# ML-Based Target Detection in SAR Images

Vamsee Krishna Tunuguntla
*Department of Electrical and Computer Engineering*
*Stevens Institute of Technology*
vtunugun@stevens.edu

Yu-Dong Yao
*Department of Electrical and Computer Engineering*
*Stevens Institute of Technology*
yyao@stevens.edu

***Abstract*—Synthetic Aperture Radar (SAR) images have emerged as a powerful tool for analyzing large areas despite adverse weather conditions such as fog and clouds. High-resolution SAR images enable the detection of potential threats and assessment of the situation. However, the vast amount of data and the unique characteristics of SAR imagery require considerable time and cost to analyze. Automatic target recognition is essential to detect and classify potential targets and reduce time and cost factors. Recent advancements in deep learning-based domain adaptation have made it possible to achieve automatic detection on an unprecedented level. Simulated data can now be used to almost eliminate the need for rare and costly real target data. SAR ATR is crucial for remote-sensing image recognition and has a wide range of applications in military surveillance, national defense, civil applications, and more. Convolutional Neural Network (CNN) has been widely used for SAR ATR, but it is challenging to train models with limited ray SAR images. The proposed method exploits the advancements in deep learning-based domain adaptation to improve the performance of SAR ATR. The results demonstrate the effectiveness of the proposed approach in reducing the cost required for SAR ATR while maintaining high accuracy. Future work could explore the use of other deep learning techniques for SAR ATR and investigate the transferability of the proposed approach to other domains.**

***Index Terms*—Synthetic Aperture Radar (SAR), Automatic Target Recognition (ATR), Deep Learning, Military Surveillance**

## I. Introduction

Synthetic Aperture Radar (SAR) is an important imaging technology in both military and civilian fields due to its all-time, all-weather, long-range, and high-resolution imaging capabilities [1], [2]. As a result, it has been widely used in various applications, including target detection [3], target recognition [4], and terrain classification. Among these applications, automatic target recognition (ATR) plays a crucial role in SAR civil and military fields. In the context of target recognition, human identification of objects of interest in SAR images is time-consuming and often impractical. Therefore, a standard SAR-ATR system consists of three main components: detection, discrimination, and classification, as proposed by the MIT Lincoln Laboratory [5].

*1) Detection:* This component aims to identify potential targets in SAR images.

*2) Discrimination:* While this component aims to differentiate between targets and non-targets.

*3) Classification:* Finally, this component automatically assigns a label to the target image based on the extracted features.

### A. Target Classification

Target classification is a complex problem due to a large amount of data and the unique characteristics of SAR imagery. Therefore, various machine-learning techniques have been used for SAR-ATR, including support vector machines, artificial neural networks, and decision trees. Recently, deep learning-based approaches, particularly Convolutional Neural Networks (CNNs), have achieved significant improvements in SAR-ATR performance.

Current SAR ATR methods can be broadly categorized into three mainstream paradigms:

*1) Template-based:* The template-based approach involves matching the target with a known template, which requires accurate registration and is sensitive to variations in the target's appearance.

*2) Model-based:* The model-based approach represents the target's features using mathematical models and compares them with the features of the input image, but it is computationally intensive and may require extensive manual feature engineering.

*3) Machine learning:* The machine learning approach is currently the most popular and effective method for SAR ATR. It involves training a classifier using a large amount of labeled data to automatically recognize targets in new images.

In recent years, deep learning has emerged as a promising technique in various technical fields, including SAR target classification [6]. Convolutional neural networks (CNN) can learn complex features from original SAR images with convolution and pooling layers. VGGNet, ResNet, and DenseNet are popular CNN architectures used as backbone structures in SAR image interpretation methods [7]. Deep-learning-based SAR image target recognition methods mainly focus on aircraft, ship, and vehicle targets [8], and require abundant target data.

### B. Moving and Stationary Target Acquisition and Recognition (MSTAR)

In recent years, the Moving and Stationary Target Acquisition and Recognition (MSTAR) dataset has become a widely used benchmark for SAR ATR. Many novel SAR ATR algorithms and systems have been proposed and designed based on the MSTAR dataset, and most of them have performed well in various applications. However, SAR ATR remains a highly challenging task. The MSTAR dataset [9] serves as a benchmark for evaluating and comparing SAR vehicle target classification algorithms. The advantage of deep learning is that it can automatically learn high-dimensional features of the target, eliminating the need for complex artificial feature extraction. Additionally, it achieves higher target recognition accuracy in the same training and testing environment. However, deep learning models lack interpretability compared to traditional methods, and their robustness is insufficient.

### C. Related Work

Traditional SAR target recognition methods are focused on electromagnetic scattering features. These methods describe the scattering center of the target and establish a correspondence between sets of scattering centers to evaluate their similarity. Chiang et al. [10] utilized a posteriori probability to evaluate the similarity of two sets of ASCs and proposed a Bayesian approach to establish a one-to-one correspondence between them. Dungan et al. [11] used the square Hausdorff distance with the least trimming to match the scattering center sets. Fu et al. [12] proposed a novel scattering structure feature-based, template-matching-based aircraft recognition method to improve classification accuracy in SAR images. Ding et al. [13] for SAR ATR, a method of hierarchical fusion of the global and local features was proposed.

In contrast, deep learning methods directly learn high-dimensional target features through the network without considering electromagnetic scattering characteristics. Chen et al. [14] a fully all-convolutional network (A-ConvNet) that substitutes convolutional layers for fully connected layers was first proposed for SAR vehicle target classification. Shang et al. [15] also proposed a vehicle target classification network structure known as M-net. A two-step approach to training the network can speed up the process and make it more stable at the same time. A transfer-learning-based approach is proposed in [16] that enables the transfer of knowledge from sufficient unlabeled SAR scene images to labeled SAR target data. The aforementioned research findings demonstrate that the deep learning method can achieve higher classification accuracy. However, the limited number of training samples and a large number of network parameters make it easy to overfit. To increase the network's robustness, therefore, data augmentation and parameter transfer are required.

The structure of this paper is as follows. In Section II, we introduce the problem statement and the system model. In Section III, we present the proposed solution for the problem statement. In Section IV we discuss the numerical results and analyze the evaluation of the performance of the proposed model. Finally, Section V mentions the conclusion of the research problem.

## II. Problem Statement

Target classification in SAR images is a complex problem due to several factors. One of the major challenges is a large amount of data, which requires efficient algorithms for processing and analysis. SAR images typically have high dimensionality, making it difficult to extract relevant features for target detection and classification. In addition, the unique characteristics of SAR imagery, such as speckle noise, geometric distortion, and layover effects, make it challenging to identify targets accurately.

Another complexity arises from the diversity of target types and variations. Targets in SAR images can have different orientations, aspect angles, and sizes, making it difficult to develop a single model that can accurately detect and classify all targets. Additionally, targets can be partially obscured by natural terrain or other objects, making it difficult to distinguish them from their surroundings.

### A. System Model

To address these challenges, various machine-learning techniques have been used for SAR target detection and classification. The proposed system model for target detection and classification of SAR images using deep learning algorithms consists of three main components:

*1) Detection:* The detection component aims to identify potential targets in SAR images using CNN-based models such as SARNet and VGGNet.

*2) Discrimination:* The discrimination component aims to differentiate between targets and non-targets using feature extraction techniques such as transfer learning and data augmentation.

*3) Classification:* Finally, the classification component automatically assigns a label to the target image based on the extracted features using machine learning models.

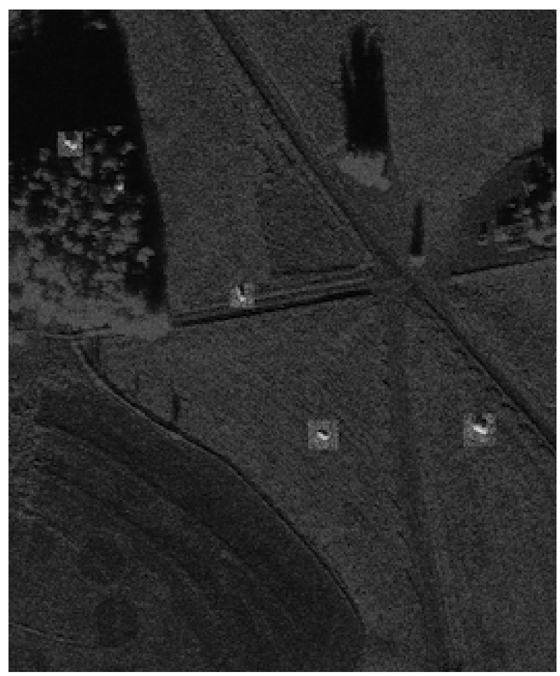

Fig. 1. SAR Image sample with multiple armor vehicles.

We are considering MSTAR dataset which consists of SAR images of different types of vehicles, collected using an X-band sensor in spotlight mode, with a 1-foot resolution and 15-degree depression angle. The MSTAR dataset includes five

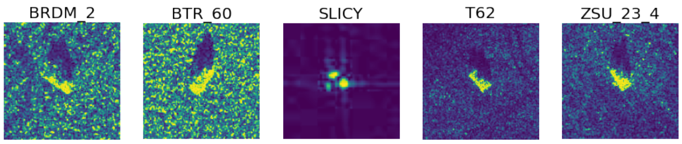


Fig. 2. Target vehicle samples represented in SAR images.

types of ground armor vehicles, namely 'BRDM-2', 'BTR-60', 'SLICY', 'T62', and 'ZSU-23-4' as shown in Fig. 2. Also, we considered the MathWorks dataset, which includes 300 training images with multiple targets and their respective locations in each image for the ATR. A training sample from the MathWorks dataset is shown in Fig. 1.

The system will be trained and evaluated using the datasets, which contain a diverse range of target types and variations. The proposed system aims to achieve high target recognition accuracy while addressing the interpretability and robustness issues of deep learning models in SAR target classification.

## III. Proposed Solution

In this project, our goal is to employ deep learning models, namely CNN, SARNet, and VGG, for target classification. We will implement one of the classified models to detect and label the target types on the SAR sample image.

### A. Data Preprocessing

Before training the models, we need to preprocess the data to ensure that it is in a suitable format for the models. This involves tasks such as resizing the images, normalizing the pixel values, and splitting the data into training, validation, and test sets. Additionally, we need to perform exploratory data analysis to gain insights into the data and identify any potential issues that may affect the model's performance.

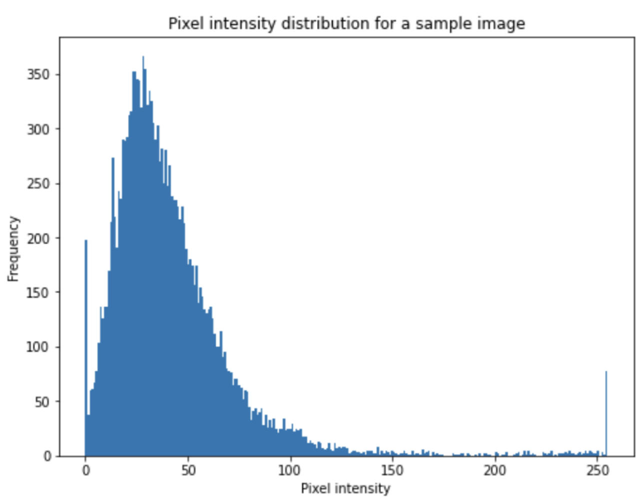


Fig. 3. Pixel intensity distribution for a sample image.

Here we have transformed the image to the desired shape (128, 128, 3). Exploratory data analysis has been performed to check pixel intensity distribution as well as to check the number of images per category. Fig. 3 displays the pixel intensity distribution of a sample image using a histogram. The histogram indicates that the majority of the pixels are under bin 100. Fig. 4 resembles the number of SAR images with respect to each target class.

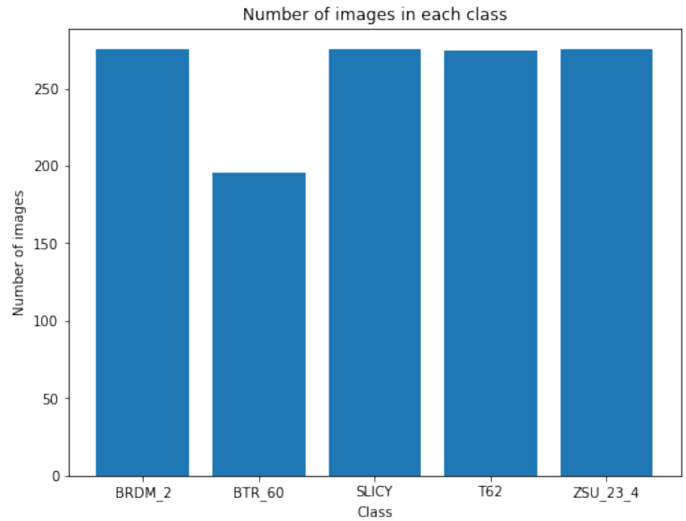


Fig. 4. Number of SAR images per each target class.

By preprocessing the data and performing exploratory data analysis, we can ensure that our models are trained on high-quality data, and by evaluating their performance, we can determine which model is best suited for this task. As the data processing was done, we start training the models. We will use the MSTAR dataset to train the models and the validation set to tune their hyperparameters and prevent overfitting. We will evaluate the performance of the models using evaluation metrics i.e., prediction accuracy on the test set.

### B. Models

We have considered 3 models (CNN, SARNet, and VGG) to perform the classification task.

*1) CNN Model:* In this task, we used a Convolutional Neural Network (CNN) model for target classification in Synthetic Aperture Radar (SAR) images. CNNs are widely used in computer vision tasks due to their ability to automatically learn complex features from raw data [17].

The CNN model used in this study consists of a series of convolutional layers with an increasing number of filters, followed by max pooling layers to reduce the spatial dimensions of the output feature maps. Specifically, the model contains three convolutional layers with 32, 64, and 128 filters respectively, each using a 3x3 kernel and ReLU activation. The output of each convolutional layer is then passed through a max pooling layer with a 2x2 pool size to reduce the spatial dimensions by a factor of 2. After the convolutional layers, a Flatten layer is added to convert the output of the convolutional layers into a 1D array that can be passed to fully connected layers. A Dense layer with 64 units and ReLU activation is added as a fully connected layer, and finally, a Dense layer with 5 units and softmax activation is added as the output layer which performs the final classification of the target. The model architecture for the described CNN model is shown in Fig. 5.

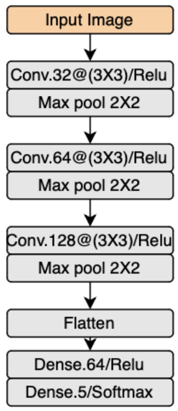


Fig. 5. Convolutional Neural Network (CNN) model architecture.

*2) SARNet Model:* SARNet is a deep learning model that is specifically designed for Synthetic Aperture Radar (SAR) image classification [18]. The model consists of a series of convolutional and pooling layers followed by fully connected layers for classification. The key feature of SARNet is that it incorporates a SAR-specific preprocessing step that enhances the quality of the SAR image before classification. The pre-processing step included multi-looking and speckle filtering to improve the image quality.

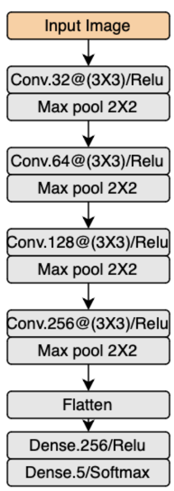


Fig. 6. SAR Network (SARNet) model architecture.

The model consists of a series of convolutional layers with an increasing number of filters, followed by max pooling layers to reduce the spatial dimensions of the output feature maps. Specifically, the model contains four convolutional layers with 32, 64, 128, and 256 filters respectively, each using a 3x3 kernel and ReLU activation. The output of each convolutional layer is then passed through a max pooling layer with a 2x2 pool size to reduce the spatial dimensions by a factor of 2. After the convolutional layers, a Flatten layer is added to convert the output of the convolutional layers into a 1D array that can be passed to fully connected layers. A Dense layer with 256 units and ReLU activation is added as a fully connected layer, and finally, a Dense layer with target classes units and softmax activation is added as the output layer. The model architecture for the described SARNet model is shown in Fig. 6.

*3) VGG Model:* The VGG (Visual Geometry Group) model is a deep convolutional neural network architecture that has achieved the state of the art results in various image classification tasks. In recent years, it has also been applied to SAR image classification tasks due to its excellent performance in feature extraction [19].

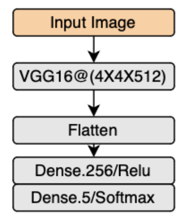


Fig. 7. VGG (Visual Geometry group) model architecture.

For our SAR image classification task, we used the VGG16 model, which is included without its top layer. A Flatten layer is added to convert the output of the convolutional layers into a 1D array that can be passed to fully connected layers. A Dense layer with 256 units and ReLU activation are added as a fully connected layer, and finally, a Dense layer with target classes units and softmax activation is added as the output layer. The input to the network is a SAR image, and the output is a vector of probabilities representing the likelihood of each target class. The model architecture for the described VGG model is shown in Fig. 7.

*4) Automatic Target Recognition (ATR):* We performed detection and classification tasks on the Mathworks dataset using transfer learning techniques and the SARNet model, originally developed for target classification. The images were preprocessed and hot code was enabled for the targets.

The proposed CNN model architecture for target detection consists of multiple layers. The first layer is a 2D convolutional layer with 32 filters of size 3x3 and a Rectified Linear Unit (ReLU) activation function. The input shape of the layer is (128, 128, 1), representing the dimensions of the input image.

A 2D max pooling layer with a pool size of 2x2 follows the first convolutional layer to reduce the spatial dimensions of the feature maps. A second 2D convolutional layer with 64 filters of size 3x3 and a ReLU activation function is added after the first max pooling layer. Another 2D max pooling layer is added after the second convolutional layer.

After the max pooling layer, the feature maps are flattened into a 1D vector, which is passed through a fully connected layer with 64 neurons and a ReLU activation function. Finally, a dense output layer with 5 neurons and a softmax activation function is added to classify the targets into one of the five possible categories.

The proposed model architecture combines convolutional and fully connected layers to extract features and classify targets in the MathWorks dataset. Categorical cross-entropy loss and the Adam optimizer are used for model optimization.

## IV. Numerical Results and Analysis

*1) CNN Model:* The model was trained using an MSTAR training dataset and validated. Fig. 8 represents the confusion matrix of the convolutional neural network (CNN) model comparing the actual and predicted target values.

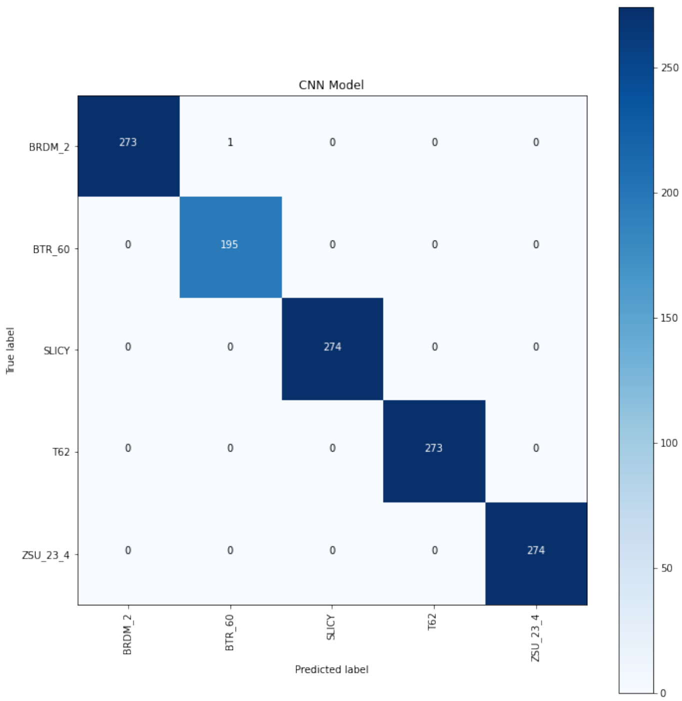


Fig. 8. Convolutional Neural Network (CNN) model confusion matrix.

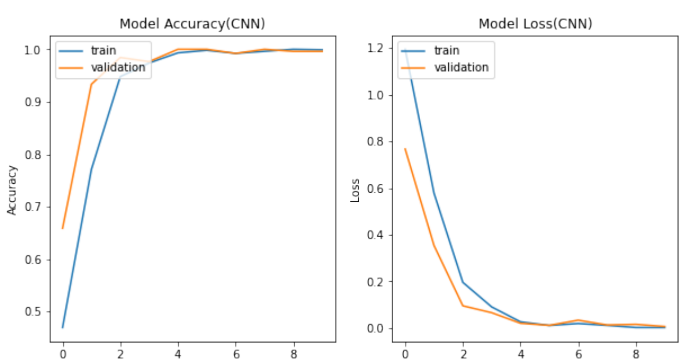


Fig. 9. Visualization graphs for training and validation (CNN).

Furthermore, it was observed from Fig. 9 that the model performed with optimum accuracy within 10 epochs. This indicates that the model learned the patterns and relationships from the preprocessed data. The model achieved an accuracy of 99.6 on the test dataset.

*2) SARNet Model:* The SARNet model was trained and validated using an MSTAR dataset. The confusion matrix in Fig. 10 shows the comparison between the actual and predicted target values.

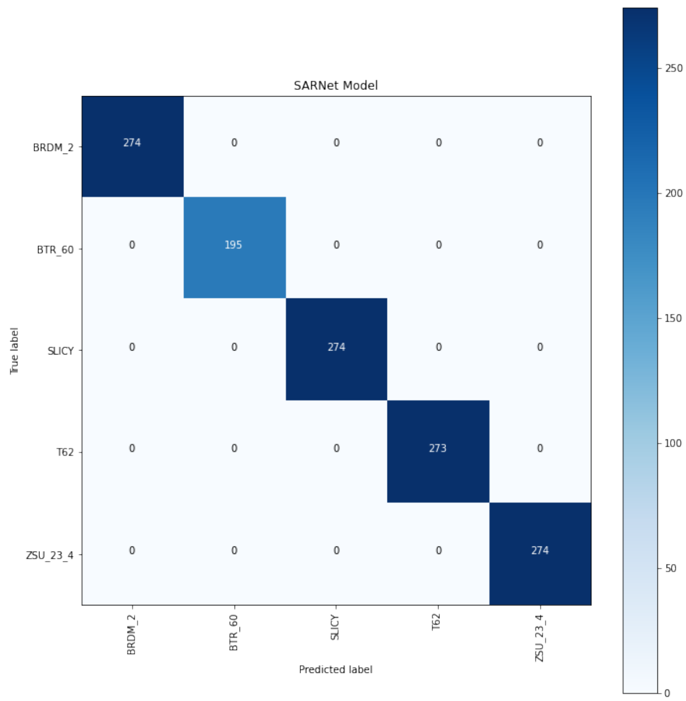


Fig. 10. SAR Network (SARNet) model confusion matrix.

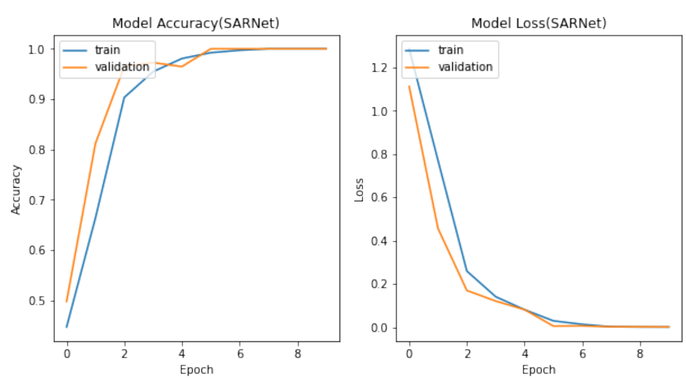


Fig. 11. Visualization graphs for training and validation (SARNet).

Fig. 11 indicates that the model achieved optimal accuracy within 10 epochs, suggesting that it learned patterns and relationships from preprocessed data. The model achieved 100 accuracy on the test dataset due to model-specific features.

*3) VGG Model:* The VGG model was trained and tested on the MSTAR dataset and achieved an accuracy of 72.1 on the test data. It demonstrated optimal accuracy within 10 epochs, suggesting that it learned the patterns from the preprocessed data.

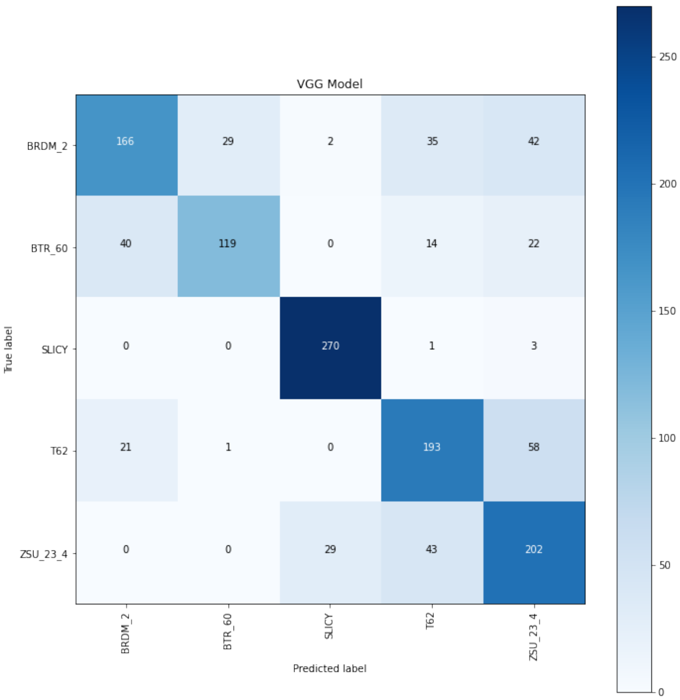


Fig. 12. VGG model confusion matrix.

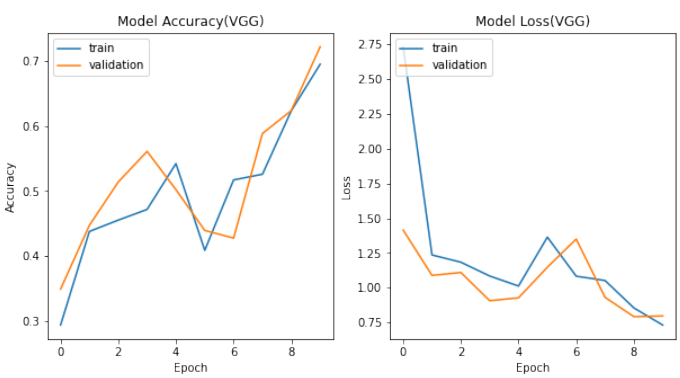


Fig. 13. Visualization graphs for training and validation (VGG).

However, compared to the other models, CNN and SARNet, VGG underperformed due to overfitting and slower convergence. This is evident from the decrease in accuracy in the initial epochs before it reaches its optimum accuracy. The evaluation of VGG's performance was done using a confusion matrix and is shown in Fig. 12, while Fig. 13 demonstrates the accuracy and loss of the model on the training and validation dataset.

*4) Automatic Target Recognition (ATR):* The SARNet model was integrated with the Mathworks dataset for target recognition and classification. Fig. 14 shows that the model achieved optimal accuracy within 10 epochs, indicating that it learned the data patterns effectively. The model was tested on a separate dataset and achieved an accuracy of 99.2.

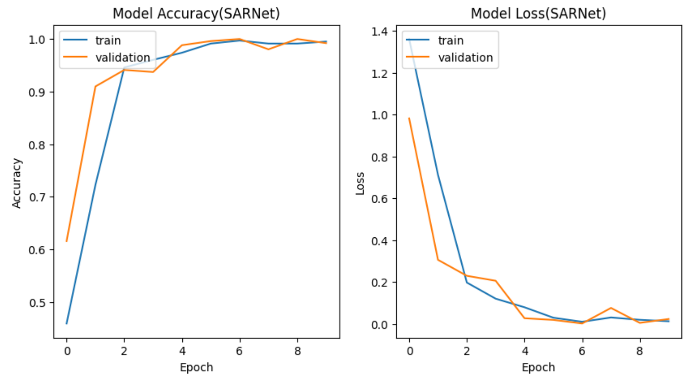


Fig. 14. Visualization graphs for training and validation.

Fig. 15 shows the final output of the ATR. After detecting and classifying the target object, the model utilizes OpenCV to label the targets on a SAR image.

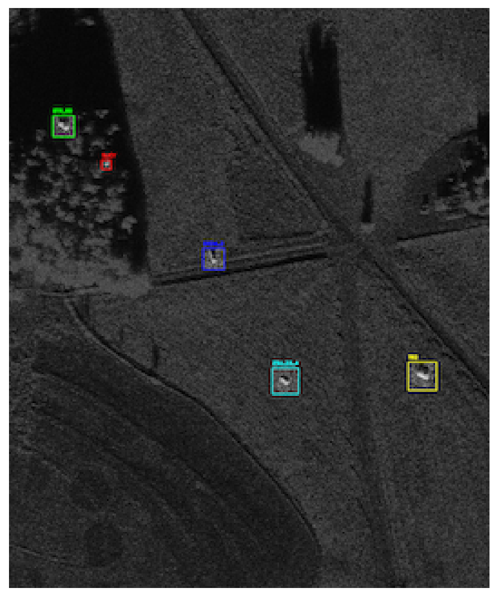

Fig. 15. Expected SAR Image sample with targets and labels.

## V. Conclusions

In this study, a novel approach for target detection and classification in Synthetic Aperture Radar (SAR) images was proposed, utilizing convolutional neural networks (CNNs) and transfer learning techniques. The SARNet and VGG models were employed for target classification, while a CNN network was used for target detection. The proposed approach was evaluated on the MSTAR and MathWorks datasets, demonstrating accurate and efficient target detection and classification. Further improvements can be made by expanding the training dataset, exploring different transfer learning techniques and model architectures, and optimizing hyperparameters. Overall, the proposed approach holds promise for SAR image analysis in various fields and can be extended for more complex datasets.

## REFERENCES


[1] A. Moreira, P. Prats-Iraola, M. Younis, G. Krieger, I. Hajnsek, and K. P. Papathanassiou, "A tutorial on synthetic aperture radar," IEEE Geosci. Remote Sens. Mag., vol. 1, no. 1, pp. 6-43, Mar. 2013.

[2] O. Kechagias-Stamatis and N. Aouf, "Automatic target recognition on synthetic aperture radar imagery: A survey," IEEE Aerospace and Electronic Systems Magazine, vol. 36, no. 3, pp. 56-81, Mar. 2021.

[3] X. Hua, Y. Ono, L. Peng, Y. Cheng, and H. Wang, "Target detection within nonhomogeneous clutter via total Bregman divergence-based matrix information geometry detectors," in IEEE Transactions on Signal Processing, vol. 69, pp. 4326-4340, Jul. 2021.

[4] J. Liu, M. Xing, H. Yu, and G. Sun, "EFTL: Complex convolutional networks with electromagnetic feature transfer learning for SAR target recognition," IEEE Trans. Geosci. Remote Sens., to be published, doi: 10.1109/TGRS.2021.3083261.

[5] D. E. Dudgeon and R. T. Lacoss, "An overview of automatic target recognition," Lincoln Lab. J., vol. 6, no. 1, pp. 3-10, Jan. 1993.

[6] X. Zhu et al., "Deep learning meets SAR: Concepts, models, pitfalls, and perspectives," IEEE Geosci. Remote Sens. Mag., to be published, doi: 10.1109/MGRS.2020.3046356.

[7] S. Feng, K. Ji, X. Ma, L. Zhang, and G. Kuang, "Target region segmentation in SAR vehicle chip image with ACM net," IEEE Geosci. Remote Sens. Lett., to be published, doi: 10.1109/LGRS.2021.3085188.

[8] X. Ma, K. Ji, L. Zhang, S. Feng, B. Xiong, and G. Kuang, "An open set recognition method for SAR targets based on multitask learning," IEEE Geosci. Remote Sens. Lett., to be published, doi: 10.1109/LGRS.2021.3079418.

[9] E. R. Keydel, S. W. Lee, and J. T. Moore, "MSTAR extended operating conditions: A tutorial," in Proc. Algorithms SAR Imagery, 3rd SPIE Conf., 1996, pp. 228-242.

[10] H. Chiang, R. L. Moses, and L. C. Potter, "Model-based classification of radar images," IEEE Trans. Inf. Theory, vol. 46, no. 5, pp. 1842-1854, Aug. 2000.

[11] K. E. Dungan and L. C. Potter, "Classifying transformation-variant attributed point patterns," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 43, no. 11, pp. 3805-3816, Nov. 2010.

[12] K. Fu, F. Dou, H. Li, W. Diao, X. Sun, and G. Xu, "Aircraft recognition in SAR images based on scattering structure feature and template matching," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 11, no. 11, pp. 4206-4217, Nov. 2018.

[13] B. Ding, G. Wen, C. Ma, and X. Yang, "An efficient and robust framework for SAR target recognition by hierarchically fusing global and local features," in IEEE Transactions on Image Processing, vol. 27, no. 12, pp. 5983-5995, Dec. 2018.

[14] S. Chen, H. Wang, F. Xu, and Y. Jin, "Target classification using the deep convolutional networks for SAR images," in IEEE Transactions on Geoscience and Remote Sensing, vol. 54, no. 8, pp. 4806-4817, Aug. 2016.

[15] R. Shang, J. Wang, L. Jiao, R. Stolkin, B. Hou, and Y. Li, "SAR targets classification based on deep memory convolution neural networks and transfer parameters," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 11, no. 8, pp. 2834-2846, Aug. 2018.

[16] Z. Huang, Z. Pan, and B. Lei, "Transfer learning with deep convolutional neural network for SAR target classification with limited labeled data," in IEEE Transactions on Geoscience and Remote Sensing, vol. 55, no. 3, pp. 1436-1449, March 2017.

[17] A. Krizhevsky, I. Sutskever, and G. Hinton, "Imagenet classification with deep convolutional neural networks," in Advances in Neural Information Processing Systems, 2012, pp. 1097-1105.

[18] Z. Ying et al., "TAI-SARNET: Deep transferred atrous-inception CNN for small samples SAR ATR," in Sensors, vol. 20, no. 6, p. 1724, Mar. 2020.

[19] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in Proceedings of the International Conference on Learning Representations (ICLR), San Diego, CA, USA, May 2015, pp. 1-14.