ENG 4001: Research Project

Faculty of Science, Engineering and Technology

Machine learning techniques for radar target recognition

Progress Report

Project ID: 2022s2-EEE-Ng-UG-63280

Author	
Name	**ID**
Duc Thang Nguyen	a1831044

Oct 24, 2022

Nomenclature

ISAR – Inverse Synthetic Radar

ATR – Automatic Target Recognition

CNN – Convolutional neural network

CM – Confusion matrix

Executive Summary

The project applies machine-learning approaches to radar target recognition. Several deep-learning-based techniques have been widely used in various areas, and the Convolutional neural network (CNN) is one of the most prominent approaches. In the radar community, deep learning applications in automatic target recognition (ATR) have been proposed, and the use of CNN has outstanding results in detecting and classifying targets. Additionally, the complex motions of real-world targets encourage the use of ISAR images for target recognition. This project aims to investigate and compare different deep-learning algorithms on ISAR images for ATR and how additional information about the target can affect the model's classification performance.

Background research was conducted to attain current knowledge from the industry, and potential contributions to the areas were identified. The literature reviews have been completed, critical information is included in the document from which the project aims are derived, and technical objectives are set to frame the project's progress. In the initial state, the project aim is to replicate findings in a published article to validate the findings and find space for improvement. Deep learning is a powerful tool in ATR, but deep learning applications have not been profoundly surveyed. In most research, deep learning takes place when the database has been collected through experiments or generated with simulations. Hence, to investigate deep-learning approaches, the project will concentrate on these stages of implementation where the deep-learning is utilized for classification,

For research objectives, the general analysis of radar imaging has been conducted and included in the literature review with basic information about the current well-known deep-learning techniques and ATR. The basic principle of target recognition based on radar scattering is also discussed. CNN's applications in ATR have been surveyed with various techniques, and potential gaps are summarized for further research opportunities. The implementation objectives are partly completed in which the database of four commonly found vehicles with different clutter conditions are analyzed and classified by a well-known CNN named Google net. The classification method and experiment plan are illustrated with experimental results reported by Confusion matrices. Finally, the completion plan has been conducted to frame future works for the second semester. The repetition of the completed work on different CNNs would be performed to survey their classification performances.

Contents

1. Introduction

In the radar industry, automatic target recognition (ATR) is the primary goal of any scouting system. This work applies several well-known deep-learning techniques to classify targets from Inverse Synthetic Radar (ISAR) images. The work simplifies the process by only recognizing a single target on each image, and all images have the same top view projection angle. An ISAR dataset of five commonly found types of vehicles, namely a bicycle, an autorickshaw, a truck, a mid-size car, and a full-size car, is used for experiment over several different deep learning approaches to survey their performances.

1.1 Background

In recent years, machine learning has granted great attention from the industry. The outstanding performance of the convolutional neural network (CNN) in optical image recognition [1] brings its potential to the radar industry. CNN in radar target recognition has also been studied for multiple purposes. In areas such as detection [3], [6], [7], data augmentation [4], and classification [7]-[9], different applications of CNNs are studied and verified with high accuracy and efficiency. Inverse synthetic radar (ISAR) is a radar technique for imaging targets that relies on rotating motion. Most ISAR applications are for ATR since they contain rich information about the target, such as its size, type, and position. ISAR image frame is generated based on range-Doppler operations. Range resolution produces one imaging dimension, and the orthogonal (cross-range) dimension is given by the differential Doppler resolution from target rotation. Deep-learning approaches have also been applied to classify ISAR images. Several studies have considered using this approach for target recognition [10]-[12]. This approach works well in extreme conditions with various signal-to-noise ratios (SNR). A more complex system using 3D-ISAR for point cloud formation to enrich the information of the target is performed in [13]. This project is a simplified version of [13] with the dataset provided by the work of [11].

1.2 Project aims and scope

The project aims to contribute value to the radar community by validating a previous finding from different approaches. On the one hand, this project acts as a survey of different well-known CNNs on a given dataset following a simplified version of an efficiency-verified method for findings validations. On the other hand, repeating the same experiment on the same dataset but with different deep-learning algorithms and different features from the dataset potentially archives accuracy improvement and the effect of the target’s known features on classification performance.

The study is accomplished in 3 stages, starting with a research-based stage followed by two implementation stages. In the first stage, fundamental analysis is conducted to acquire a general knowledge of the fields and the advantages of current applications. This stage results in the literature review of this document, with several related works highlighted. In the second stage, the dataset is experimented with by one deep-learning approach, and different aspects of the dataset are examined. The model built up from this stage is then applied to all other algorithms. The final stage is the repetitions of the previous one on all other algorithms with accuracy extracted for comparisons.

1.3 Technical objectives

The SMART approach is applied in defining five technical objectives to archive project aims. By this, each objective must be Specific and Measurable with a realistic goal that can be Archivable and Relevant to the project aims. Time is also a valuable resource, so Time-bound requirements for the objective must also be guaranteed. Five technical objectives are summarized:

i. **Research and systematically document fundamental knowledge of the field and related studies of ATR**
   The first objective is to build up the fundamental knowledge of radar principles and be aware of critical studies in the area. General information on radar imaging, ATR and deep-learning techniques is acquired, and technology selection is made and reported in the literature review.

ii. **Data analysis and conversion**
   In this objective, general dataset analysis is conducted, resulting in detailed descriptions of the dataset. As the project involved classifying ISAR data by deep-learning technique, the data must be converted to an appropriate form that the network can read before being trained for classification. Therefore, this object is essential to verify the preparation for experiments.

iii. **Build up experiment plan**
   Before any experiment is conducted, a plan is needed to navigate the experiment to archive the project aims and keep track of the experiment's progress. The plan must analyze the specific aim and frame the experiment to clarify the research directions step-by-step.

iv. **Dataset investigation by one CNN**
   Objective 4 is the first implementation objective in which the first CNN is trained and classified with the dataset following the plan defined in the previous objective. Every action in this objective is then analyzed and documented to frame the experiment process and adjust the plan if applicable. This document later reports the classification result by Confusion Matrices (CM).

v. **Survey of different CNNs performance**
   The final objective is the repetition of objective 5 but in a broader concept in which several CNNs are respectively applied on the given dataset to compare the performances to analyze the effect of each algorithm on the ISAR images and work out the most suitable one for further research in the area.

1.4 Document Overview

This document reports the project progress to the current date, and the plan for completion is also conducted to ensure the project's outcome is met, and all aims are archivable. Section 2 outlines the relevant research evidence to the field and summarizes potential research gaps. The experiment method is described in section 3, followed by fundamental theories detailed in section 4, and experiment results are reported in section 5 of the document. Section 6 illustrates the completion plan for the rest of the project, and the document's brief conclusion is summarized in section 7.

2. Literature review

Automatic target recognition has been widely studied for over decades. Traditionally, target recognition is based on geometric features, such as length, width, derived from the high-resolution range profile of the target. In recent years, applications of deep-learning techniques in target detection [3], [6], [7], classification [9]-[13] and data augmentation [4] have made tremendous progress. The reliable performance of deep learning in ATR has widened its community within the radar industry. For this reason, the analysis of potential deep-learning applications in general and in ATR has been conducted by reviewing relevant literature.

2.1 Deep learning approaches

ImageNet Large Scale Visual Recognition Challenge (ILSVRC) is a challenge to evaluate algorithms' performance in object detection and image classification for optical images. In recent years, deep-learning algorithms, specifically CNN, dominated the challenge with high performance, such as Alex net for first place in 2012, Google net and Res Net respectively won the 2014 and 2015 challenge.

2.1.1 GoogleNet

In the work proposed by [16], a deep convolutional neural network, which won the ILSVRC-2014 challenge with 6.7% error, is designed with increased depth and width but remains a constant computational budget. The codename for CNN is Inception, a 22-layer-deep convolutional network with the main idea of investigating how to estimate an ideal local sparse structure of a convolutional vision network. In Fig. 1, 1x1 convolutions are utilized to compute reductions before the more expensive 3x3 and 5x5 convolutions. The network included nine Inception modules and employed average pooling rather than fully linked layers for classification.

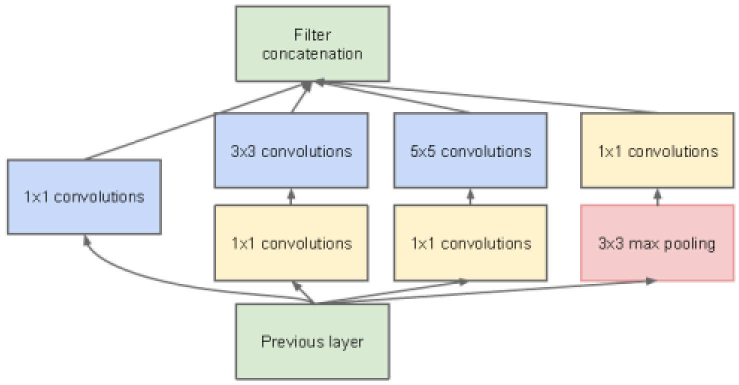

Figure 1 Inception module with dimensionality reduction [16]

2.1.2 AlexNet

AlexNet algorithm is presented in [17] with the classification error of 16.4% in ILSVRC-2012. The model uses the dataset from ILSVRC- 2010 contest for testing and validating the model, and the error rate of top-1 and top-5 in ILSVRC- 2010 are archived before being submitted for ILSVRC-2012. The network consists of eight weighted layers, five convolutional layers, three fully connected layers, a final 1000-way SoftMax, and three max-pooling layers with a total of 60 million parameters. The applications of ReLu non-linearity are available in the output of every convolutional and fully connected layer. The first convolutional layer includes 96 kernels of size 11x11 with a stride of 4 pixels. The output of the first convolutional layer is filtered with 256 kernels of size 5x5, which is the architecture of the second convolutional layer. The third, fourth and fifth convolutional have the same 3x3 size kernels with the number of kernels 384, 384 and 256, respectively. Figure 2 below shows the architecture of the Alex Net CNN.

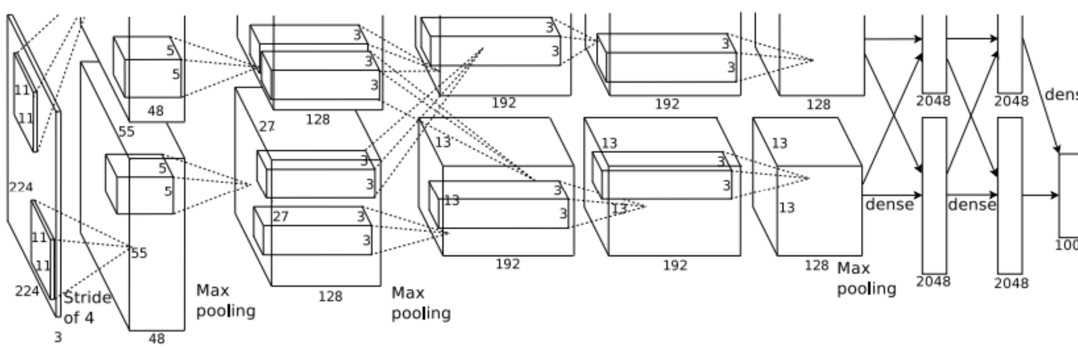

Figure 2 Alex Net architecture [17]

2.1.3 ResNet

Microsoft's Residual Networks (ResNet) [18] archived an incredible error rate of 3.6% in ILSVRC-2015 competition. Resnet is an extremely deep network with 152 layers depth which is 8 times deeper than VGG Net [19] but lower complexity compared to VGG net. The model is evaluated the dataset that contains 1000 classes with 1.28 million images used for training and 50 thousand images for validating and finally tested with 100 thousand test images. Fig.3 below depicts a building block of residual learning. The outputs of the shortcut connections are simply added to the outputs of the stacked layers after identity mapping is performed.

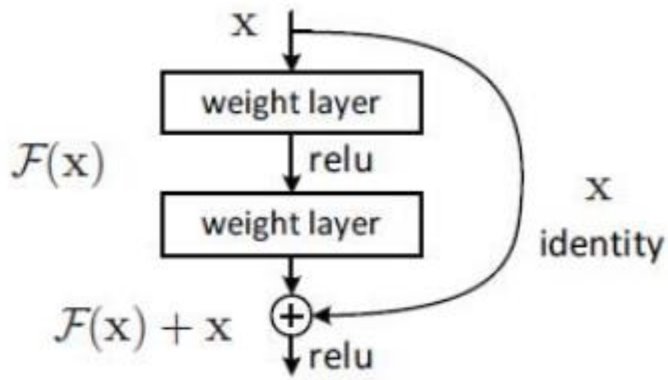

Figure 3 Residual learning building block [18]

The plain network of Res Net is inspired by the philosophy of VGG Net. Most of the convolutional network have kernels size 3x3 ended by a global average pooling layers and 1000-way fully connected layer with SoftMax. A residual network in ResNet also has same architecture but a short cut connection between each pair of 3x3 kernels is added. They employed a modified building block as a bottleneck design for the deeper nets as show in Fig 4.

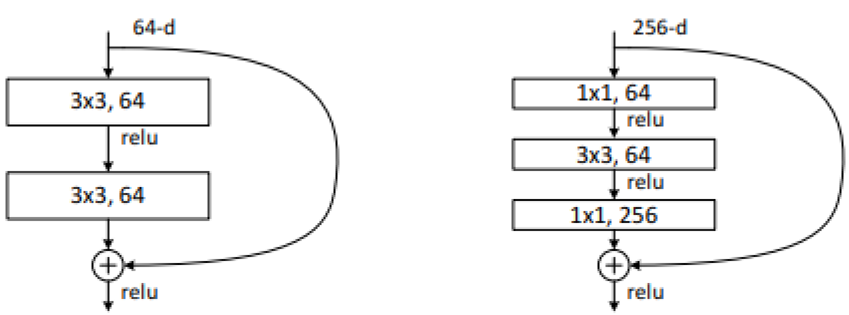

Figure 4 A deeper residual function [10]. Left: a building block for ResNet-34.

2.2 Deep learning in automatic target recognition

Due to the progressive development of deep learning techniques in target detection and classification in optical images, deep-learning techniques are widely applied in automatic target recognition. Despite of differences of radar images and optical images, deep learning remains it constant performance. Several deep-learning applications in ATR are analyzed below.

In the work conducted by [10], a deep adaptive learning technique is performed on a computationally limited platform to recognize the ISAR target. Bin Xue et al [10] proposed a deep multimodal mechanism to tackle the real-world modal diversity problem with different components named feature sampling, feature extraction, deformation handling and classification are jointly learned to handle the complex target recognition. A specific deep adaptive learning model called DOR_ISAR is applied and compared with 7 other deep-learning techniques and DOR_ISAR have the highest accuracy with lowest runtime. However, the context of comparison and dataset description is not clear and the fact that it is implemented on other platform so that the practicality of the comparison may not be guaranteed.

In [12], the simulation framework to generated realistic ISAR images is conducted and verified with measured data. Over 30000 ISAR images of five commonly found vehicles with different types of motions and clutter conditions are successfully generated and published to the community. The dataset is also tested for classification by two classical machine-learning approaches and two recent deep-learning approaches with the performance of deep-learning approaches outweighing the classical counterpart. The classification approaches are only used to validate the dataset and are not fully investigated into each approach's performance.

A 3-D ISAR technique classified using CNNs is introduced in [12], a 3-D ISAR image formation technique and an approach for classification are illustrated. A simple CNN with a single convolution layer is trained with 3-D simulated ISAR images before classifying both simulation data and real data. The classification performance is high since 3D-ISAR images enrich the information that can be derived from the target which severely improves the classification performance.

2.3 Summary and potential gaps

In summary, the performances of deep-learning techniques in recognizing optical images are outstanding applications of deep-learning in ATR has not been widely surveyed. The plenty number of radar dataset available provides huge potential for testing and validating the potential of machine-learning techniques in this area. Therefore, this could be the opportunity for research to be conducted.

3. Implementation methodology

The implementation is accomplished in 2 stages which are the preparation stage and the experiment stage. The preparation stage starts with a deep dataset analysis to understand what will be examined. The conversion of the dataset is necessary since the input layer of the CNN requires a specific type of data. The plan is also a part of the preparation stage. Training and testing come after preparation, and the implementation is officially started from this stage.

3.1 Data description

In this work, we make use of the simulation dataset released by [11]. The dataset consists of over 30000 realistic ISAR images validated with measurement data provided by Texas Instrument's AWR 1843 77GHz automotive radar. The scatterers of five commonly found vehicles, a bicycle, an Autorickshaw, a truck, a mid-size car, and a full-size car, are simulated. All vehicles are assumed to undergo four types of traffic trajectories at an intersection. The ground is aligned with the XY plane of the system with the axis aligned with lanes coming from the North (N) – South (S) direction and West (W) – EAST (E) direction, respectively, as shown in Fig 1. The height is along the Z-axis, and an ego radar is assumed to be held at the point (0, 0,0.5) m along the south road. 16 possible trajectories could be performed in this intersection (shown in Fig 6) with four left turns, four right turns, four U-turns, and four straight paths. All trajectories are simulated at a specified speed with different clutter conditions. The noise and clutter models are presented in [11] with the ground-based clutter in the range dimension and wind-based clutter in the Doppler dimension of the ISAR images realistically simulated. ISAR images of all trajectories are generated with 5 different signal-to-noise ratios (SNR) varying from -5dB to 10dB and 4 different wind velocities, 2.5 MPs, 5 MPs, 7.5 MPs, and 10 MPs, for Doppler clutter simulation.

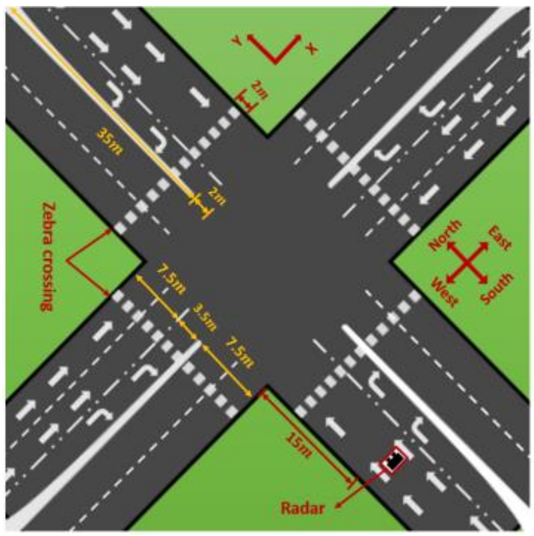

Figure 5 Road geometry of the intersection

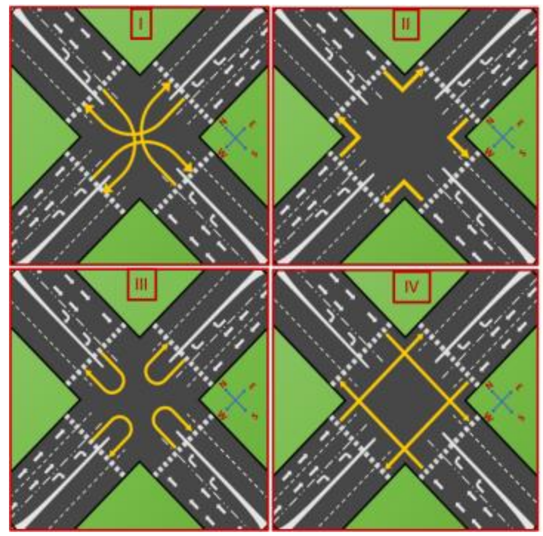

Figure 6 Trajectories undertaken by the automotive target in a four-way junction - (i) Right turn, (ii) Left turn, (iii) U-turn and (iv) Straight through.

3.2 Data conversion

The ISAR dataset provided by [11] is the targets' complex range and cross-range profile, which cannot be directly fed into the deep-learning model. The model requires the data to be in image format before the classification can be accomplished; therefore, the conversion step is needed. The ISAR data is converted to the 8-bit greyscale image, with each pixel being the normalized dB scale magnitude of the aligned range - cross-range profile.

3.3 Experiment plan

There are two primary purposes for conducting this experiment. The first purpose is to profoundly investigate the dataset in which the effect of each type of trajectory is examined to verify whether the CNN trained with one type of motion can recognize the targets undergoing different types of motion and which type of motion provides the most valuable information for classification. Moreover, the experiment also verifies the question of how more types of target motion are "learned" and can affect the classification performance. This is a crucial question to verify since real-world targets undergo numerous complex movements, not all of which can be captured and trained, so the verification of whether the model is trained with fewer motions can perform as well as the one with more trained motions. The second purpose is to survey other CNNs performances to find the most suitable CNN for this system.

i. Randomly separate the dataset to 70% for training and 30% for validating
ii. Train the CNN with one type of target trajectory (left turn) to classify the 30% validating data
iii. Repeat ii for other types of target trajectory (right turn, U-turn, straight)
iv. Train the CNN with 2, 3, and 4 types of the target trajectory
v. Repeat steps ii to iv with other CNNs

4. Theory

4.1 Radar principle

Generally, radar is an electrical system that emits electromagnetic waves with radio frequency (RF) toward interest regions and measures the reflection waves from the object detected in the wave paths. A radar system generally consists of four subsystems: a transmitter, antenna, receiver, and signal processor. The transmitter and receiver can be both connected to the antenna through a transmit/receive device or connected to a separate antenna, depending on the system. The transmitter generates the EM waves and inputs them to the antenna to introduce them to the propagation medium. The wave then propagates to the target and induces currents on the target, which reradiates these currents into the environment. The receive antenna then receives the reradiated signal. However, not all legitimate received signals are from the desired target, the surrounding environment can also reradiate the signal to the receive antenna, and those unintentional signals received are called *clutter*. Fig 7 below illustrates the essential elements of a radar system.

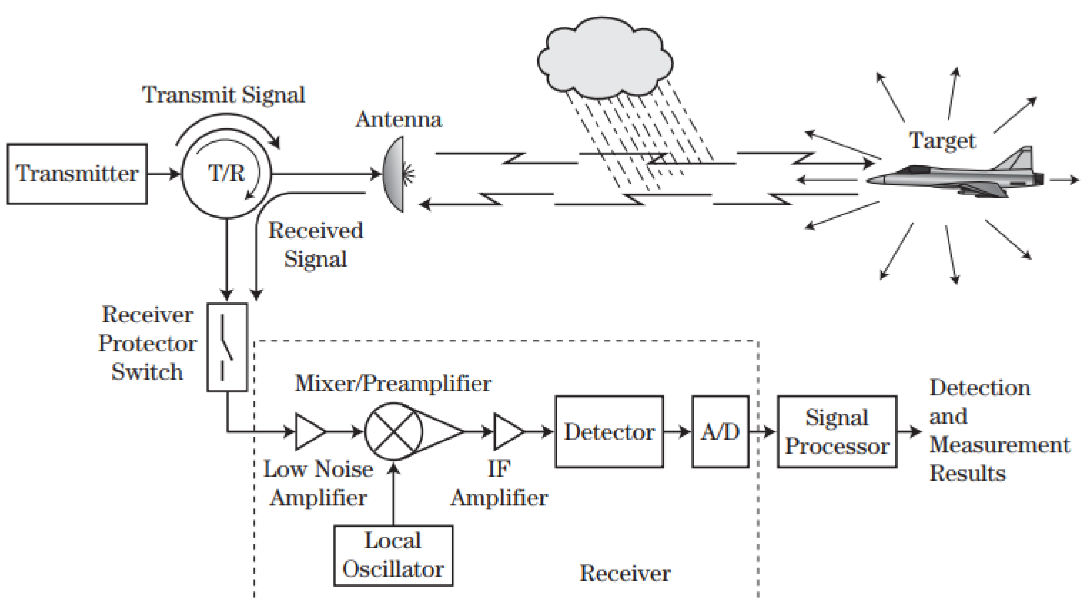

Figure 7 Major elements of the radar transmission/ reception process.

4.2 Inverse synthetic aperture Radar (ISAR)

Inverse synthetic aperture Radar (ISAR) is a radar signal processing technique that can provide a 2-dimensional electromagnetic image of the area of interest. ISAR inherits any radar-based imaging techniques that it can be employed in all weather and day/light conditions. ISAR images are obtained by coherently processing the received signals. Range-direction resolution is dependent on the transmitted pulse's bandwidth while using a synthetic antenna aperture can acquire high cross-range resolution.

In SAR, the synthetic aperture is based on the motion of the radar platform, while ISAR synthetic aperture, in contrast, is generated by the assumption of moving targets. Therefore, the name ISAR is initially derived from SAR since it is the inverse version of SAR principle. However, today ISAR synthetic aperture can possibly be generated by the radar platform motion, so the noncooperation of the ISAR target defines the primary difference between SAR and ISAR. ISAR images' formation starts with motion compensation to deal with highly nonstationary signals, followed by range and cross-range compression. Range compression is done by compressing the time-domain signals or multifrequency signals by inverse Fourier transform to produce a complex range profile. For cross-range compression, Furrier transform is only applied when targets are under smooth motion or integration time is short enough while fast maneuvering targets or sea-driven motioned ships or demand of high resolution require different techniques such as the JTFA, the range-instantaneous Doppler (RID), the enhanced image processing (EIP), and several tomography-based techniques and super-resolution techniques, such as CLEAN technique, and the Capon technique. The combination of range and cross-range compression accomplishes the ISAR image formation process.

5. Results

Following the defined plan, the model has been trained with one type of trajectory and repeated four times for four different trajectories and tested with the validating data. The results are illustrated by confusion matrices attached in the appendix. The overall classification accuracy of the Google net trained by images of the targets making a left turn, right turn, straight path, and U-turn are respectively 92.48%, 84.58%, 80.88%, and 84.58 %. The performance of the model trained by different types of trajectories does not vary much, which can inform that the information that can be "learned" from different trajectories is nearly the same, and targets can be recognized with relatively high accuracy by Google net with only one type of trajectory trained. The classification results also show that the classification process can be affected by the size of the targets since the targets with the size relative equal are getting confused with each other, such as the full-size car and the mid-size car, which have the highest amount of misclassification, or the full-size car and the autorickshaw which also have a relatively high amount of misclassification.

6. Completion Plan

As the experiment plan is partly completed, the future work aims to continue the incomplete parts of the plan. For the 12 remaining weeks of the project timeline, 2 of those are spent to continue in dataset investigation in which 2, 3, and 4 types of trajectories will be trained for the CNN and repeat the classification process. After that, the following week will be for CNNs analysis to pick out three other CNNs to be surveyed, and 6 weeks is the duration for those 3 CNNs to be systematically investigated and results documented. The last three weeks will be used to finalize documents and thesis preparations.

7. Conclusion

This document provides detailed information on the project's progress. Background information, project aims and scope, and technical objectives are introduced with clear explanations and a literature review to locate the project contribution within the area. The experimental framework has also been formalized with results reported in the document. Finally, the plan for future work is practically conducted with specific requirements.

References

[1] M. Pak and S. Kim, “A Review of Deep Learning in Image Recognition,” in 2017 4th International Conference on Computer Applications and Information Processing Technology (CAIPT). IEEE, 2017, pp. 1–3

[2] R. Wu, S. Yan, Y. Shan, Q. Dang, and G. Sun, “Deep Image: Scaling up Image Recognition,” ArXiv Preprint ArXiv:1501.02876, vol. 7, no. 8, 2015

[3] L. Wang, J. Tang and Q. Liao, "A Study on Radar Target Detection Based on Deep Neural Networks," in IEEE Sensors Letters, vol. 3, no. 3, pp. 1-4, March 2019, Art no. 7000504, doi: 10.1109/LSENS.2019.2896072.

[4] J. Ding, B. Chen, H. Liu and M. Huang, "Convolutional Neural Network With Data Augmentation for SAR Target Recognition," in IEEE Geoscience and Remote Sensing Letters, vol. 13, no. 3, pp. 364-368, March 2016, DOI: 10.1109/LGRS.2015.2513754.

[5] Wan, J., Chen, B., Xu, B. *et al.* Convolutional neural networks for radar HRRP target recognition and rejection. *EURASIP J. Adv. Signal Process.* **2019**, 5 (2019). https://doi.org/10.1186/s13634-019-0603-y

[6] J. Williams, L. Rosenberg, V. Stamatescu, and T.-T. Cao, “Maritime Radar Target Detection Using Convolutional Neural Networks,” in IEEE Radar Conference.

[7] M. Upadhyay, S. K. Murthy and A. A. B. Raj, "Intelligent System for Real time detection and classification of Aerial Targets using CNN," 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS), 2021, pp. 1676-1681, doi: 10.1109/ICICCS51141.2021.9432136.

[8] Z. Liu, M. Waqas, J. Yang, A. Rashid and Z. Han, "A Multi-Task CNN for Maritime Target Detection," in IEEE Signal Processing Letters, vol. 28, pp. 434-438, 2021, doi: 10.1109/LSP.2021.3056901.

[9] M. Ma, J. Chen, W. Liu, and W. Yang, “Ship classification and detection based on CNN using GF-3 Sar Images,” *MDPI*, 14-Dec-2018. [Online]. Available: https://www.mdpi.com/2072-4292/10/12/2043#cite. [Accessed: 21-Oct-2022].

[10] B. Xue, W. Yi, F. Jing, and S. Wu, “Complex Isar target recognition using deep adaptive learning,” Engineering Applications of Artificial Intelligence, vol. 97, p. 104025, 2021.

[11] N. Pandey and S. S. Ram, “Classification of automotive targets using inverse synthetic aperture radar images,” *IEEE Transactions on Intelligent Vehicles*, pp. 1–1, 2022.

[12] S. Musman, D. Kerr, and C. Bachmann, “Automatic recognition of isar ship images,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 32, no. 4, pp. 1392–1404, 1996.

[13] C. Pui, B. Ng, L. Rosenberg and T. Cao, "Target Classification for 3D-ISAR using CNNs".

[14] R. M. A. (ed.), S. J. A. (ed.), and H. W. A. (ed.), “Chapter 1 . Introduction ,” in *Principles of modern radar: Vol I: Basic principles*, Raleigh, NC: SciTech Pub., 2010, pp. 4–5.

[16] C. Szegedy et al., "Going deeper with convolutions," 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, pp. 1-9, doi: 10.1109/CVPR.2015.7298594.

[17] A. Krizhevsky, I. Sutskever, and G.E. Hinton, "ImageNet classification with deep convolutional neural networks," in Advances in neural information processing systems, pp. 1097-1105, 2012.

[18] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770-778, 2016.

[19] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. In ICLR, 2015

Appendices

Classification results by confusion matrix

1. Google net is trained with images of targets turning left

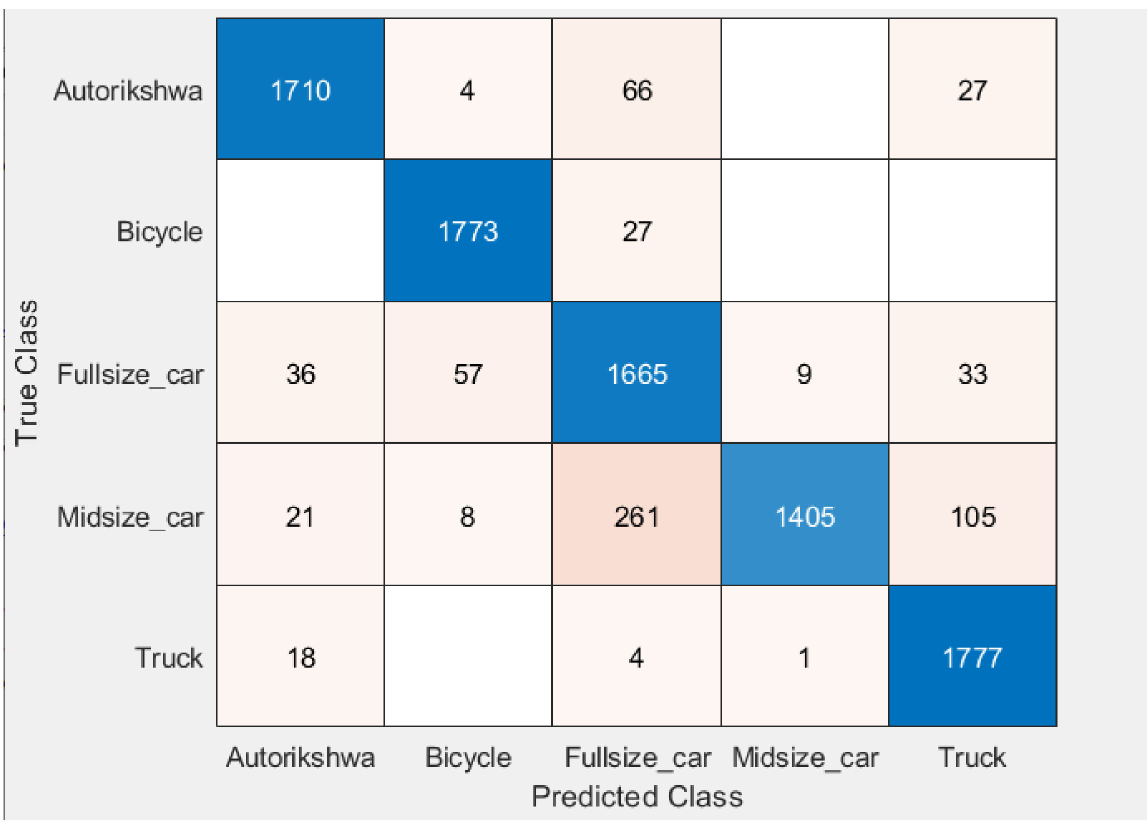

Figure 8 Classification results when Google net is trained with images of turning left trajectory

2. Google net is trained with images of targets turning right

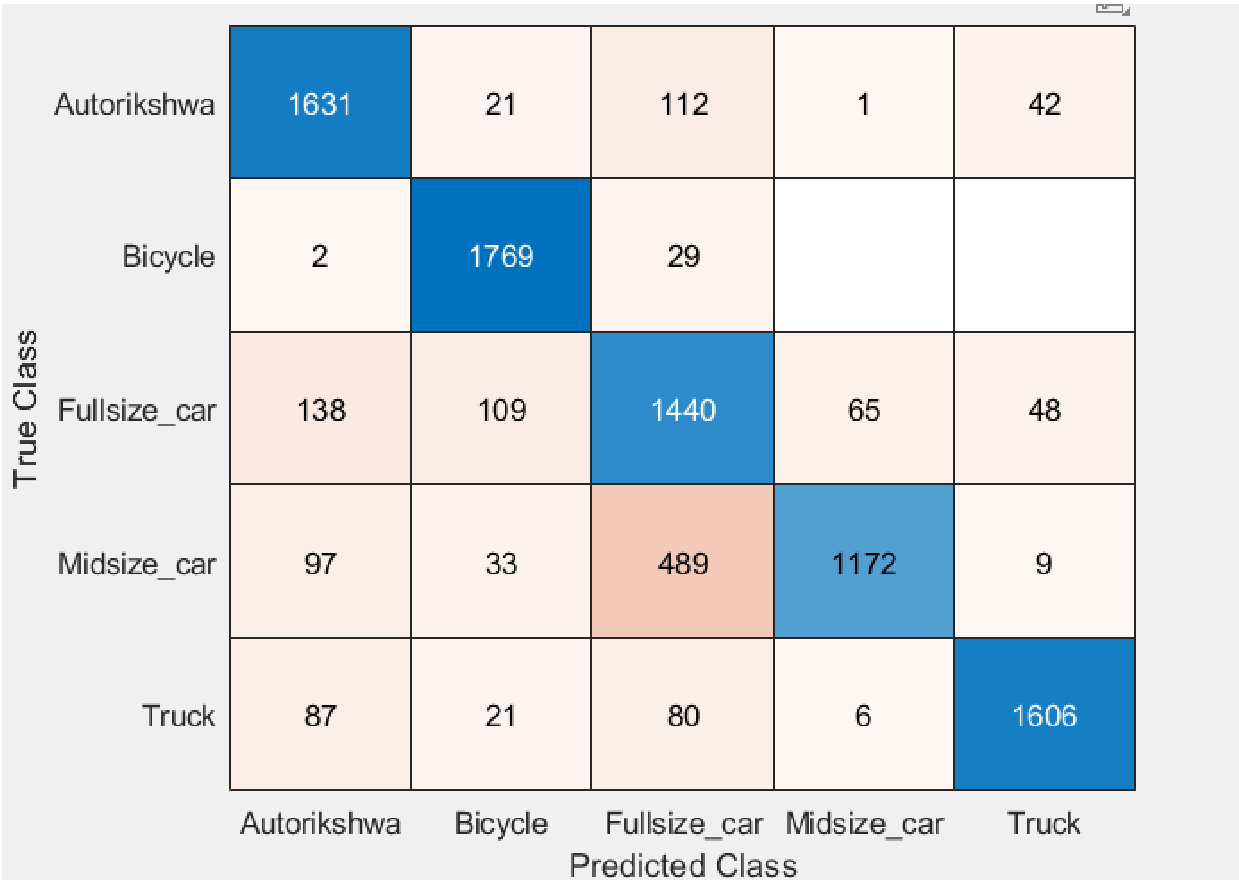

Figure 9 Classification results when Google net is trained with images of turning right trajectory

3. Google net is trained with images of targets going straight

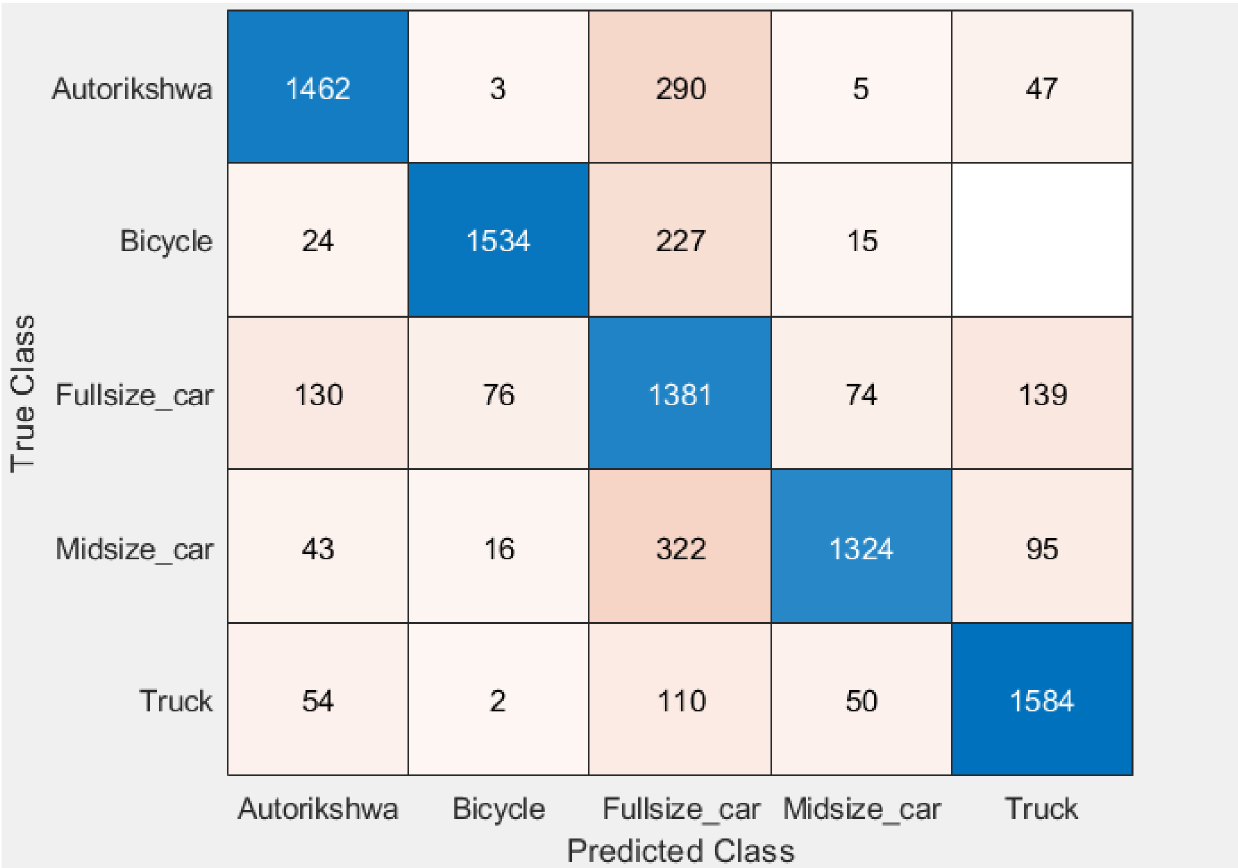

Figure 10 Classification results when Google net is trained with images of going straight trajectory

4. Googlenet is trained with images of targets making U-turn

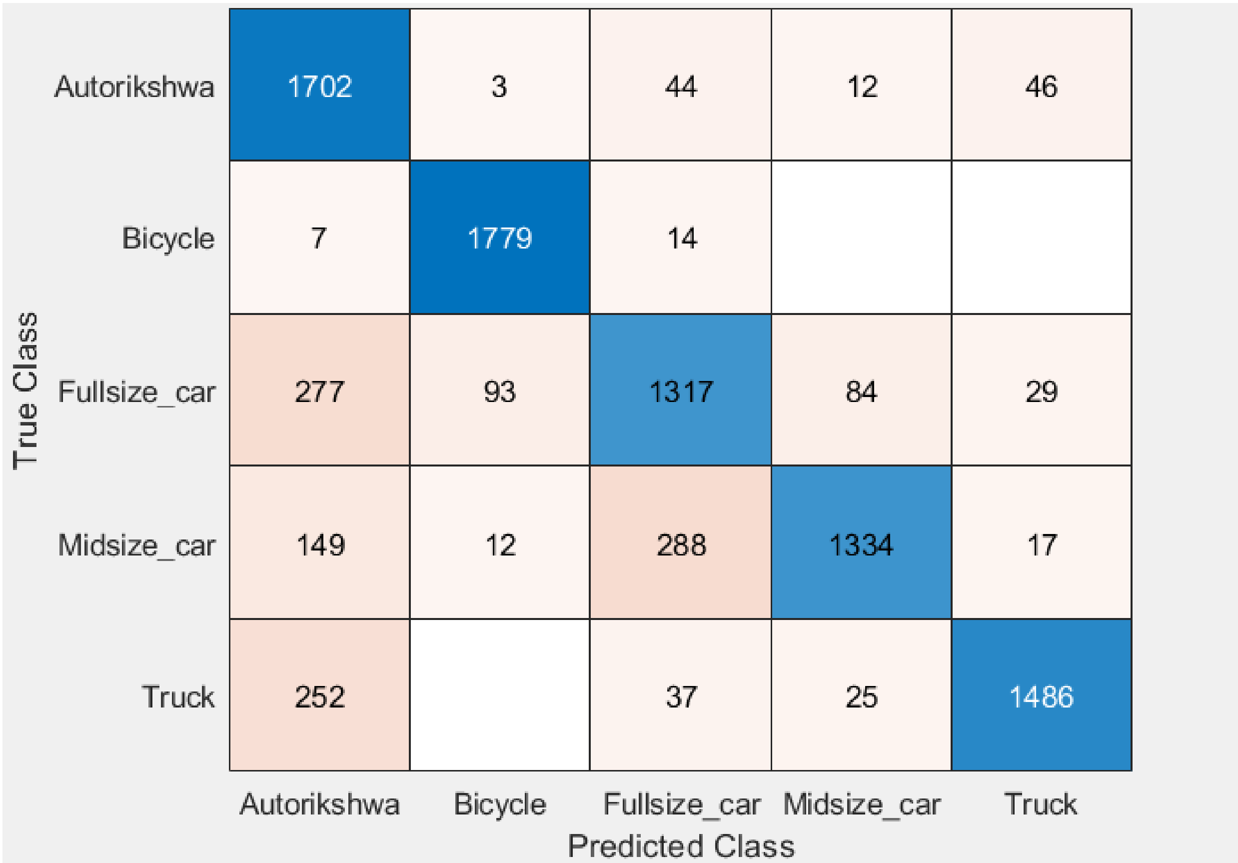

Figure 11 Classification results when Google net is trained with images of U-turn trajectory