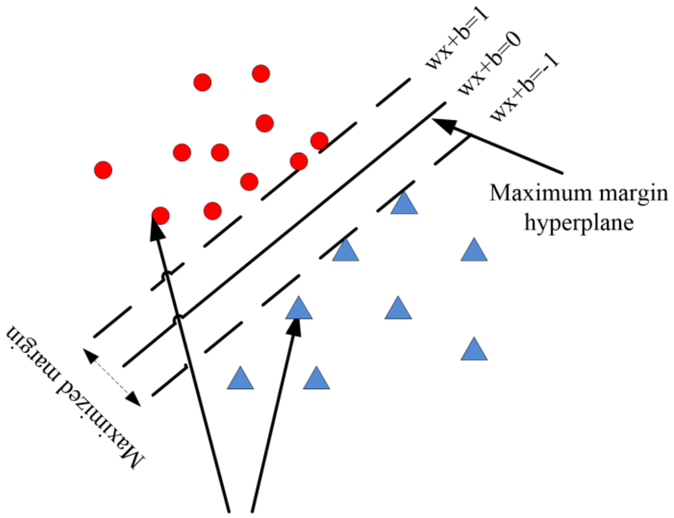

Fig. 2. The principle of SVM [4].

A linear model is used by SVM to realize nonlinear class boundaries. SVM transforms input data into a high dimensional space using a nonlinear mapping, so it enhances linearly separable than the actual space. But transformations lead to overfitting issues and high computational complexity. To overcome these problems, the support vectors and maximal margin hyperplane are used. The hyperplane can be set in such a way so that it becomes far away from classes and support vector instances that are nearest to the maximum margin hyperplane. The important concept is that the support vectors can be used to define maximum margin hyperplane. The remaining instances are not relevant. In the transformed high dimensional space, the maximum margin hyperplane defined by

$$p = c + \sum \beta_i q_i K\left(d(i), d\right) \quad (1)$$

$$q_i = \begin{cases} 1, \text{ if training instance } d(i) \text{ is in the class} \\ -1, \text{otherwise} \end{cases} \quad (2)$$

Where parameters c and β_i need to be learned, kernel function represented by $K\left(d(i), d\right)$. Out of different available kernel functions such as radial basis function, linear, laplace, polynomial and gaussian, the polynomial kernel is more popular which can be represented by $K\left(d(i), d\right) = (d(i) \bullet d)^n$. The formula can be cracked by a constrained quadratic optimization algorithm [22].

2.2 Proposed convolutional neural network model

Unlike in machine learning techniques, the beauty of CNNs is automatic feature reduction [6]. To design a suitable network architecture is the big question. Many building layers (blocks) can be used. The convolutional blocks include a set of 2D kernels with a specific size (usually the same). The parameter stride is used to controls the decimation of the outputs. To control the convolutional outputs and boundary effects, zero padding can be used. Generally, nonlinear activation function connected with convolutional blocks such as rectified linear unit (ReLU), tangent hyperbolic (tanh), sigmoid etc. ReLU is more expensive in terms of computational cost. Still, it is more popular due to its faster learning capability. In this paper, the authors proposed a CNN (see Fig. 3). The architecture includes convolutional layers, pooling layers, and fully connected layers. The two-dimensional convolutional layer involves several kernels. The convolutional kernel convolves two-dimensional matrix (input image) as a sliding window. The pooling layer (also known as sampling layer) convolve input with the sliding window to reduce running time and feature reduction purpose [18]. The performance of any neural network can also be optimized with the help of available optimization techniques [8].

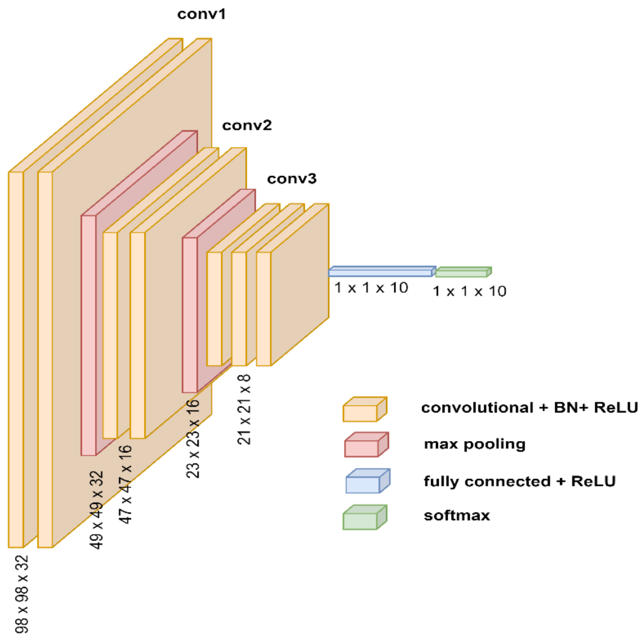

Fig. 3. The proposed CNN.

There are two more useful blocks. The first is average or max pull blocks, which can be utilized in data reduction. It works as a convolutional block. In convolutional block, a direct multiplication can be done with kernel whereas, to control volume data, the average value is returned in case of average or max pull blocks. To prevent overfitting issues, another block can be used and that is known as dropout. The dropout block drops some of the units and their respective connections from the network through a training process.

3 Result and discussion

This section is divided into three phases. The first section describes the dataset that has been used in this work, and the second section describes the result achieved with SVM classifier. Finally, the third section describes the result achieved with the proposed method.

3.1 Dataset

In this work, the authors used the MSTAR (moving and stationary target acquisition recognition dataset [10], which is freely available. The dataset contains ten military target classes of SAR images with 0.3m×0.3m resolutions images (see Fig. 4).

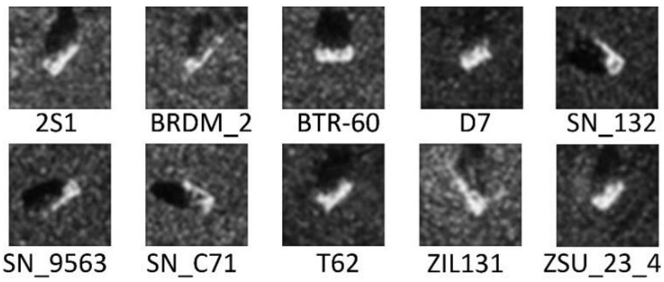

Fig. 4. Samples of ten class target images from MSTAR dataset.

The dataset contains images as mentioned in the Table 1. Training images captured on 17° and testing images captured on 15° azimuth angle.

3.2 Using SVM classifier

Images sample used in these experiments have 128×128 and then during data pre-processing these images are converted to a size of 100×100. In data pre-processing, the authors used principal component analysis (PCA) to reduce dimension (feature extraction). In the experiment, 10000 (100×100) feature set is converted to 3172 elements. Then it has been tested with the SVM classifier. The classifier performance is measured in the form of accuracy, which is defined as the number of correct classifications divided by the total number of tested samples. The overall accuracy achieved is 95.8%.

Table 1. Class wise size of training and testing dataset.

Class	Number of training sample (depression angle : 17°)	Number of testing sample (depression angle : 15°)
BTR60	256	195
T72	195	233
2S1	299	274
D7	299	274
ZIL131	299	273
BTR70	196	232
T62	299	273
BRDM2	298	274
BMP2	196	233
ZSU234	299	274
Total Classes:10	Total:2636	Total: 2535

Authors have been used different classifiers such as Naïve Bayes, K-Nearest Neighbor, Logistic Regression and these classifiers produced the accuracy as 84.92%, 94.6% and 93.8%, respectively. So, the authors have selected the SVM classifier to compare with the proposed novel CNN architecture method.

3.3 Using proposed CNN architecture

The architecture of the proposed CNN is described (see Fig. 3). The proposed architecture involves three layers with kernel size 3×3. The max-pooling gives us higher efficiency than average pooling. Batch normalization is used to speed up training and handling covariance shift. The activation function ReLU is being used in the proposed architecture to handle vanishing gradients. For training and testing purposes, authors have used different datasets. For the training purpose, the images captured at 17° azimuth angle have been used, whereas, for testing purposes, 15° azimuth angle images have been used as described in Table 1.

After execution of 100 epochs, the proposed model has achieved 99.60% of validation accuracy and 98.69% testing accuracy. The epoch versus the total number of iteration graph is shown in Fig. 5 and validation loss versus a total number of iteration graph is shown in Fig. 6. The confusion matrix is shown in Fig. 7. The grid in the diagonal position denoted the number of correct predictions against the real label. The other value in the confusion matrics denoted the misclassified targets.

From the results, it is observed that the convolutional neural network-based proposed architecture performed well compared to the SVM classifier. The overall accuracy achieved by SVM classifier is 95.8%, as compared to the proposed CNN architecture accuracy of 98.69%. The proposed CNN architecture are also compared with some other approaches and found having better accuracy than Inkawhich et. al [14] (95%), Borodinov et. al [1] (93.09%), Gorovyi et al [5] (94.1%) and Yu et. al [21] (95.3%).

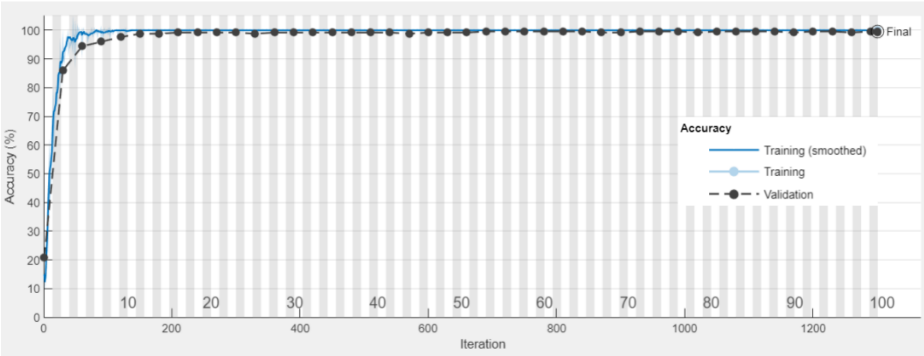

Fig. 5. Epoch vs. iteration graph.

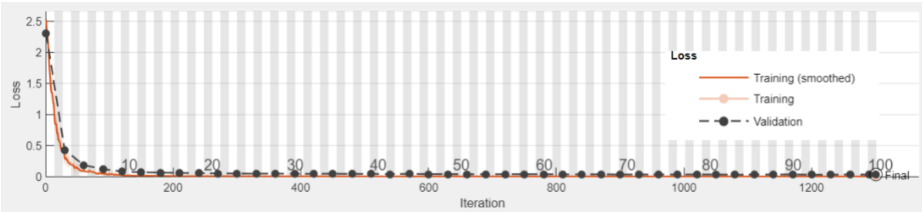

Fig. 6. Validation loss vs. iteration.

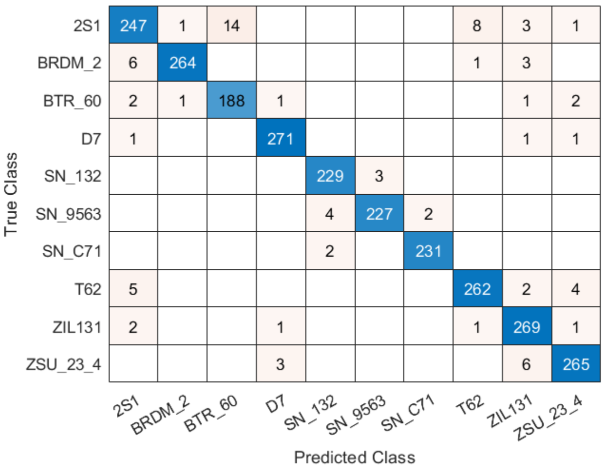

Fig. 7. Confusion matrix.

4 Conclusion

In this work, the SAR object detection methodology was analyzed with two methods which are based on SVM and CNN, and both the methods have given promising classification results. The results of the evaluation of these tow methods shows that SVM classifier can perform better along with a feature reduction approach which can be a difficult task in the case of high-resolution images such as SAR. A feature reduction technique PCA was used to tune the SVM classifier. The second proposed model is a convolutional neural network that has the capability to reduce features automatically. Based on the achieved results, the proposed CNN model produced a better accuracy of 98.69%, as compared with the PCA based SVM classifier with an accuracy 95.8%.

References

1. Borodinov, A.A., Myasnikov, V. V.: Classification of radar images with different methods of image preprocessing. CEUR Workshop Proc. 2210, 6–13 (2018). https://doi.org/10.18287/1613-0073-2018-2210-6-13.
2. Coman, C., Thaens, R.: A deep learning SAR target classification experiment on MSTAR dataset. Proc. Int. Radar Symp. 2018-June, 1–6 (2018). https://doi.org/10.23919/IRS.2018.8448048.
3. El-Darymli, K. et al.: Automatic Target Recognition in Synthetic Aperture Radar Imagery: A State-of-the-Art Review. IEEE Access. 4, c, 6014–6058 (2016). https://doi.org/10.1109/ACCESS.2016.2611492.

4. Gholami, R., Fakhari, N.: Support Vector Machine: Principles, Parameters, and Applications. Elsevier Inc. (2017). https://doi.org/10.1016/B978-0-12-811318-9.00027-2.
5. Gorovyi, I.M., Sharapov, D.S.: Comparative analysis of convolutional neural networks and support vector machines for automatic target recognition. MRRS 2017 - 2017 IEEE Microwaves, Radar Remote Sens. Symp. Proc. 63–66 (2017). https://doi.org/10.1109/MRRS.2017.8075027.
6. Huang, Z. et al.: What, Where, and How to Transfer in SAR Target Recognition Based on Deep CNNs. IEEE Trans. Geosci. Remote Sens. 58, 4, 2324–2336 (2020). https://doi.org/10.1109/TGRS.2019.2947634.
7. Joachims, T.: Text categorization with support vector machines: Learning with many relevant features. Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics). 1398, 137–142 (1998). https://doi.org/10.1007/s13928716.
8. Joshi, C. et al.: A Fuzzy Logic based feature engineering approach for Botnet detection using ANN. J. King Saud Univ. - Comput. Inf. Sci. xxxx, (2021). https://doi.org/10.1016/j.jksuci.2021.06.018.
9. Katiyar, A., Santhi, V.: Region based speckle noise reduction approach using fuzzy techniques. 2017 Int. Conf. Intell. Comput. Instrum. Control Technol. ICICICT 2017. 2018-Janua, 673–676 (2018). https://doi.org/10.1109/ICICICT1.2017.8342643.
10. Keydel, E.R. et al.: against measured data over a broad range of challenging real world battlefield scenarios . These Extended Operating Conditions Table 1 : Summary OfMSTAR Extended Operating Conditions ntra-Cass v ' arability Obscuratin (Occkision and / orLiyover): up to . 228–242 (1975).
11. Kumar, B. et al.: A Multi-Objective Enhanced Fruit Fly Optimization (MO-EFOA) Framework for Despeckling SAR Images using DTCWT based Local Adaptive Thresholding. October 2020, (2021).
12. Moreira, A. et al.: A tutorial on synthetic aperture radar. IEEE Geosci. Remote Sens. Mag. 1, 1, 6–43 (2013). https://doi.org/10.1109/MGRS.2013.2248301.
13. Nasr, S. et al.: Face recognition system using bag of features and multi-class SVM for robot applications. 2017 Int. Conf. Control. Autom. Diagnosis, ICCAD 2017. 263–268 (2017). https://doi.org/10.1109/CADIAG.2017.8075668.
14. Nathan Inkawhich Matthew J. Inkawhich , Eric K. Davis , Uttam K. Majumder, S.M., Erin Tripp , Chris Capraro, and Y.C.: Bridging a Gap in SAR-ATR: Training on Fully Synthetic and Testing on Measured Data. IEEE J. Sel. Top. Appl. EARTH Obs. Remote Sens. 14, (2021).
15. Rai, H.M., Chatterjee, K.: Hybrid adaptive algorithm based on wavelet transform and independent component analysis for denoising of MRI images. Meas. J. Int. Meas. Confed. 144, 72–82 (2019). https://doi.org/10.1016/j.measurement.2019.05.028.
16. Rajapriyadharshini, R., Benadict Raja, J.: SAR image denoising via clustering based Linear Discriminant Analysis. ICIIECS 2015 - 2015 IEEE Int. Conf. Innov. Information, Embed. Commun. Syst. 0–5 (2015). https://doi.org/10.1109/ICIIECS.2015.7193136.
17. Shi, X. et al.: Automatic target recognition for synthetic aperture radar images based on super-resolution generative adversarial network and deep convolutional neural network. Remote Sens. 11, 2, (2019). https://doi.org/10.3390/rs11020135.
18. Singh, A. et al.: Credit Card Fraud Detection under Extreme Imbalanced Data: A Comparative Study of Data-level Algorithms. J. Exp. Theor. Artif. Intell. 00, 00, 1–28 (2021). https://doi.org/10.1080/0952813X.2021.1907795.
19. Subba Reddy, T., Harikiran, J.: Hyperspectral image classification using support vector machines. IAES Int. J. Artif. Intell. 9, 4, 684–690 (2020). https://doi.org/10.11591/ijai.v9.i4.pp684-690.
20. Wang, C. et al.: When deep learning meets multi-task learning in SAR ATR: Simultaneous target recognition and segmentation. Remote Sens. 12, 23, 1–19 (2020). https://doi.org/10.3390/rs12233863.
21. Yu, H. et al.: PCANet based nonlocal means method for speckle noise removal in ultrasound images. PLoS One. 13, 10, 1–19 (2018). https://doi.org/10.1371/journal.pone.0205390.
22. Zhao, Q., Principe, J.C.: Support vector machines for SAR automatic target recognition. IEEE Trans. Aerosp. Electron. Syst. 37, 2, 643–654 (2001). https://doi.org/10.1109/7.937475.

***