Pakistan Journal of Engineering and Technology, PakJET
ISSN (p): 2664-2042, ISSN (e): 2664-2050
Volume: 6, Number: 1, Pages: 28- 33, Year: 2023

Radar Based Pedestrian Classification using Deep Learning Approach

Musawir Ghani*, Gulbadan Sikander, and Shahzad Anwar

Department of Mechatronics Engineering, University of Engineering & Technology, Peshawar, Pakistan
Corresponding author: Musawir Ghani (E-mail: musawirghani@uetpeshawar.edu.pk).

Received: 25/11/2022, Revised: 27/01/2023, Accepted: 07/02/2023

Abstract- **Developments in sensor fusion systems led to extensive research in autonomous vehicles, and object detection is a crucial aspect of vehicle operation. Detecting obstacles can be difficult due to the wide range of potential obstructions, the characteristics of each sensor, and the influence of the surrounding environment. In this paper, the authors use automotive radar data and various neural networks to classify various objects (vehicles, single and multiple people, and bicycles). Combined with the vehicle radar, this research proposed a rapid radar algorithmic implementation for locating, monitoring and extracting micro-Doppler. The authors evaluate three distinct neural network architectures for the five recorded classes of targets: the basic CNN, the residual network, and the combined neural architecture of convolution and recurrent layers. Considerable accuracy is 95.6% immediately before identification of the radar spectrogram from (~0.55 s to produce 0.5 s long spectrogram).**

Index Terms- Object Detection, Micro-Doppler, Spectrogram, Deep Learning, Radar.

I. INTRODUCTION

Pedestrians play a crucial role in the rapidly evolving transportation sector brought on by technological advancements. The self-driving cars need a sensor that can pick up on foot traffic [1]. For this reason, developing these new technologies must concentrate on pedestrian safety. This is especially true of autonomous driving, which has the potential to enhance the safety of pedestrians significantly. Daily, pedestrian accidents are among the leading causes of death in our increasingly dense traffic environment. One and a half million and thirty-five thousand people die in automobile accidents yearly; of these, 0.31 million are pedestrians [2]. As research in this vast field grows, more than 5,000 papers on pedestrian detection and trajectory have been published [3].

A major barrier in pedestrian detection and trajectory prediction planning is dealing with the complexity and variability of the real-world environment. Pedestrian behaviours can vary significantly based on factors such as age, culture, and personal preferences, making it difficult to predict their movements accurately**.** The task of planning a route can also be made harder by obstacles, other pedestrians, and vehicles. Artificial intelligence (AI) has revolutionized many fields, including pedestrian detection (see Fig. 1) and trajectory planning. In this context, Artificial intelligence (AI) is the ability of computers to carry out activities typically associated with human intellect.

FIGURE 1. A representation of pedestrian detection system in the outdoor environment [4]

These activities include pattern recognition, learning from experience, and adjusting to novel circumstances. In pedestrian detection and trajectory planning, AI algorithms are used to analyze video feeds or sensor data to identify and track the movements of pedestrians in real time. This information is then used to generate safe and efficient trajectories for pedestrians, vehicles, and other environmental objects (see Table I).

TABLE I
COMMON SENSORS FOR AUTONOMOUS VEHICLES

Sensor	***Method***
Sonar	Sound waves
Lidar	Laser beam
Radar	Radio waves
Camera	Visible Light

The ability to detect hazards in the road ahead and other vehicles, pedestrians, and animals are crucial for fully autonomous driving. Many other sensing methods have been proposed [5]. Depending on the image classification technique, cameras are relatively inexpensive against other types of sensors, and they're good for classifying objects based on colour and texture data. However, they have some drawbacks, such as a limited field of view and poor visibility in dim lighting.

However, even though a LiDAR system uses rotatable arrays of lasers to produce a 3-Dimensional representation concerning the surrounding area of a self-driving car, these sensors are still very costly. They may necessitate substantial computer power to compensate for the unpleasant effects of weather and sunlight (foggy, snowy, and rainy conditions).

Even though Radar sensors can be unaffected by environmental factors like light and weather, they can use existing techniques such as range-Doppler and classification processing [5]. However, there are still many open research problems, including how these methods may be adapted to the unique needs of the automotive industry and how the best processing can be created to combine data from many radar channels and heterogeneous sensors. Preserving the safety of pedestrians and other vulnerable road users has been the focus of extensive study of radar technology for vehicles.

This paper describes and explains a modular parallelizing technique for multiple moving detection and tracking, processing radar data in near real-time and classifying objects based on their characteristics. The study explored three distinct neural network architectures: a scaled-down version of VGG16 using the same building blocks as the incredibly extensive ResNET-50 layout; and a hybrid of the two[6]. It takes short cuts a unique combination of a convolutional neural network (CNN) and a long short-term memory (LSTM) architecture that links together network nodes that avoid overfitting and enhance generalization, that can learn a time series representation from feature extraction on MD spectrogram segments (sequences of data). In contrast to many prior attempts in the literature, the LSTM network component of this technique does not treat the radar data as a series of static spectrogram images [7, 8], on the other hand, as temporal data sequences. Instead of viewing the problem as picture classification, this method may be easily converted to radar data by using the information already there in a sequence of radar waveforms, as proven by the first findings on a small experimental dataset. The experimental samples employed in this work are meant to demonstrate this approach's viability, despite the dataset's small size. In addition to providing a lightweight radar signal processing implemented for radar medical diagnosis-based categorization, it offers a proof-of-concept assessment of several neural network architectures that can be used to extract features without the need for manual adjustment of external input parameters.

II. LITERATURE REVIEW

This section will give an overview of pedestrian detection algorithms and the types of algorithms used for pedestrian detection, including Haar cascades, HOG, and YOLO. A summary of the use of pedestrian detection in various applications such as surveillance, autonomous vehicles, and smart cities Challenges to pedestrian detection are discussed, such as occlusion, scale variation, and lighting conditions. Future and present developments in pedestrian detection research are discussed, focusing on deep learning-based methods and multi-modal fusion. A summary of the datasets and metrics used to assess pedestrian detection algorithm performance. An overview of the conclusions from the literature review and ideas for additional research.

There are some popular algorithms used for pedestrian detection. The choice of algorithm depends on the application's specific requirements, such as accuracy, speed, and computational resources.

A. Computer vision-based algorithms

1) Haar cascades: Uses Haar-like[9] features and Adaboost to detect pedestrians.

2) HOG (Histograms of Oriented Gradients): Extracts features based on the gradient orientation of the image and uses SVM for classification.

B. Deep Learning based algorithm:

1) Convolutional neural networks are used for classification in R-CNN (Regions with Convolutional Neural Networks) [10], which proposes object regions.

2) YOLO (You Only Look Once): A real-time object detection algorithm divides the image into grids and predicts each grid's bounding boxes and class probabilities [11].

3) Retina-Net: A one-stage object detection algorithm that uses focal loss to handle the class imbalance problem [12].

The technology is continuously evolving and finding new applications for pedestrian detection in various fields. Pedestrian detection is used in surveillance systems to monitor public spaces and detect suspicious behaviour. Pedestrian detection is a critical component in developing autonomous vehicles, helping them detect and avoid pedestrians on the road. Pedestrian detection is used in smart cities to improve pedestrian safety, manage traffic flow, and enhance the overall experience.

A.B. Khalifa et al. [13] proposed a study of an innovative framework for trajectory classification that can detect pedestrians in real-world environments where this is a significant challenge. To account for the shake introduced by camera movement, the proposed technique simulates the change in the backdrop between two frames. Then, it establishes a method of categorization that distinguishes between the image's foreground and background. With a clear foreground, it can easily pick out people walking around in the background. Using the publicly available benchmark dataset CVC 14, which comprises both daylight and evening visible and far infrared video sequences, this research verified the efficacy of the recommended approach. The recommended method effectively recognizes pedestrians' presence in the picture by capturing their dynamic nature between frames, as evidenced by experimental data. It needed a lot of computing labour to set up the experiment.

S. Rajendra et al. [14] recommended pedestrian detection and avoidance system for autonomous vehicles based on stereo vision. We employed a pair of cameras separated by a certain distance to take pictures of the area. The algorithm measures the distance to the pedestrian once it has identified it. The Automatic Emergency Braking System (AEBS) controller algorithm would engage AEB if the calculated distance was less than 3.3m. The experiment results show the suggested strategy's potential for improving prediction accuracy and lowering risk. The maximum speed during experimental testing is 30 km/h, the legal limit for highway travel. This research must ensure the technology works well at different velocities and weather.

T. Srinivas Reddy[15] presents that most existing methods might detect and track objects in a single camera's view or across several cameras. Its trackers are effective but need a lot of computer resources, making them inefficient. In this study, various classifiers are used to implement a technique for detecting pedestrians, and the resulting detection rates are high enough to use for at least some applications. The system monitors and records a wide range of human activities. The gradient system generates a histogram that is used to track the movement and acceleration of objects. Object-tracking apps would benefit greatly from this method as well. Unfortunately, the experimental design is inefficient and demanding of processing resources.

S. Mehai[16] uses a 16-layer Velodyne Puck LITE LiDAR and machine learning to detect pedestrians. The proposed mechanism linearly interpolates across layers to make up for the LiDAR's low resolution, essentially introducing 15 pseudo layers to aid in obtaining timely detection at practical distances. A Support Vector Machine (SVM) is used to categorize the potential pedestrians. The algorithm's accuracy is tested using real LiDAR frames acquired in various traffic settings.

In this study, this research offers linear interpolation to lower the threshold for item separation necessary for accurate detection and extraction. When there are no other items within a 0.38-meter radius of the LiDAR, the algorithm can detect a single object as far away as 21.54 meters.

Javad Enayati [17] suggested using a sensor fusion approach to deploy a tri-radar setup (long range, medium range, and short range) to spot objects of varying sizes when driving with advanced driver-assistance systems. All the usual suspects, such as cars, people, and animals, can be spotted in various contexts. The study proved that the proposed configuration could identify various targets in various traffic conditions, including pedestrians, small animals, and moving cars.

A comparison of the accuracy of three widely used categorization algorithms is shown in Table II. The VGG-like CNN-0.5s long dataset is the least accurate for car-person-2 classification and the most accurate for the long.

TABLE II.
TEST ACCURACY OF THE TWO-NETWORK ARCHITECTURE

Evaluation/ network type	**VGG-like CNN (0.5 s long datasets)**	**VGG- (2 s long datasets)**	**CNNLSTM (0.5 s long datasets)**	**CNNLSTM (2 s long datasets)**
Car-Person-2 people Classification	78%	81%	84%	80%
Car-Person bicycle Classification	83%	79%	83%	93%

III. METHODOLOGY

Numerous methods are identified for the detection and classification of pedestrians and bicyclists. CNN is being developed. We also chose CNN for pedestrian classification.

A. Dataset

We are using the helper functions helperBackScatterSignals and helper Doppler Signatures. We constructed a dataset with 40,000 signals: 20,000 pedestrians, 20,000 bicycles, and 12,500 automobiles. The data sets are created from the signals and the vehicle noise.

Micro-Doppler signatures were computed, and 5000 signatures were generated for each of the five scenes in the first data set by combining pedestrian and bicycle signals and adding Gaussian noise (without automobile noise).

As shown in Fig. 2, 80% of the signatures in each category (or 4000 signatures total) are put aside for the training data set, and 20% of the signatures (or 1000 signatures total) are set aside for the test data set.

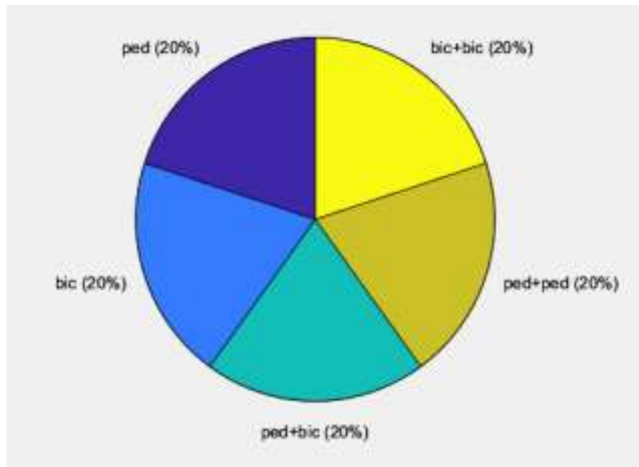

FIGURE 2. Distribution of five classes of data

B. Architecture Layers

(a) Fully connected layers.

It's directly connected to every layer below it, activation-wise. To make classifications, a completely connected layer is used.

b) Pooling layer

As a result, the image's quality drops, reducing the clarity of the translated message by softening the shift and distortion.

c) Convolutional layer

It serves as an edge, corner, and endpoint extractor.

C. Processing Steps

In Fig. 3, we can see a schematic of the various processing stages.

- Apply a fourth-order Butterworth transform [19] after performing a Fast Fourier Transform (FFT) on the raw digital data. High-pass filter with infinite impulse response and a cutoff of 0.04 Hz is used to filter out passive items (i.e., the presence of items with a Doppler signature at 0 Hz or near to it).

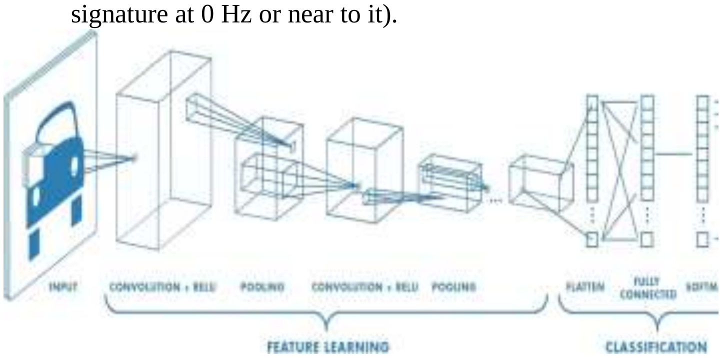

FIGURE 3. Components of CNN architecture [18]

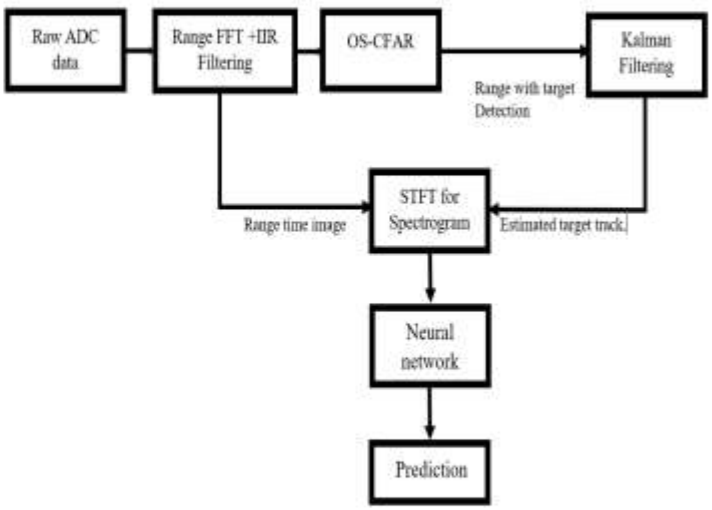

FIGURE 4. Proposed Methodology

- The goal is to target detection using the OS-CFAR algorithm with reduced noise and clutter (see Fig. 4).
- For a given frame, detect the targets' locations (the range bins they occupy) and save this information in a detection matrix.
- Frame-by-frame, the algorithm takes in the detection matrix using a hybrid of the Hungarian and constant-acceleration Kalman filtering. The former would generate a more accurate estimate of the target's location and keep outputting forecasts in the event of temporary frame loss or corruption. Using Kalman filter estimates, the latter would consistently label detected objects with their proper names. Using markers for each track, the system may additionally account for new objects entering or departing the radar's area of vision.
- Combining the target signature's location in the range bins where the object track position estimates are made with the range-time frames allows us to build MD signature segments. Based on the categorization method, the total length of the MD signature can be adjusted by stringing together different frames.
- Incorporate the MD spectrograms into the training and testing of neural network-based classifiers.

IV. RESULTS

The method employed to classify Pedestrians and Bicyclist separately and combined. Since the micro-Doppler signatures of pedestrians and bicyclists are different, it is easy to determine whether a given realization represents a pedestrian or a biker. Classifying many pedestrians or bicycles, especially in the presence of Gaussian noise or vehicle noise, is typically challenging.

The five categories having 20% signatures, including ped, bic, ped+bic, bic+bic, ped+ped spectrogram, were obtained (see Fig. 5).

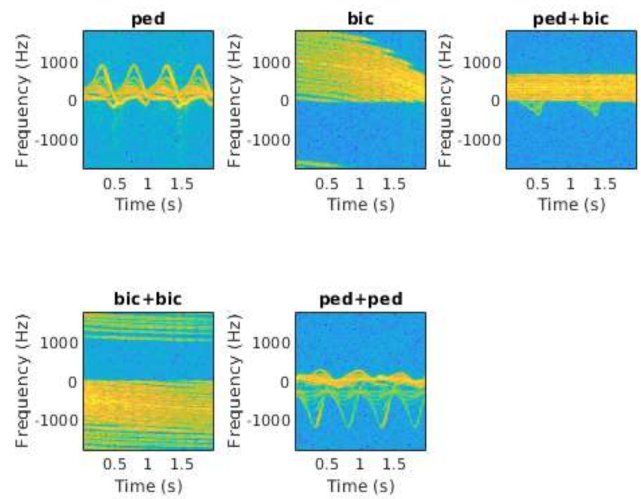

FIGURE 5. Spectrogram of Five Categories

The three items have quite different signatures, as shown by the spectrograms, normalized (STFT absolute values). Micro-Doppler signals from the sway of arms and legs and the spinning of wheels can be seen in detail in the spectrograms of both the pedestrian and the bicycle.

The anticipated labels for test and train data analysis are obtained by training a network to classify. The network reaches about 95.6% accuracy (see Fig. 6).

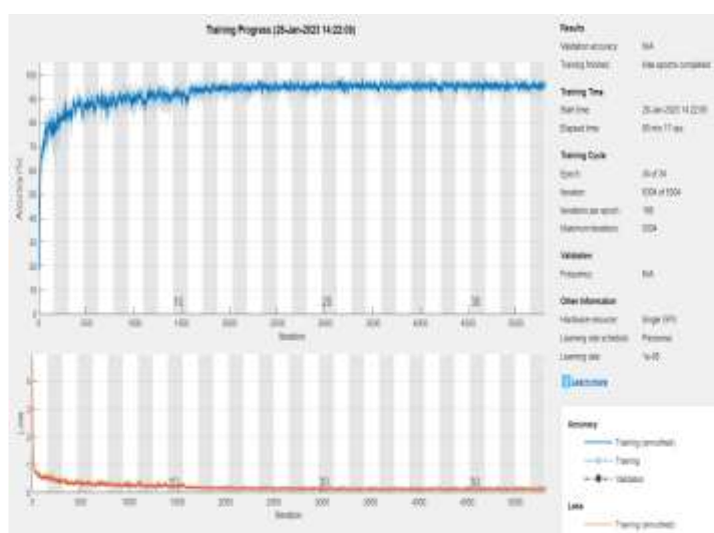

FIGURE 6. Training phase

For each group, the confusion matrix provides specific data on the predictions' accuracy. The trained network's confusion matrix accurately predicts the labels of the signals in the test data set (see Fig. 7).

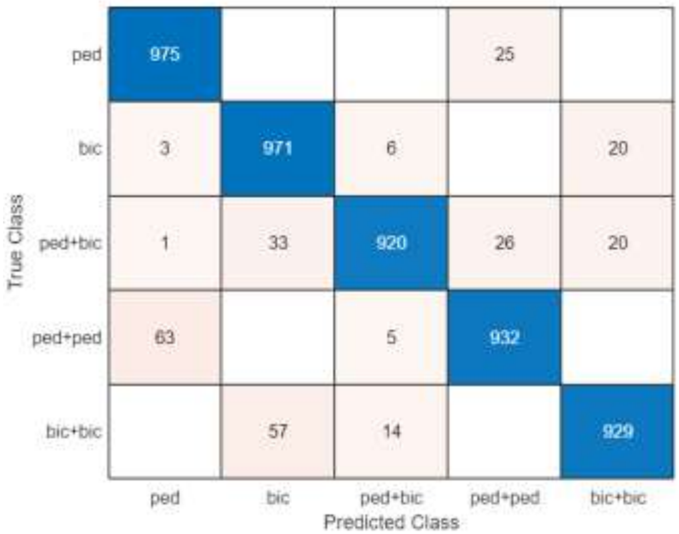

FIGURE 7. Confusion Chart

V. CONCLUSION

Classification issues in automobile radar have been addressed, and the results of these experiments utilize several neural network topologies. Despite being supported by a small amount of experimental data, these proof-of-concept outcomes demonstrated advantages. The system has a 95.6% success rate in its classifications. Furthermore, the significance of selecting appropriate characteristics of radar systems and radar signal processing the best potential input data to the networks (for use in target detection and tracking), as well as the potential drawbacks (overgeneralization and overfitting) of the network, are discussed. Whenever the radar type is framed image identification task between several spectrograms, residual networks tend to deliver superior results compared to simpler convolutional networks. It has also been suggested to combine recurrent and convolutional networks. However, this research describes a scenario where the network may be overfitted in reduced conditions if the first convolutional layer is too deep to distinguish slight changes in target spectrograms across different categories.

Moreover, instead of restricting similar to a challenge of image classification, one may think about simply recurrent network design and skipping the convoluted part, so recasting the considering the radar classification issue as a data sequence classification (series of radar pulses). This would be a novel way to investigate the data, as it would permit the investigation of data from radar domains apart from MD patterns, including range profile sequences or unprocessed complex data.

VI. FUTURE WORK

Regardless, future efforts should emphasize building a more extensive training dataset based on experimental results and validating the selected neural networks; this is especially important for very deep networks with a high number of tuning parameters. The problem of gathering and correctly labelling it is well known; deep Learning can improve from radar data, and approaches like transfer learning and pretraining are being explored.

The clutter cancellation filter, which for simplicity has been assumed to be stationary in this study, could benefit from further research into making it adaptive, considering how fast and in which way the radar-carrying vehicle is moving. Another promising topic, data fusion from different sensors, is a promising area for studies aiming to boost classification performance and inter-classifier Learning (such as cameras, Lidar, or other sensors).

Lastly, research on the different network designs should focus on how their cognitive abilities and performance are formed and how well they can be predicted while ensuring they follow the rules for standardization and safety in the auto industry.

FUNDING STATEMENT

The authors declare they have no conflicts of interest to report regarding the present study.

CONFLICTS OF INTEREST

The authors state they have no conflicting interests related to this study.

ACKNOWLEDGMENT

This work received support from the Artificial Intelligence in Health Care Lab National Center of Artificial Intelligence. Gratitude is extended to our colleagues at work.

REFERENCES

[1] Parekh, D., et al., *A review on autonomous vehicles: Progress, methods and challenges.* Electronics, vol. **11, no.** 14, pp. 2162, 2022.

[2] Organization, WH, *Global status report on road safety 2018: Summary (No. WHO/NMH/NVI/18.20).* World Health Organization, 2018.

[3] Tian, D., Han, Y., Wang, B., Guan, T., Wei, W. A Review of Intelligent Driving Pedestrian Detection Based on Deep Learning. Computational Intelligence and Neuroscience, 2021. https://doi.org/10.1155/2021/5410049.

[4] Brunetti, A., et al., *Computer vision and deep learning techniques for pedestrian detection and tracking: A survey.* Neurocomputing, vol. **300,** pp. 17-33, 2018.

[5] Sahin, Furkan E. "Long-range, high-resolution camera optical design for assisted and autonomous driving." In photonics, vol. 6, no. 2, p. 73. 2019.

[6] He, K., et al. *Deep residual Learning for image recognition*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.

[7] Jokanović, B. and M. Amin, *Fall detection using deep Learning in range-Doppler radars.* IEEE Transactions on Aerospace and Electronic Systems, vol. **54, no.** 1, pp. 180-189, 2017.

[8] Trommel, R., et al. *Multi-target human gait classification using deep convolutional neural networks on micro-Doppler spectrograms*. in *2016 European Radar Conference (EuRAD)*. 2016. IEEE.

[9] Rithik, D.S. and J. NarsimhaRao, *VEHICLE AND PEDESTRIAN DETECTION USING HAAR CASCADE.*

[10] Xie, J., et al. *Count-and similarity-aware R-CNN for pedestrian detection*. in *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XVII 16*. 2020. Springer.

[11] Hsu, W.-Y. and W.-Y. Lin, *Ratio-and-scale-aware YOLO for pedestrian detection.* IEEE transactions on image processing, vol. **30, p**p. 934-947, 2020.
[12] Raza, M.A., et al. *BirdView Retina-Net: Small-Scale Object Detector for Unmanned Aerial Vehicles*. in *2021 16th International Conference on Emerging Technologies (ICET)*. 2021. IEEE.
[13] Khalifa, A.B., et al., *pedestrian detection using a moving camera: A novel framework for foreground detection.* Cognitive Systems Research, vol. **60, p**p. 77-96, 2020.
[14] Rajendar, S., et al., *Prediction of stopping distance for autonomous emergency braking using stereo camera pedestrian detection.* Materials Today: Proceedings, vol. **51, p**p. 1224-1228, 2022.
[15] Reddy, T.S., *Video Surveillance System for Pedestrian Crossing using Matlab.* International Journal of Future Generation Communication and Networking, vol. **13, no.** 4, pp. 1810-1816, 2020.
[16] Mihai, S., et al. *Towards autonomous driving: A machine learning-based pedestrian detection system using 16-layer LiDAR.* in *2020 13th International Conference on Communications (COMM)*. 2020. IEEE.
[17] Enayati, J., P. Asef, and Y. Jonnalagadda, *A Novel Triple Radar Arrangement for Level 2 ADAS Detection System in Autonomous Vehicles.* 2022.
[18] P. Doll, RG, and F. Ai, , *Mask R-CNN ar*.
[19] Siddiqi, M.H. and I. Alrashdi, *Sharpening and Detection of Road Studs for Intelligent Vehicles at Nighttime.* Security and Communication Networks, 2022. **2022**.