Implementation of Real-Time Object Classification And Detection Model for Pedestrian Safety

Arnab Hati
MSc Data Analytics
(School of Computing)
National College of Ireland
Dublin, Ireland
x22107321@student.ncirl.ie

Rajshri Pawar
MSc Data Analytics
(School of Computing)
National College of Ireland
Dublin, Ireland
x22126571@student.ncirl.ie

Taniya Bagh
MSc Data Analytics
(School of Computing)
National College of Ireland
Dublin, Ireland
x22120831@student.ncirl.ie

***Abstract*—In machine vision applications, the detection of objects plays an important role in enhancing safety in various real-world scenarios. This paper basically focuses on pedestrians in both images and videos. As it plays an important role in different domains like autonomous vehicles, surveillance systems, and human-computer interaction. Based on machine learning and deep learning, object detection is a method to identify various objects in images and videos. The goal of object identification using computer intelligence is a technique to get an accurate performance based on image or video observation. In this study, we are trying to propose a system that can detect and classify the different types of objects, basically pedestrians and their surroundings. The model will be implemented using deep learning algorithms such as Convolutional neural networks, Region-based CNN, and YOLO with the help of images and videos. The algorithms are based on real-life scenarios, accuracy, and efficiency. The dataset is playing an important role in this study as it includes all the important attributes required. This study will include all the challenges with object detection on pedestrians like different lighting conditions, occlusions, varying scales, and complex backgrounds. The end results of this experiment will be focusing on the positive and negative of each algorithm. Moreover, understanding of the real-time application of the algorithms which are provided, considering their complexity and resource requirements.**

Keywords—*YOLO, Convolutional Neural Network, Region-based CNN, Computer intelligence, Complex backgrounds.*

I. INTRODUCTION

Many different areas have seen major changes because of the development of computer vision classification and detection, which has made it possible for bots to classify, detect and analyze visual data from their surroundings.

The recognition and translation of various factors inside an image or video are made easier by object classification and detection, one of the important roles in this study.

The importance of object detection extends beyond a variety of applications, including robots, augmented reality, surveillance, and autonomous vehicles.

With the wide-ranging fields of transportation systems, human-computer interaction, and public safety, pedestrian object classification and detection have a particular interest in the field of object detection.

To ensure the safety of pedestrians and to improve the functionality of automated systems that interact with or react to human presence, effective pedestrian identification in a variety of settings is essential.

This paper explores the complexities of object classification and detection with a particular focus on pedestrian detection in pictures and videos. Due to the dynamic nature of human movement, a wide variety of looks, occlusions, and complicated backgrounds, pedestrian classification, and detection present special obstacles. These difficulties call for the investigation of new approaches that can successfully handle these complexities and deliver consistent outcomes in a variety of real-world contexts.

Major improvements have been achieved recently in the field of object detection, with deep learning-based methods changing the field. The foundation of modern object classification and detection approaches, convolutional neural networks (CNNs), and similar architectures have proven to offer unmatched ability in capturing complicated visual patterns. Large-scale datasets with annotations and computer resources have also made it easier to train complicated models, which has allowed them to achieve remarkable levels of accuracy.

The goal of this study is to thoroughly explore and analyze advanced object classification and detection algorithms specifically designed for pedestrian safety. This study tries to clarify the advantages, drawbacks, and practical considerations of each technique by measuring their performance across a wide range of scenarios and datasets. In addition, the study shifts its attention to video-based pedestrian identification in recognition of the potential for time aspect to improve classification and detection accuracy and dependability.

The purpose of this analysis is to offer helpful insights for researchers, professionals, and developers who are working to construct reliable pedestrian detection systems. A solid understanding of the current approaches and their applicability is crucial as the integration of such systems into applications that require safety becomes more common. In the end, our research advances pedestrian classification and detection technologies, promoting safer surroundings and more complex human-machine interactions. The work flow diagram has been mentioned in figure 1.

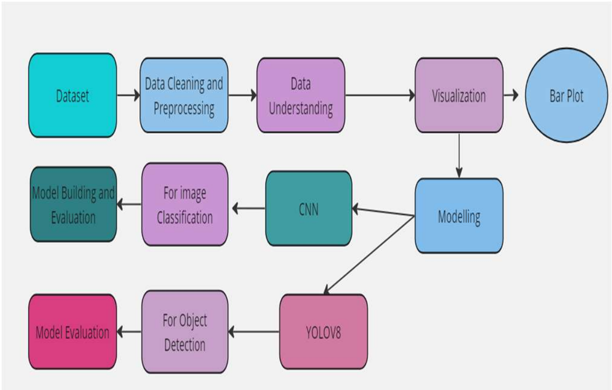

Fig 1 Work Flow Chart

II. LITERATURE REVIEW

The drawback of the currently utilized systems for object detection is described in [1] where it discusses the limitations of DNN-based systems in terms of complexity in data and accuracy achieved. Since real-time object detection has feature extraction as a key feature for achieving accurate and precise results, it proposes new techniques based on deep learning algorithms such as R-CNN and YOLO since it helps the system to learn automatically about the object which needs to be detected among many images.

The implementation of face mask detection using deep CNN models named YOLOV5s and YOLOV5I which are further the latest models, is described in [2]. The systems used for the deployment of these models are NVIDIA Jetson Nano and another platform was Jetson Xavier NX. Furthermore, the performance of these models was evaluated through the mAP value. On comparing the values, YOLOV5I gave a higher value of mAP (92.49) than YOLOV5s (86.43).

The work of a UAV (Unmanned aerial vehicle) named quadrotor which is further classified as VTOL that is Vertical Takeoff Landing for object detection has been explained in the paper [3]. The limitations of this UAV such as non-linearities, under-actuated, static stability, and many more make it more complex and increase the need for the proper controller. For controlling the drone (quadrotor) PIX-HAWK which is a flight controller was used along with GPS to achieve moderate stabilization while in the air, for polit mode PX4 software was used. The second part consists of the object detection model, YOLOV5 was used for the same. The architecture was divided into three parts to achieve the final model namely CSPDarknet which is used for feature extraction, PANet which is used for feature fusion, and YOLO layer which provides the expected results (object detection).

In the paper [4], the author describes the implementation of deep learning with mixed reality devices such as Robots to increase their capability of human interactions. To achieve this, they have applied the YOLO model on the client side with HoloLens (AR technology) and Ubuntu for real-time object detection and based on the results it was proven to be a high-speed object detection system. The YOLO models used for this purpose were YOLOV3 and YOLOV4. Due to the drawback faced due to the hardware GPU, the detection speed of YOLOV4 was only 13 fps but the accuracy achieved was much higher than that of YOLOV3.

The implementation and advantage of the YOLOV3 deep learning model have been described in [5] which was further used for the detection of dynamic targets in real-time. Since it needs large image data to process and provide real-time results, therefore instead of traditional methods deep learning method is used for the implementation of such models. This model consists of convolutional layers that are fully connected to each other. Based on the results, the accuracy achieved is very high which helped to achieve a detection model. The YOLOV3 model also helped in optimizing the limitations faced by the detection of dynamic targets.

The optimization of YOLOV3 for achieving a detection system for vehicle targets in an autonomous vehicle environment with higher accuracy has been described in [6]. The researchers address the limitations caused by the multiple parameters present in the mentioned YOLO models and optimize the model using the deep separable CNN method. In this method, for CNN the reduced parameter is used to enhance its ability to learn. The model uses models such as MobileNetV2 and YOLO to create a model named the MobileNet Yolo model for vehicle target detection. Moreover, the function associated with the loss while implementing this model is used for the implementation of another model which will be further used for addressing the regression problems. The newly created model shows an improvement in accuracy which was raised by 78% to detect the targets associated with the vehicles and on the other hand, the model size decreased by 26%.

The advancement and application of computer vision in traffic safety, ADAS, and self-driving vehicles are explained in [7]. In this paper, researchers have discussed the application of YOLOV4 to detect potholes, people, and animals to avoid any mishappenings. The backbone network of the model for detecting objects was stage partial CNN. The hardware such as NVIDIA GTX 1060 GPU was used to increase the speed of the computation. The precision scores that were achieved for the above-mentioned objects ranged between 0.97 to 0.99.

The use of an advanced version of YOLOV3 named DenseYOLOV3 for the detection of traffic status systems has been described in [8]. The process includes the extraction of information associated with on-line traffic on roads. For detection purposes, they have used a combination of deep learning and a generic object detection system. Along with DenseYOLOV3 mIOU has been used to achieve the proposed system with a precision rate of 86% and recall rate of 87%.

The prevention of accidents caused due to jaywalking using deep learning models has been shown in [9]. There have been techniques associated with computer vision for path analysis, detecting objects in motion, forecasting the moving object, and many more but the drawback that it faces is it is specific to the area of a single road. In this paper, YOLOV4 along with DeepSORT which is a type of multi-object tracking system has been used to create a model which can detect jaywalkers on multiple roads simultaneously.

The implementation of YOLO over traditional signal controllers to control traffic signal lights has been described in [10]. The advantage of YOLO over CNN is also shown in the paper. The data is collected as images using NVIDIA Jetson which is further stored in a cloud object storage named AWS S3 bucket and then these images are processed as part of code associated with YOLO to provide the output.

The use of YOLO with Convolutional Neural Networks to identify the blur images has been described in [11]. The conventional method of blur image detection consists of edge detection, but it fails when there is no edge present in the images therefore the authors have developed a model where the YOLO model is used for object detection and the layer of CNN is trained with another dataset to identify the bigger objects.

The working of R-CNN has been explained in [12]. A brief survey was done on region-based CNN along with the extended technologies associated with RCNN such as Fast RCNN and another technology named Faster RCNN. Upon technical evaluation, it was seen that faster RCNN shows a lot better accuracy and faster detection speed.
The implementation of CNN for classifying the image of leaf has been done in [13]. The model was used to recognize the properties of the leaf in the aspect of its geometrical features and statistical values, but it lacked to detect the features related to topology, therefore, an integrated model that consists of VGG16 (a CNN model) and PI was implemented to address all the three aspects of the leaf and the accuracy of the model achieved was 98% approx.

The classification of the use of land based on images captured by remote satellites using a Convolutional Neural Network has been carried out in [14]. The study has been carried out in a manually made dataset containing images taken from Google Earth. The model contains three layers, 2 convolutional and a third with three hidden layers fully connected. The performance of the evaluated using the BCE method and accuracy was 93% approx. with test data.

The implementation of Faster R-CNN for advanced traffic safety has been explained in [15]. The authors have described the importance of object detection in terms of the driver's safety and the surroundings of the vehicle. For model implementation, the Faster R-CNN model was used since it is known for accurate results and faster detection of small to medium objects present in the surroundings. The accuracy of the model was 86% approx. Also, FPR and FNR are also calculated through the confusion matrix for the evaluation of the model and the values were 15.97% and 12.27% respectively.

The comparison of SSD300, Tiny-YOLO V3, and RetinaNet has been carried out in [16] based on the detection of images with diameters of 2–3-inch diameter which is 20*20 to 40*50 pixels at a distance of 3 to 4 m. The best-fitting model was YOLO V3 with an accuracy of 60%.

To detect lanes CNN model along with a computer is used to create a model and was described in [17]. The aim is to enhance the capabilities of lane detection. To collect data, sensors such as LiDar, front cameras, GPS etc are used. TensorFlow is used for modeling neural networks.

The implementation of deep learning algorithms for object detection, tracking the targets, classification of videos, and many more have been discussed in [18]. These algorithms are used for analyzing the flow of pedestrians. The model was implemented in three stages which included a pipeline for deep learning, visualization of data, and a database to monitor the flow. Previously Deep SORT algorithm and YOLOV5 were used for the model but due to shortcomings caused by the dynamic features of the pedestrians, it becomes extremely tedious to establish an effective pedestrians flow static system. To overcome these drawbacks Idbased model was created which included PP-YOLOE for target detection, for image classification model PP-LCN. For the video-based model, CSRNet and PP-TSM models are used, and for the skeleton-based model, HRNet and ST-GCN models are used for detecting actions.

The optimization of the algorithm based on ssd for detecting moving objects including pedestrians has been discussed in [19]. To improvise the target detection and classification of the image CBAM module is used in the layer used for the detection. The introduction of this feature fusion has enhanced the ability of the feature map and K++ clustering is used along with IOU for setting the points of the parameters. These implementations of the algorithms resulted in an increase in accuracy and helped in achieving a better and faster training model which further helped in achieving an efficient and more accurate SSD model.

In [20], a study is carried out to propose a 3D CNN used for the classification of HSI. Previously only 2D CNN models were used for feature extraction, but this paper proposed a deep-learning feature extraction with the data of HIS split into 3D samples and then fed as an input to the proposed CNN model for getting the expected results. The performance upon evaluation was found to be more accurate than the 2D models.

Pedestrian safety is one of the major concerns everywhere around the world and leads to the need for a system that will give accurate yet faster results for detection. To achieve such a model, researchers in [21] have proposed a model consisting of CNN with a single shot detection framework for activating maps of different places and OpenCV for achieving a reliable and efficient model that will be further used in the real world to enhance pedestrian safety. The final implementation of the model was done by using MobileNet and Single shot Detector which makes it more compatible with real-world applications.

In [22], the authors have introduced the limitations associated with the traditional method for detecting objects, mostly pedestrians since their behaviour is dynamic which makes it difficult to detect in a short span of time mostly when they are present at some distance. To address these issues optimization and advancement in the deep learning model that is YOLOV4 is implemented. The accuracy of the model is increased based on the feature fusion and feature extraction associated with the appearances of the people. The YOLOV4 model is combined with a quantization method to achieve high accuracy of the model with faster speed for identifying the features of the pedestrians. The accuracy of the model achieved was 77.25% and the speed associated with the detection was 34.33 fps.

III. METHODOLOGY

KDD stands for " Knowledge Discovery in Databases," and it's a 6-step process for data mining projects. We'll be following KDD for this project, and the six steps of KDD methodology are outlined in Figure 2.

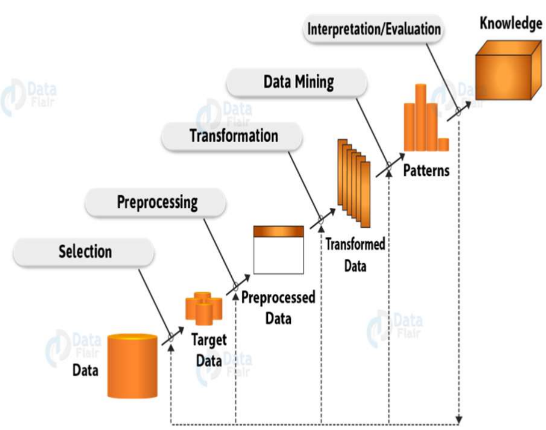

Fig.2 Steps involved in KDD [23]

1. *Goal setting and Business Understanding:*
   This paper examines the classification of objects through deep learning techniques and then analyzes the detection of an object in an image. The goal is to distinguish between different objects, and then to identify the different objects in the image through deep learning.

2. *Data selection and integration:*
   The dataset is available for public consumption on the Kaggle platform. It comprises images of eight distinct objects. As the project concerns human safety, it is classified into two distinct categories. The majority class of the dataset consists of 1800 instances, while the minority class of the dataset comprises 250 instances shown in Fig 3. Other objects such as buses, lorries, pedestrians, cars, and motorcycles are also included in the dataset.

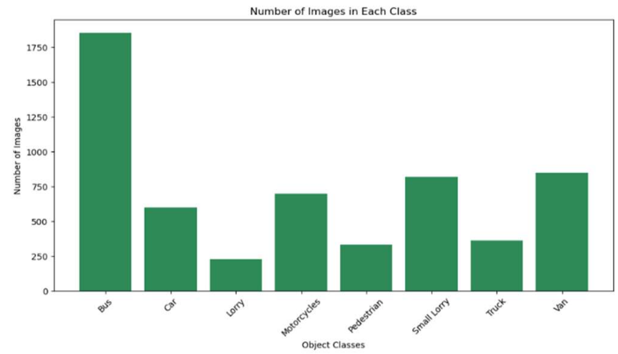

Fig 3 Imbalance in Dataset

3. *Data Cleaning and Pre-processing:*

   To address the issue of class imbalance, a data upsampling process was conducted. By randomly re-sampling from the initial dataset, additional data instances were generated. These images were aggregated and used to train and validate the model.

4. *Data Transformation:*
   The function of feature scaling is to normalize the values in an array between 0 to 1. This can be done by dividing the NumPy arrays by 255. The process of data augmentation is the creation of processed images from the original image by applying properties such as zoom, tilting, rotation, flipping, and zooming. In this research, we have added rotation, tilting, and zooming properties to the training dataset.

5. *Modeling:*
   The purpose of this paper is to evaluate the performance of a customized build CNN. The object detection process was performed with the help of YoloV8, and the object was successfully identified.

A. *CNN implementation:*

The images were stored in 8 distinct folders. The images were loaded from this folder. The images were inspected as outlined in Fig. 3 and the label was applied.

Fig 4. Different Images

The dataset was partitioned into training and testing data in a ratio of 80:20. Additionally, the dataset was found to be heavily skewed. The training dataset is partitioned into a ratio of 90:10. Ten percent of the data is utilized to hyper-perambulate the model, and then 90 percent of the data is utilized for training shown in fig 5.

```
In [12]: # Split dataset into train and test sets
         train_images, test_images, train_labels, test_labels = train_test_split(train_images, train_labels, test_size=0.2, ran

In [13]: # Split the training data into optimization data (10%) and remaining data (90%)
         opt_train_images, train_images, opt_train_labels, train_labels = train_test_split(train_images, train_labels, test_siz
```

Fig 5. Split the dataset.

The images were available in various sizes. Therefore, all the images are resized to 224 x 224 px. Normalize the images by dividing each image by 255 applying both the train and test dataset. In order to generate a CNN model suitable for classification, a CNN model was constructed, consisting of three convolution layers, dense layers, and an output layer. The "Relu" function was used as an activation function in the convolution layers and dense layers, while the "Softmax" function was employed as an activation function for the output layer. Compiling the model, the "Adam" optimizer was employed to calculate the loss, and a callback was used to prematurely stop the epoch, adjust the learning rate, and save the most suitable model. Fig 8 shows the output classified images. It can be observed that the model is slightly underfitting, which can be remedied by refining the preprocessing methods and refining the model further in future work. It can be observed that the model is slightly underfitting, which can be remedied by refining the preprocessing methods and refining the model further in future work.

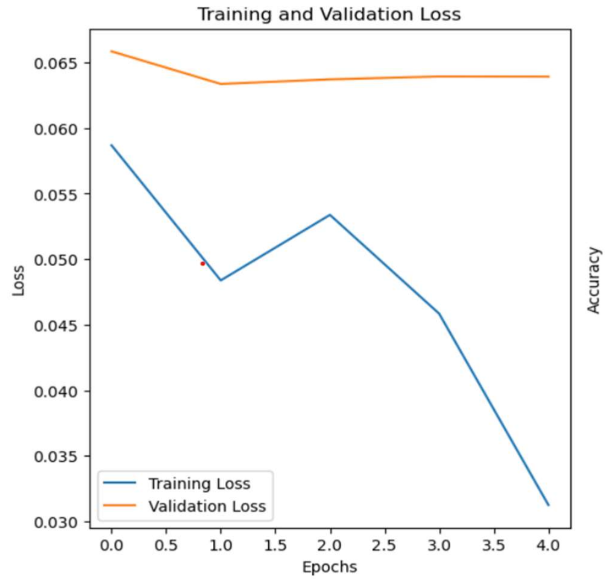

Fig 6

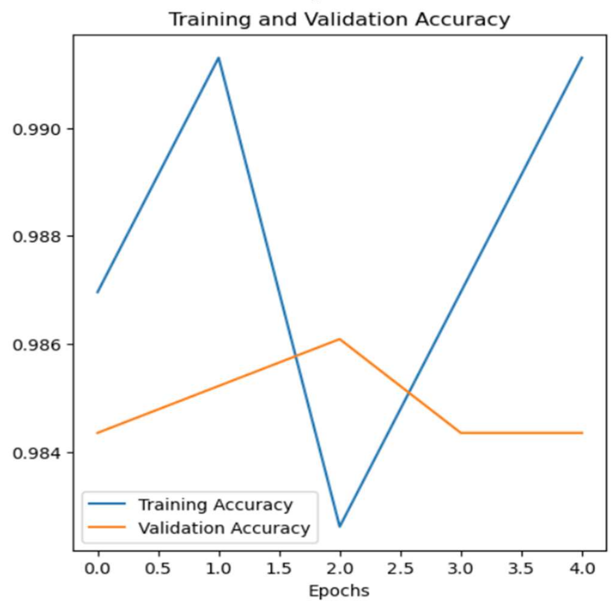

Fig 7.

Fig 8. The classified Image

B. Implementation YOLOV8

The YOLOV8 model is a pre-trained model. This model is an advanced version of the YOLO model which is further developed using Pytorch framework. The internal architecture of the model consists of a group of CNNs with convolutional layers of up to 53 in number and uses cross-stage connections that are partial in nature to enhance the information flow between the layers. For the implementation of this model, ultralytics package consisting of YOLOV8 is installed which further helps in the creation of a model consisting of neural networks. The YOLOV8 model contains an extension as .pt (yolov8.pt) since the model is trained and created using a framework named Pytorch. This model is classified into three different models depending on the application and then subdivided into five different types depending on the size of the model. The model can be used for Classification, detection, and segmentation. In our project, we are going to work with the detection type YOLOV8 model since we are going to detect objects to enhance the safety of pedestrians.

The model used here is yolov8m.pt, a middle-sized object detection model. The steps involved in the detection are- training the dataset, dataset prediction, and then lastly, export which is the result. Since this model is already trained in advance therefore directly the step involving data prediction is carried out in which the data i.e image is fed as input to the model and as an output it provides an array that is further used to give results that are non-other than the predicted object with bounded boxes around them as shown in the fig 9.

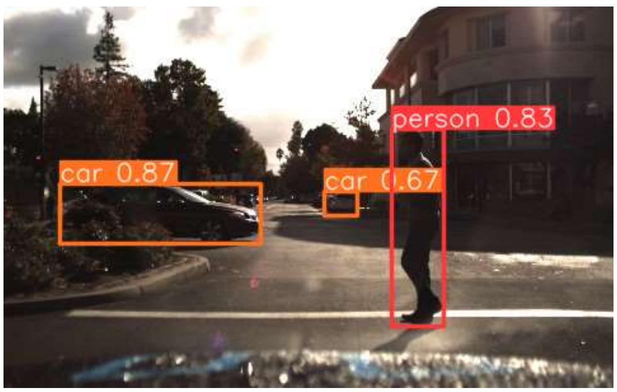

Fig 9. Implementation of YOLOV8

IV. RESULT AND EVALUATION

The evaluation of the model was done on the basis of their performances which was further evaluated by calculating the statistical values such as Accuracy, Precision, Recall, and F1 score.

Accuracy – The calculated value shows the performance of the model in the entire class present that is the proportion of the total predictions that are correct to the predictions present in total.

$$Accuracy = \frac{(TP + TN)}{(TP+FP+TN+F\ \ )}$$

Precision- The calculation of precision is carried out to determine the reliability of the model, the higher the value the more stable the model.

$$Precision = \frac{TP}{(TP+FP)}$$

Recall- This attribute determines the positive cases present in a dataset.

$$Recall = \frac{TP}{(TP+F\ \ )}$$

F1 Score- This is calculated using the values obtained from precision and recall determining the cumulative positive cases that are correct in nature is achieved.

$$F1 = 2 \times \frac{(Precision \times Recall)}{(Precision + Recall)}$$

Where TN is True Negative, TP is True Positive FN is False Negative, FP is False Positive.

A. For CNN

The confusion matrix with the values of accuracy, precision, recall, and F1 Score for the CNN model is mentioned in Figure 10- 13.

Fig.10 Confusion matrix of CNN model

```
Test Accuracy: 0.9321739077568054
```

```
Train Accuracy: 0.9538
```

Fig.11 Test and Train accuracy of CNN model

```
Precision: 0.9342884547216802
Recall: 0.9321739130434783
```

Fig.12 Precision and Recall Values of CNN model

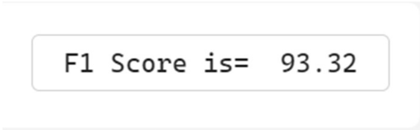

Fig.13 Precision and Recall Values of CNN model

Based on the values of the accuracy of the model, precision, recall, and F1 score it can be seen that the model implemented is working as anticipated.

V. CONCLUSION

We have investigated object detection in this research, with a focus on pedestrian detection in photos and videos. With the help of YOLOv8 for object recognition and Convolutional Neural Networks (CNNs) for classification, we have taken on the critical task of recognizing pedestrians in a variety of real-world circumstances.

It has been shown that using CNNs for classification tasks has changed the field of computer vision. Accurate and effective classification of pedestrians has been made possible by CNNs' capacity to learn complicated characteristics and patterns from images, setting the stage for further item recognition techniques. CNNs effectively distinguish pedestrians from other objects, providing a strong foundation for more complex detection algorithms.

Our ability to recognize objects has been improved even more with the use of YOLOv8, an advanced object detection framework. The architecture of YOLOv8 combines exactness and speed, making it suitable for real-time applications like pedestrian detection. Its value in improving surveillance and public safety systems is highlighted by its capacity to detect pedestrians with high precision—even in difficult situations with occlusions and complicated backgrounds.

The effectiveness of the combined technique was proved by our experimental evaluation, which showed remarkable accuracy rates in both classification and object detection tasks. We were able to recognize pedestrians within photos using CNNs integrated for classification, and we were able to locate pedestrians within video frames effectively using YOLOv8's skill at object recognition.

However, we agree that there are still issues with pedestrian detection. Obstacles that require continual research and invention include variations in pedestrian appearances, lighting, and occlusions. Although our method produced commendable results, additional improvements in precision and effectiveness might be achieved by modifying the models and investigating hybrid architectures.

In conclusion, a reliable and successful method for detecting pedestrians has been proven by the addition of CNNs for classification and YOLOv8 for object identification. By providing insights into the implementation of cutting-edge approaches for improving pedestrian safety, surveillance systems, and other safety-critical applications, this work adds to the database of knowledge in computer vision. The given strategy serves as a starting point for future research projects aimed at ever more advanced and trustworthy pedestrian detection systems as technology advances and new problems are encountered.

VI. FUTURE WORK

Although the integration of CNNs for classification and YOLOv8 for pedestrian object detection has been successfully investigated in this study, there are still a number of opportunities for additional research and advancement like Multi-Modal Fusion, Real-Time Optimization, Domain Adaptation, Privacy Preservation, Edge Computing, Incremental Learning etc. By focusing on these areas, researchers can develop pedestrian detection technology, eliminating present drawbacks and building the foundation for future systems that are more reliable, flexible, and privacy-conscious.

REFERENCES

[1] P. K. Goswami and G. Goswami, "A Comprehensive Review on Real Time Object Detection using Deep Learing Model," 2022 11th International Conference on System Modeling & Advancement in Research Trends (SMART), Moradabad, India, 2022, pp. 1499-1502, doi: 10.1109/SMART55829.2022.10046972.

[2] S. Ruparelia, M. Jethva and R. Gajjar, "Real-time Face Mask Detection System on Edge using Deep Learning and Hardware Accelerators," 2021 2nd International Conference on Communication, Computing and Industry 4.0 (C2I4), Bangalore, India, 2021, pp. 1-6, doi: 10.1109/C2I454156.2021.9689421.

[3] M. Issame and A. Benyounes, "Real Time Object Detection With Drone Using Deep Learning Algorithm," 2022 International Conference of Advanced Technology in Electronic and Electrical Engineering (ICATEEE), M'sila, Algeria, 2022, pp. 1-6, doi: 10.1109/ICATEEE57445.2022.10093104.

[4] J. Guo, P. Chen, Y. Jiang, H. Yokoi and S. Togo, "Real-time Object Detection with Deep Learning for Robot Vision on Mixed Reality Device," 2021 IEEE 3rd Global Conference on Life Sciences and Technologies (LifeTech), Nara, Japan, 2021, pp. 82-83, doi: 10.1109/LifeTech52111.2021.9391811.

[5] Y. Yu, "Research on Real-time Dynamic Object Detection Based on YOLOv3 Deep Learning Network," 2023 IEEE 3rd International Conference on Electronic Technology, Communication and Information (ICETCI), Changchun, China, 2023, pp. 1624-1629, doi: 10.1109/ICETCI57876.2023.10176887.

[6] N. Wang and Y. Jia, "Research on Vehicle Object Detection Based on Deep Learning," 2023 4th International Conference on Computer Vision, Image and Deep Learning (CVIDL), Zhuhai, China, 2023, pp. 412-415, doi: 10.1109/CVIDL58838.2023.10166204.

[7] N. Sharma and R. D. Garg, "Real-Time Computer Vision for Transportation Safety using Deep Learning and IoT," 2022 International Conference on Engineering and Emerging Technologies (ICEET), Kuala Lumpur, Malaysia, 2022, pp. 1-5, doi: 10.1109/ICEET56468.2022.10007226.

[8] C. -T. Lam, B. Ng and C. -W. Chan, "Real-Time Traffic Status Detection from on-Line Images Using Generic Object Detection System with Deep Learning," 2019 IEEE 19th International Conference on Communication Technology (ICCT), Xi'an, China, 2019, pp. 1506-1510, doi: 10.1109/ICCT46805.2019.8947064.

[9] S. Mostafi, W. Zhao, S. Sukreep, K. Elgazzar and A. Azim, "Real-Time Jaywalking Detection and Notification System using Deep Learning and Multi-Object Tracking," GLOBECOM 2022 - 2022 IEEE Global Communications Conference, Rio de Janeiro, Brazil, 2022, pp. 1164-1168, doi: 10.1109/GLOBECOM48099.2022.10000957.

[10] S. K. Saini and M. Singh Ghumman, "Automated Traffic Management System Using Deep Learning Based Object Detection," 2022 International Conference on Machine Learning and Cybernetics (ICMLC), Japan, 2022, pp. 1-5, doi: 10.1109/ICMLC56445.2022.9941332.

[11] A. Senapati et al., "Identification of blurred objects in real time Video using deep learning neural networks," 2022 13th International Conference on Computing Communication and Networking Technologies (ICCCNT), Kharagpur, India, 2022, pp. 1-4, doi: 10.1109/ICCCNT54827.2022.9984429.

[12] O. Hmidani and E. M. Ismaili Alaoui, "A comprehensive survey of the R-CNN family for object detection," 2022 5th International Conference on Advanced Communication Technologies and Networking (CommNet), Marrakech, Morocco, 2022, pp. 1-6, doi: 10.1109/CommNet56067.2022.9993862.

[13] Y. Zhang, "Improved Leaf Image Classification Using Topological Features and CNN With Attention Module," 2021 IEEE 6th International Conference on Signal and Image Processing (ICSIP), Nanjing, China, 2021, pp. 311-315, doi: 10.1109/ICSIP52628.2021.9689038.

[14] P. S, A. P. A, R. R, R. K. P, R. K and P. C. J, "CNN Based Automated Land Use Classification from Remotely Sensed Image," 2021 Innovations in Power and Advanced Computing Technologies (i-PACT), Kuala Lumpur, Malaysia, 2021, pp. 1-4, doi: 10.1109/i-PACT52855.2021.9696772.

[15] A. Datta, T. Islam Meghla, T. Khatun, M. Hasan Bhuiya, S. Rahman Shuvo and M. Mahfujur Rahman, "Road Object Detection in Bangladesh using Faster R-CNN: A Deep Learning Approach," 2020 IEEE International Women in Engineering (WIE) Conference on Electrical and Computer Engineering (WIECON-ECE), Bhubaneswar, India, 2020, pp. 348-351, doi: 10.1109/WIECON-ECE52138.2020.9397954.

[16] R. Gopal, S. Kuinthodu, M. Balamurugan and M. Atique, "Tiny Object Detection:Comparative Study using Single Stage CNN Object Detectors," 2019 IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE), Bangalore, India, 2019, pp. 1-3, doi: 10.1109/WIECON-ECE48653.2019.9019951.

[17] M. Saranya, N. Archana, J. Reshma, S. Sangeetha and M. Varalakshmi, "OBJECT DETECTION AND LANE CHANGING FOR SELF DRIVING CAR USING CNN," 2022 International Conference on Communication, Computing and Internet of Things (IC3IoT), Chennai, India, 2022, pp. 1-7, doi: 10.1109/IC3IOT53935.2022.9767882.

[18] Y. Bu, "Pedestrian flow monitoring system based on Deep learning pipeline," 2023 4th International Conference on Computer Vision, Image and Deep Learning (CVIDL), Zhuhai, China, 2023, pp. 395-398, doi: 10.1109/CVIDL58838.2023.10167254.

[19] E. Sun, X. Ma and M. Li, "Improved SSD based pedestrian detection algorithm for forklift active warning system," 2022 IEEE 5th Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC), Chongqing, China, 2022, pp. 1523-1528, doi: 10.1109/IMCEC55388.2022.10019947.

[20] M. Kanthi, T. H. Sarma and C. S. Bindu, "A 3d-Deep CNN Based Feature Extraction and Hyperspectral Image Classification," 2020 IEEE India Geoscience and Remote Sensing Symposium (InGARSS), Ahmedabad, India, 2020, pp. 229-232, doi: 10.1109/InGARSS48198.2020.9358920.

[21] Z. Ahmed, R. Iniyavan and M. M. P., "Enhanced Vulnerable Pedestrian Detection using Deep Learning," 2019 International Conference on Communication and Signal Processing (ICCSP), Chennai, India, 2019, pp. 0971-0974, doi: 10.1109/ICCSP.2019.8697978.

[22] B. Zhao, X. Chen, X. Zhang, H. Li, C. Chen and S. Yang, "A Receptive Fields Enhancement YOLO V4 on Pedestrian Detection," 2022 4th International Conference on Frontiers Technology of Information and Computer (ICFTIC), Qingdao, China, 2022, pp. 1066-1072, doi: 10.1109/ICFTIC57696.2022.10075128.

[23] https://data-flair.training/blogs/data-mining-and-knowledge-discovery/