SSRG International Journal of Electronics and Communication Engineering
ISSN: 2348-8549/ https://doi.org/10.14445/23488549/IJECE-V12I8P108

Volume 12 Issue 8, 91-101, August 2025

Original Article

FGRCNN: A Newly Developed Hybrid Network for Multiple Object Detection Using Traffic Images

Dhanalakshmi Manickaraj[1], Palvinder Singh Mann[2]

[1]*Computer Science & IT, Gujarat Technological University, Gujarat, India.*
[2]*Department of Computer Science & IT, Central University of Jammu, Jammu, Jammu & Kashmir, India.*

[2]*Corresponding Author : palvinder.csit@cujammu.ac.in*

Received: 04 June 2025 Revised: 06 July 2025 Accepted: 07 August 2025 Published: 30 August 2025

Abstract *- Detection of more than one object from traffic images in real time using a Social GAN-based Fast Regional Convolutional Neural Network is presented in this paper. To detect the objects, the images obtained from the Multi-Object Tracking (MOT) dataset had undergone pre-processing by employing a Gaussian filtering technique, and enhanced images were subjected to object detection by merging two networks, Social GAN and Fast Region-based CNN (FRCNN). To assess the efficiency of the approach, experimental tests were applied on diverse image datasets, and the results were validated. The Social-Gan-based FRCNN provided better accuracy of 86.5%, TPR of 94.7%, TNR of 90.9%, PPV of 88.8% and NPV of 55.8%..*

Keywords *- Multiple object detection, Traffic images, Fast regional based social GAN.*

1. Introduction

The present scenarios stipulate automation to control traffic with certain operations and the ability to do the same in intricate weather conditions and at night [1]. Automatic detection of several objects, like pedestrians and vehicles, through traffic scenarios depicts a fundamental principle for several intelligent transportation models. The rational management of traffic and control, assisted vehicle movement, and pedestrian movement can lessen the rate of accidents and congestion on roads. Hence, substantial efforts have been devised over prior decades [2]. Automatic driving is termed as an imperative domain of Artificial Intelligence (AI) applications. AI-based techniques are extensively utilized. Detection of objects and ranging are imperative processes in the domain of automatic driving. The precise discovery and accuracy of objects in front of the driving vehicle are imperative requirements and fundamental assurance for feasible autonomous driving. Object detection is a technique utilized for determining accurate object locations. It can be used in various commercial scenarios, including live video monitoring, studying consumer actions, face recognition, traffic flow analysis, and autonomous driving technologies [3]. Accurate and efficient object detection is a key focus of object recognition, but achieving this is challenging. Detection of multiple objects in real-time is challenging because of factors like changes in pose, occlusions, varying lighting, and differences in object sizes [4]. Many works on detecting objects range from automatic driving and have provided huge accomplishments. However, the majority of research stayed in the theoretical phase of the technique, which falls to unify multi-object detection. Tracking and detecting multiple objects from videos is an imperative process and is utilized in many application domains that involve detecting events, automatic driving and navigation of robots [5]. Detection of objects and tracking are performed with machine learning models. The videos are attained in real-time, comprise different objects, and are produced online using sensory models (cameras). Image processing models can enhance the segmentation of images in videos; however, their effectiveness is generally confined to straightforward traffic flow scenarios. Meanwhile, IP schemes are not able to detect occluded objects. These problems are solved with Multi-View Stereo (MVS) because of the huge enhancement of deep models and benchmark databases. By integrating Machine Learning (ML) techniques with technically advanced Artificial Intelligence (AI), MVS is capable of identifying and monitoring more than one objects in real-time visual data [6]. Deep Learning (DL), a subset of AI, is employed in advanced technologies to enhance accuracy and achieve superior performance. In ML and computer vision, DL models such as CNN have accomplished improved effectiveness [7]. CNN is a technique of DL which particularly deals with image detection as it poses the ability to minimize the size of different inputs by convolution. Hence, CNN-assisted techniques are devised for autonomous driving like classification and detection of vehicles and pedestrians [8-10] and environmental awareness [11].

Presently, the majority of CNN-assisted networks provide elevated accuracy and real-time efficacy in detecting objects. DL-based object identification schemes use deep CNNs for mining features through images and videos and categorize objects with these features. Such detectors fall into two categories: single-stage [12] and double-stage detectors [13]. For double-stage models, the regions where the object's presence is anticipated are initially predicted. Next, features are extracted from these specified regions and used for classification, while bounding-box regression refines the detection of likely object candidates. In single-stage models, bounding boxes are generated directly from the input images without requiring predefined regions. While double-stage detectors offer higher accuracy in object detection and classification, single-stage detectors achieve results faster [14]. The development of Fast RCNN and Faster RCNN has greatly advanced two-stage object detection methods [15, 16]. Despite the efficiency of CNN when adapting to object detection for traffic scenes, one of the major conundrums is that the classical CNN-based techniques are scale sensitive. At the same time, it is common to scale different objects using traffic images [31]. Object detection in traffic environments has seen rapid advancement through deep learning techniques in recent years. However, most existing approaches primarily focus on spatial features of objects, often overlooking the dynamic interactions and contextual relationships among them. This gap becomes more evident in crowded or complex traffic scenes, where multiple objects such as vehicles and pedestrians interact in unpredictable ways. Traditional models like Faster R-CNN and YOLO, though powerful, fall short in effectively capturing these social interactions, leading to reduced accuracy in such scenarios. Therefore, there is a critical need for a detection framework that combines precise object localization with an understanding of object behaviour and interaction. The gap has been addressed in this study by proposing a hybrid model that integrates Social-GAN with FRCNN to enhance both contextual awareness and detection accuracy in real-time traffic monitoring applications.

1.1. Literature Survey

Over the years, various techniques have been developed for this purpose. This review highlights some of these methods, discussing their approaches, merits, and limitations. Gunasekaran, P., et al. [17] developed a new vehicle detection scheme that adapted YOLOv2 for performing multiple object detection. It helped to address the complexities of existing vehicle detection, such as small detection accuracy and slow speed. However, this method was not able to detect the objects on an unstable platform. Mohamed, I.S. et.al. [18] devised a Portable Appearance Extension (PAE) for detecting objects and mined appearance embeddings with a shared model. However, this method heavily impacts dropping frames because of inference. Pramanik, A., et al. [19] developed two new techniques, named Multi-Class Deep SORT (MCD-SORT) and granulated RCNN (G-RCNN). Object detection poses two phases: one is the localization of objects, and the second is classification. The method provided less computational speed.

Yang, J., et al. [20] utilized two networks in order to detect objects. A multi-GPU-based real-time synchronization technique was used to enable parallel training and object detection. The method also utilized center-selective ranging module for determining distant objects. However, this method acquired poor accuracy. Lian J. et al. [21] introduced one method for small object detection in traffic environments using a fusion of attention-based features. Initially, a multi-scale type channel attention block was designed to capture crucial feature map information at both local and global levels. This method needed network training, which took a lot of time. Murthy, J.S., et al. [22] devise a new scheme known as YOLOv5 for detecting objects present in the traffic images. Here, the method showed enhanced speed in detecting objects in real-time. However, this method did not unify the Electronic Control Unit (ECU) contained in the vehicles.

Li et al. [23] proposed an innovative approach called TrackNet, which builds upon the Faster R-CNN architecture to identify a 3D bounding tube that encloses a moving object throughout a video sequence. In trackNet, the Tube Proposal Network (TPN) was incorporated for object prediction; however, it struggled with limited capability in achieving precise localization. He, X., et al. [24] devised an enhanced technique, YOLO-MXANet, for detecting video objects. Here, the loss function was utilized to promote the position accuracy of tiny objects. The method used Mosaic and Mixup to enhance the robustness of training. The Multi-scale Feature Enhancement Fusion (MFEF) was used to fuse the mined features, and activation was used to optimise the Convolution-Batchnorm-Leaky ReLU (CBL) module. However, this method was not able to deploy on mobile devices. Numerous object detection methods, such as YOLOv2, G-RCNN, and YOLO-MXANet, have been proposed to tackle real-time detection tasks. While YOLOv2 and other one-stage detectors offer the benefit of rapid processing, they often compromise detection accuracy in densely populated or occluded environments.

Models like G-RCNN combined with MCD-SORT show improvements in tracking but encounter challenges under motion blur and complex object overlaps. Similarly, techniques such as MS-CAB focus on enhancing the detection of smaller objects using multi-scale attention, yet are not optimal for applications in real-time due to high computational demands. The proposed Social-GAN-based FRCNN addresses these shortcomings by integrating spatial localization from FRCNN with contextual interaction modeling from Social-GAN.

Table 1. Comparative analysis of existing methods

Model	Strengths	Limitations
YOLOv2 [25]	Fast inference, real-time capable	Lower accuracy in dense or occluded scenes
G-RCNN + MCD-SORT [26]	Good tracking performance	Inference lag, lower speed; frame drops in real-time
YOLO-MXANet [27]	Enhances the detection of small objects	Cannot be deployed on mobile/embedded platforms
MS-CAB [28]	Better multi-scale feature fusion	High training time; lacks social interaction modeling

This combination enables the model to better understand object relationships and environmental context, thereby improving detection precision in crowded and dynamic traffic scenes.

Moreover, the inclusion of Gaussian filtering during preprocessing contributes to cleaner image inputs, enhancing overall detection performance. These innovations collectively distinguish the proposed method from existing approaches, making it a robust and efficient solution for real-world traffic monitoring.

A thorough comparison of the proposed Social-GAN-based FRCNN model with various state-of-the-art object detection techniques widely applied in traffic image analysis. This is illustrated in Table 1. The comparison focuses on critical aspects such as detection speed, accuracy, ability to handle occlusion, support for real-time deployment, and contextual understanding of object interactions. While existing models like YOLOv2 and YOLO-MXANet offer fast detection, they often compromise accuracy in complex environments.

Others, like G-RCNN and MS-CAB, show improvements in certain areas but fall short in real-time applicability or contextual reasoning. In contrast, the proposed model effectively combines the strengths of both Social-GAN and FRCNN to overcome these limitations.

It enhances object detection by incorporating social behaviour modelling and high-precision region proposals, thus enabling more accurate and reliable detection in dynamic traffic scenes. This comparison underscores the unique contributions of the proposed approach, along with the shortcomings of existing methods.

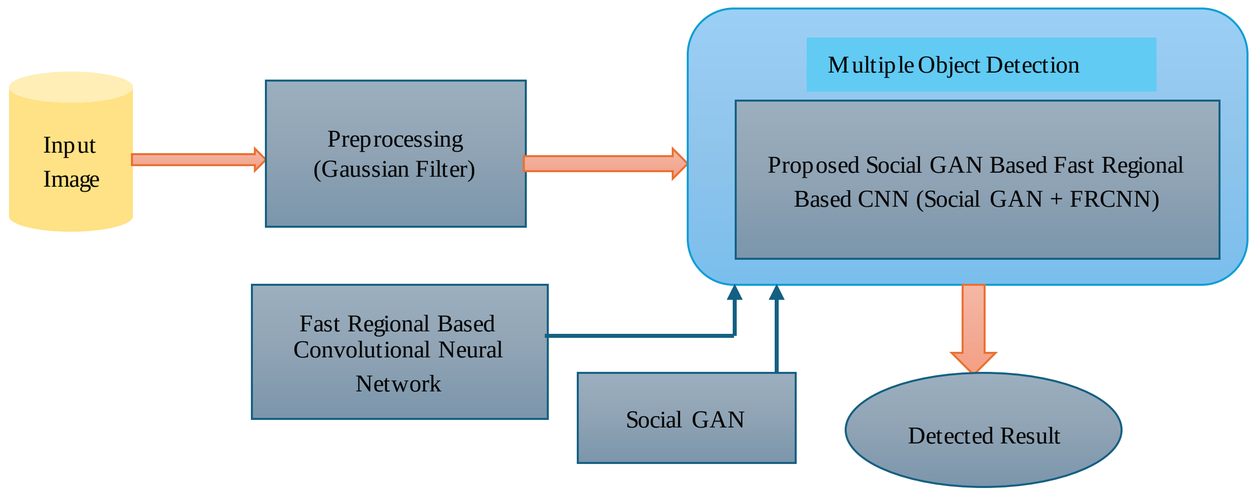

Fig. 1 Outlook of multiple object detection with social-Gan-based FRCNN

1.2. Problem Statement

Detecting objects in traffic images plays a crucial role in traffic surveillance, self-driving technologies, and smart city initiatives. Detecting multiple objects in real-time within traffic images is difficult due to factors like varying object sizes, complex backgrounds, and dynamic motion patterns in traffic environments. Existing methods often face challenges in balancing accuracy and processing speed, particularly in real-world scenarios.

Furthermore, traditional object detection models lack the ability to effectively handle interactions between multiple objects and fail to provide reliable detection performance under diverse conditions. Thus, there is a need for an approach that not only enhances detection accuracy but also ensures robust real-time performance for traffic monitoring systems. The proposed Social-GAN-based FRCNN effectively addresses the limitations highlighted in the problem statement. To handle interactions between multiple objects in traffic scenes, the integration of Social GAN models the social dynamics between objects, enabling better detection in crowded or complex environments. By leveraging both FRCNN and Social GAN, the method improves contextual perception and achieves higher precision in detecting objects across different scales and complexities. The proposed approach also incorporates Gaussian filtering in the preprocessing stage, which reduces image noise. This ensures that the detection process begins with enhanced image quality.

Moreover, the improved architecture of FRCNN enhances computational performance, enabling its effective deployment for real-time traffic monitoring.

1.3. Contribution

The key contribution of the research is presented as follows:

- An innovative model is presented for multiple object detection in traffic-based images using the proposed SG-based FRCNN.
- The SG-based FRCNN integrates FRCNN and Social-Gan by adding a regression layer, which is accomplished through the fusion of regression and Fractional Calculus (FC).

This remaining section of this paper is organized as follows: Section 2 outlines the methodology, Section 3 presents the results, Section 4 provides the discussion and Section 5 closes the research.

2. Research Method

This paper introduces an innovative model for multi-object detection in traffic images, utilizing the proposed Social-GAN-enhanced FRCNN. Initially, the image acquired from the MOT benchmark dataset illustrated in [29] is considered as an input for the whole processing. The input image is fed into the pre-processing module, where the external noise and distortions present in the images are eliminated using a Gaussian filter [30]. Finally, multiple object detection is carried out by merging two networks, namely FRCNN [31] and Social-Gan (SG) [32]. Figure 1 illustrates the outlook of multiple object detection using the proposed Social-Gan-based FRCNN.

2.1. Acquire a Set of Traffic Images

The initial set is to obtain the input images of the traffic scenario from the benchmark Multi-Object Tracking (MOT) dataset, which contains images captured from video sequences consisting of several objects like pedestrians, vehicles, traffic signals, and traffic signs existing on the road. Social-Gan-based FRCNN is programmed in Python. The evaluation is performed using the Multiple Object Tracking Benchmark [29]. It contains a large collection of images that pose detection for all sets of sequences and provides multiple object tracking. It contains various indoor and outdoor scenarios of public places, with pedestrians as interesting objects. It comprises various variants released each year, like MOT15, MOT17, MOT20, etc.

Considering traffic scenes, the technologies related to the visual perception of intelligent vehicles can aid automated driving models for perceiving complicated platforms precisely in time, which is a basic need to avoid collisions and offer safer driving. Due to the quick design of computer vision technologies, the perception of vehicle visuals is gradually being employed in automated driving. Assume a traffic database R with o traffic images manipulated as

$$R = \{\hbar_1, \hbar_2, \cdots, \hbar_l, \cdots, \hbar_o\} \tag{1}$$

Where o signifies total traffic images, $\hbar_\ell$ stands for ℓ^{th} traffic image.

2.2. Pre-Processing using a Gaussian Filter

Here, the input image $\hbar_\ell$ is provided to the Gaussian filter [27] for filtering noise. The aim is to enhance image quality by making it ready to perform enhanced processing by discarding or lessening the unrelated and superfluous components from the background. The traffic images are generally complex to examine and interpret. Thus, pre-processing tends to be a crucial part of enhancing the quality of images and helps to prepare the image for performing better multi-object detection. It is implemented through a Gaussian filter that aids in eliminating noise from the image. The two-dimensional Gaussian filter is manipulated as,

$$F(l, m) = \frac{1}{\sqrt{2\pi\alpha}} \exp\left(-\frac{(l^2 + m^2)}{2\alpha^2}\right) \tag{2}$$

Where, α^2denote variance considering the Gaussian filter, and the size of the filter kernel $n(-n \le l, m \le n)$ It is discovered by discarding values less than 5% of the maximal kernel value. The one-dimensional Gaussian filter is provided as,

$$F(l) = \frac{1}{\sqrt{2\pi\alpha}} \exp\left(\frac{-l^2}{2\alpha^2}\right) \tag{3}$$

A Gaussian filter is applied to reduce noise, with its variance effectively smoothing out noise and adjusting image components to regulate pixel brightness. It also adjusts the displacement of edge position and phantom edges. The filtered image obtained from the Gaussian filter is noted as B.

2.3. Multiple Object Detection Using Social-Gan-based FRCNN

With the rise of video surveillance, object detection has become an increasingly important focus in the computer vision industry. However, attaining elevated performance and object recognition in real-time is termed to be the foremost problem in embedded infrastructures. Hence, feasible and precise multiple object detection is essential because of rising security issues in various domains. Here, the aim is to provide a Social-Gan-based FRCNN model for multiple object detection from a scene.

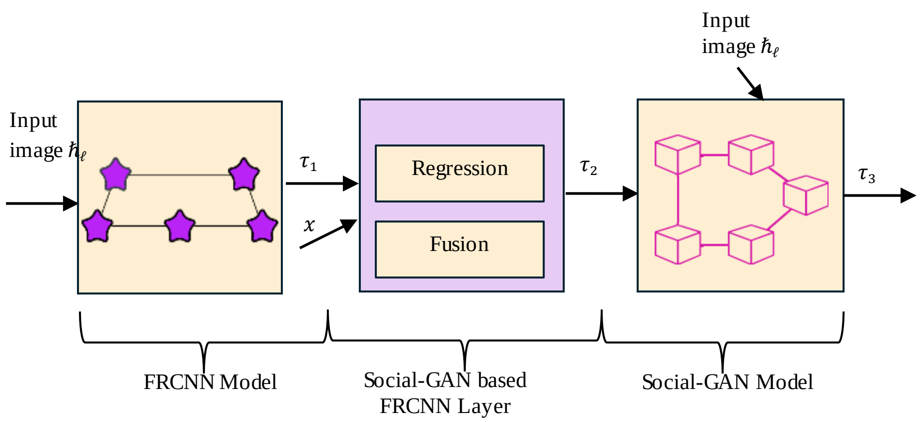

Fig. 2 Structure of the multiple object detection model using the Social-Gan-based FRCNN model

This information related to detection can help the surveillance cameras transmit real-time information about detected objects. This information can be helpful in discovering the existence of an object in a region and identifying a specific object in an area. Figure 2 details the structure of the proposed Social-GAN-based FRCNN model.

2.4. Algorithmic Representation of Proposed Social-GAN-Based FRCNN

To enhance clarity regarding the operational flow of the proposed model, the following pseudocode outlines the sequence of steps in the Social-GAN integrated Fast Regional CNN (FGRCNN).

Input: Traffic image dataset from MOT Benchmark
Output: Detected multiple objects with bounding boxes
1. Load the MOT dataset with traffic scenes
2. For each image in the dataset:
 a. Apply a Gaussian filter for noise reduction
 b. Resize image to 416×416 pixels
3. Initialize FRCNN for region proposal and feature extraction
4. Initialize Social-GAN for modeling object interactions
5. For each frame:
 a. Extract candidate object regions using FRCNN
 b. Use Social-GAN to analyze object interactions and behavior
 c. Fuse FRCNN and Social-GAN outputs
 d. Perform bounding box regression and object classification
6. Display or store the detected objects with confidence scores.

The flowchart representing the sequential workflow of the proposed Social-GAN-based FRCNN model is illustrated in Figure 3.

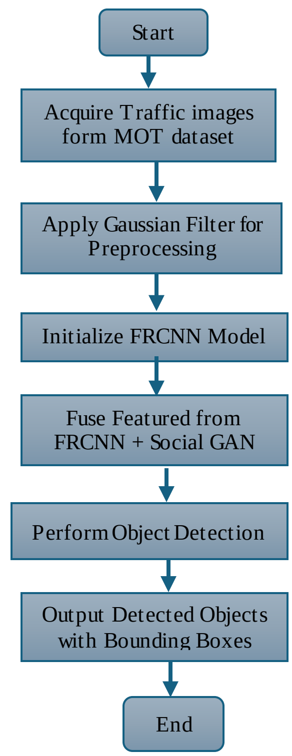

Fig. 3 Flowchart of the proposed FGRCNN (Social-GAN + FRCNN) model for multiple object detection in traffic images

3. Results and Discussion

The potential of Social-Gan-based FRCNN with former schemes is enumerated, considering different performance criteria by varying learning sets. The evaluation is performed using a Multiple Object Tracking Benchmark. It contains a large collection of images, posing detection for all sets of sequences and providing multiple object tracking. It contains various indoor and outdoor scenarios of public places, with pedestrians as interesting objects. It comprises various sets of variants released each year, like MOT15, MOT17, MOT20 and so on.

3.1. Experimental Setup

To ensure reproducibility of the experimental results, the proposed Social-GAN-based FRCNN model was implemented in Python 3.9 using key libraries including TensorFlow 2.8, Keras, OpenCV, and NumPy. The experiments were conducted on a system running Windows 11 (64-bit) equipped with an Intel Core i7-12700K processor, 32 GB RAM, and an NVIDIA RTX 3080 GPU with 10 GB VRAM, with CUDA 11.2 enabled for GPU acceleration. The input dataset comprised variants of the Multiple Object Tracking (MOT) Benchmark-specifically MOT15, MOT17, and MOT20-covering diverse traffic scenes under different lighting and occlusion conditions. All input images were resized to 416×416 pixels and preprocessed using a Gaussian filter to remove noise and enhance edge features. The dataset was split into 90% for training and 10% for testing. The model was trained for 50 epochs with a batch size 16 using the Adam optimizer and a learning rate of 0.0001. The loss function combined cross-entropy loss for classification with localization loss for bounding box regression. A five-fold cross-validation (K = 5) strategy was adopted to ensure statistical reliability, and performance metrics such as accuracy, TPR, TNR, PPV, and NPV were computed across all folds.

3.1.1. Training and Testing Procedure

Input Preprocessing: All images were resized to 416×416 pixels and processed with a Gaussian filter to minimize noise and enhance edge details. A data split of 90% for training and 10% for testing was applied. The training process spanned 50 epochs with a batch size set to 16, and the Adam optimizer with a 0.0001 learning rate was utilized to ensure smooth convergence. The training phase optimized a composite loss function that included cross-entropy loss for classification and localization loss for bounding box regression. Standard metrics, including Accuracy, TPR, TNR, PPV, and NPV, were considered to evaluate the model's performance. All experiments were repeated over five random training splits (K-Fold = 5) to ensure statistical validity. Average values across folds are reported in the results section.

Fig. 4 Experimental upshots using the input image of Social-Gan-based FRCNN

3.2. Experimental Outcomes

The experimental outcomes are explained in this section. Figure 4 portrays the input image, while Figure 5 notates the preprocessed image obtained using a Gaussian filter. Figure 6 portrays an image of objects detected. The schemes engaged in the assessment are YOLOv2 G-RCNN+MCD-SORT, YOLO-MXANet, MS-CAB, and proposed Social-Gan+FRCNN. The evaluation of the proposed Social-Gan+FRCNN considering the learning set and K-group is enumerated considering different performance criteria.

Fig. 5 Pre-processed image of Social-Gan-based FRCNN

Fig. 6 Object detected image of Social-Gan based FRCNN

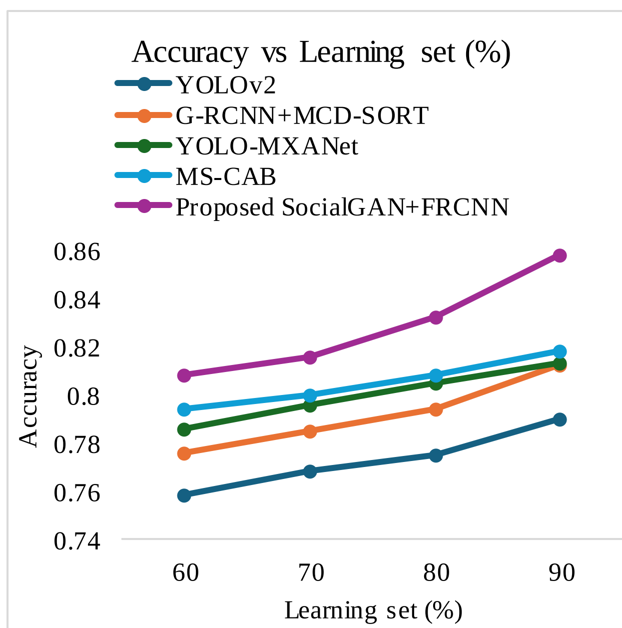

Fig. 7 Efficiency assessment considering learning set using accuracy of Social-Gan+FRCNN

3.2.1. Estimation Considering Learning Set

Figure 7 deliberates the Estimation of the proposed Social-Gan+FRCNN considering the learning set. The accuracy-related graph is specified in the figure considering the learning set as 90%, the accuracy produced by YOLOv2 is 0.788, G-RCNN+MCD-SORT is 0.799, YOLO-MXANet is 0.809, MS-CAB is 0.817, and Social-Gan+FRCNN is 0.858. The TPR-assisted graph is depicted in Figure 8. Using the learning set as 90%, the augmented TPR of 0.937 is generated by Social-Gan+FRCNN, whilst the TPRs of YOLOv2, G-RCNN+MCD-SORT, YOLO-MXANet, and MS-CAB are 0.848, 0.858, 0.866, and 0.898.

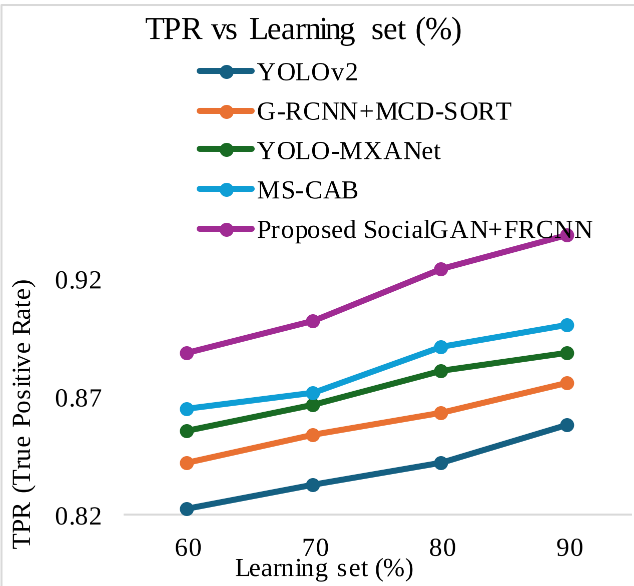

Fig. 8 Using TPR of Social-Gan+FRCNN

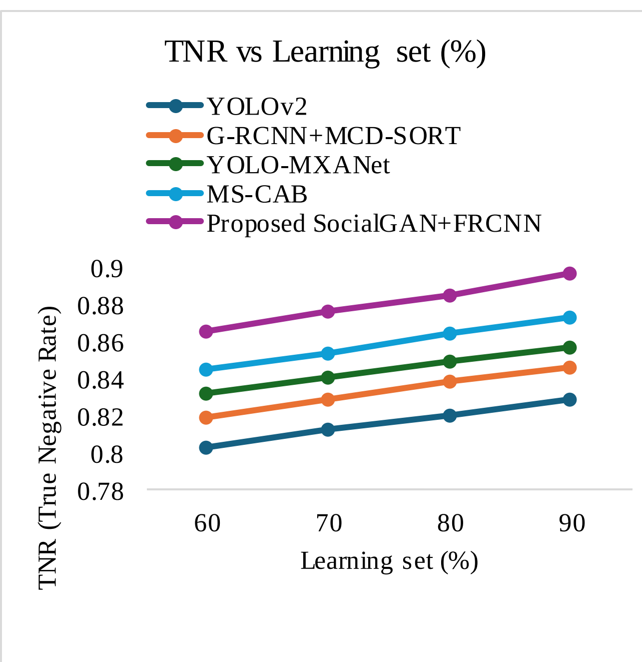

Fig. 9 Using TNR of Social-Gan+FRCNN

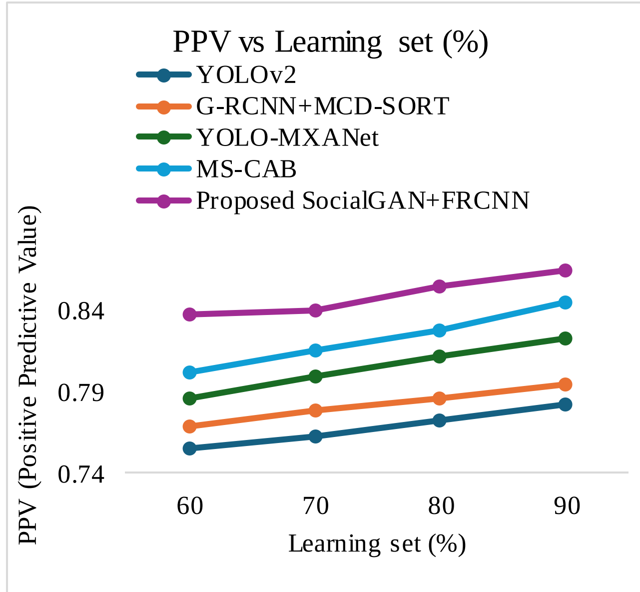

Fig. 10 Using PPV of Social-Gan+FRCNN

The graph concerning TNR is specified in Figure 9. Taking learning set as 90%, the TNR produced is 0.817 for YOLOv2, 0.838 for G-RCNN+MCD-SORT, 0.848 for YOLO-MXANet, 0.877 for MS-CAB and 0.898 for Social-Gan+FRCNN. The PPV-based graph is shown in Figure 10.

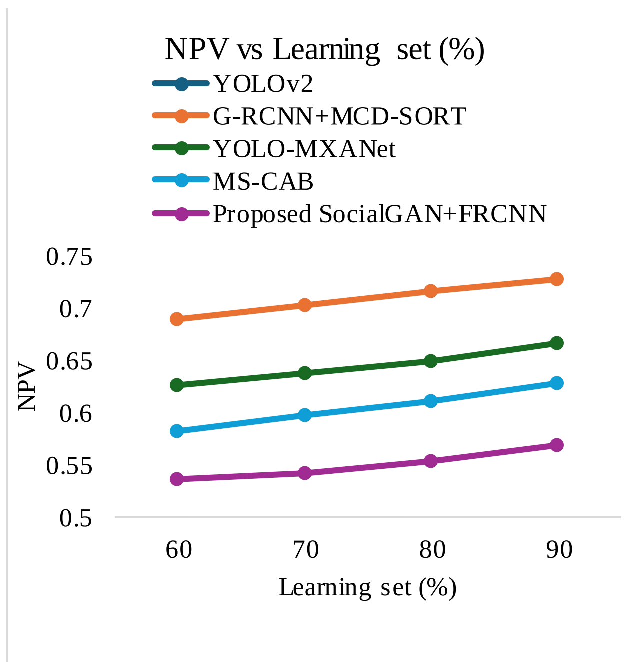

Fig. 11 Using NPV of Social-Gan+FRCNN

Adapting the learning set as 90%, the PPV generated by YOLOv2, G-RCNN+MCD-SORT, YOLO-MXANet, MS-CAB are 0.777, 0.798, 0.818, 0.848, while the PPV of the proposed Social-Gan+FRCNN is 0.865. The NPV graph is notated in Figure 11.

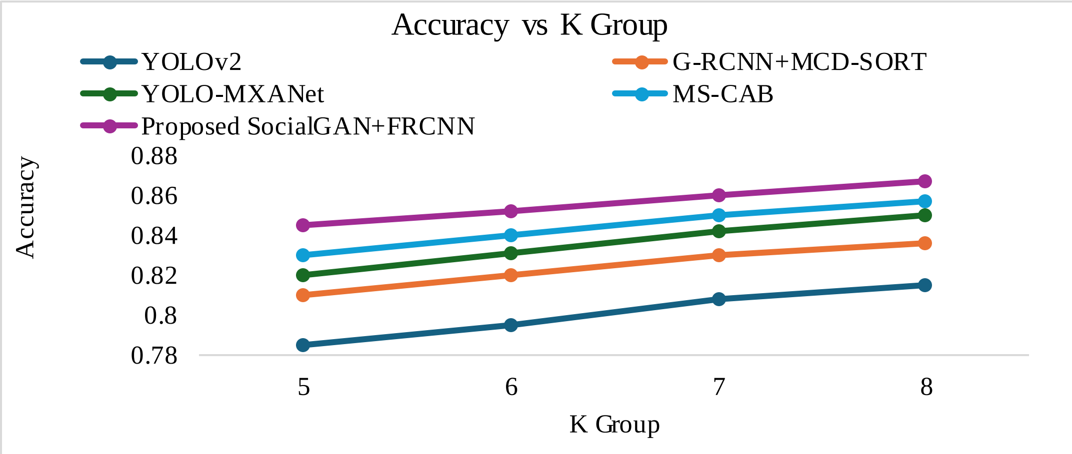

Fig. 12 Efficiency assessment considering K-group using Accuracy of Social-Gan+FRCNN

The NPVs generated are 0.787, 0.737, 0.688, and 0.628 for YOLOv2, G-RCNN+MCD-SORT, YOLO-MXANet, and MS-CAB, whereas for Social-Gan+FRCNN is 0.568 by considering the learning set as 90%. The Social-GAN improves the model's capacity to interpret relationships and interactions among objects, which is vital for handling complex traffic situations. This improves the performance of the proposed model. The remaining section shows the Estimation considering the K-Group.

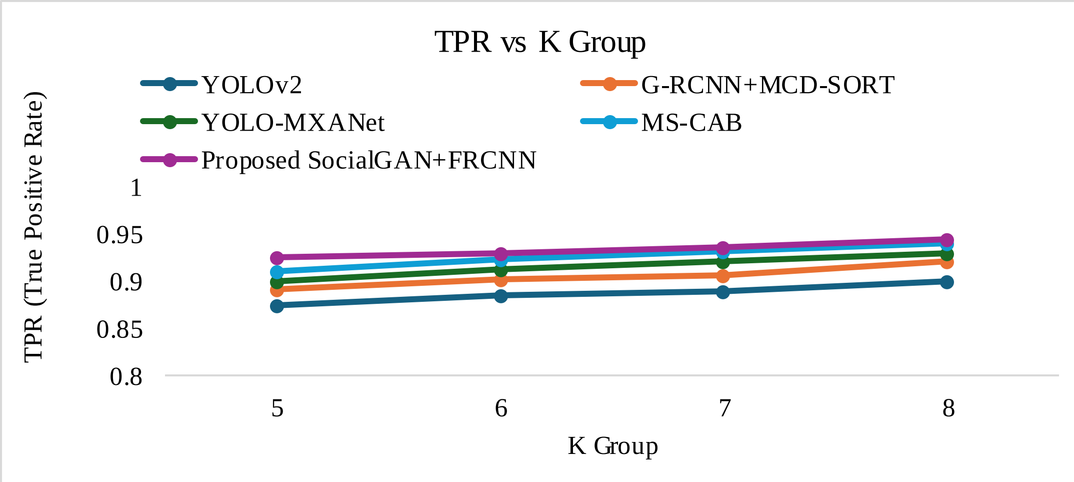

Fig. 13 Efficiency assessment considering K-group using TPR of Social-Gan+FRCNN

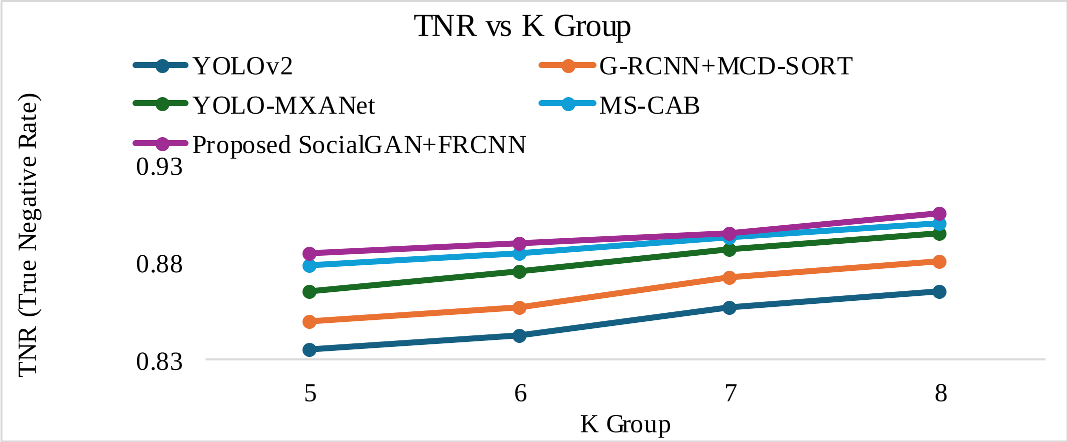

Fig. 14 Efficiency assessment considering K-group using TNR of Social-Gan+FRCNN

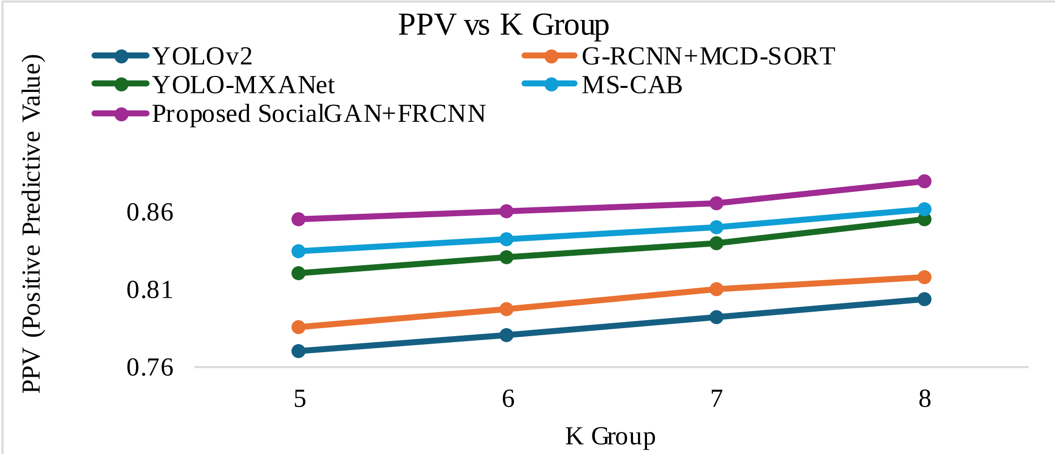

Fig. 15 Efficiency assessment considering K-group using PPV of Social-Gan+FRCNN

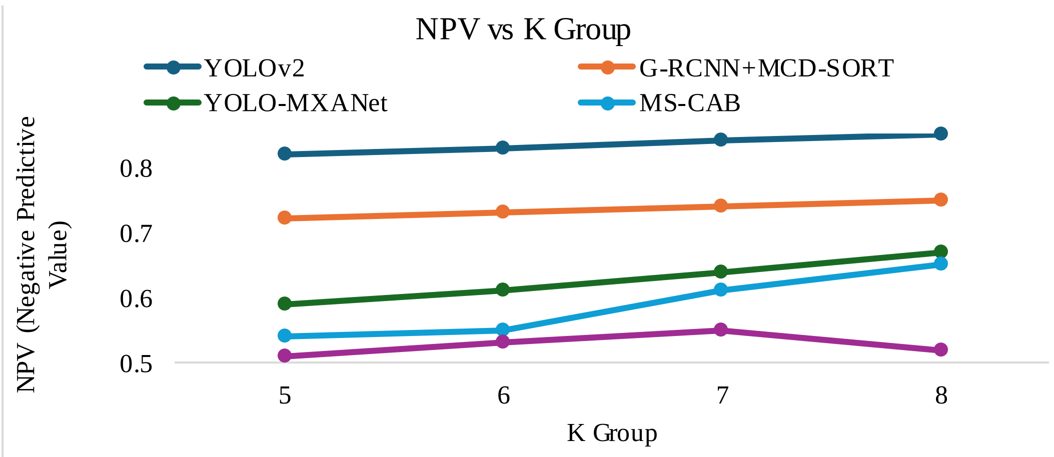

Fig. 16 Efficiency assessment considering K-group using NPV of Social-Gan + FRCNN

Figures 12-16 deliberate the Estimation of the proposed Social-Gan+FRCNN considering the learning set. The accuracy-related graph is specified in Figure 12. Considering K-Group as 8, the accuracy produced by YOLOv2 is 0.809, G-RCNN+MCD-SORT is 0.817, YOLO-MXANet is 0.838, MS-CAB is 0.848, and Social-Gan+FRCNN is 0.865.

The TPR-assisted graph is depicted in Figure 13. With K-Group as 8, the elevated TPR of 0.947 is generated by Social-Gan+FRCNN, while the TPRs of YOLOv2, G-RCNN+MCD-SORT, YOLO-MXANet, and MS-CAB are 0.865, 0.886, 0.898, and 0.918. The graph concerning TNR is specified in Figure 14. Taking K-Group as 8, the TNR generated is 0.828 for YOLOv2, 0.849 for G-RCNN+MCD-SORT, 0.858 for YOLO-MXANet, 0.877 for MS-CAB and 0.909 for Social-Gan+FRCNN. The PPV-based graph is shown in Figure 15. Adapting K-Group as 8, the PPV produced by YOLOv2, G-RCNN+MCD-SORT, YOLO-MXANet, MS-CAB are 0.800, 0.827, 0.848, 0.868, while that of Social-Gan+FRCNN is 0.888. The NPV graph is notated in Figure 16. The NPVs generated are 0.847, 0.758, 0.658, 0.629 for YOLOv2, G-RCNN+MCD-SORT, YOLO-MXANet, MS-CAB, while the NPV of Social-Gan+FRCNN is 0.558, assuming K-Group as 8. Overall, Social-GAN+FRCNN outperforms other methods across accuracy, TPR, TNR, and PPV metrics, demonstrating its robustness and effectiveness in object detection tasks.

3.3. Comparative Discussion

Table 2 deliberates the Estimation considering various performance aspects by changing the learning set and K-group. With the learning set, the augmented accuracy of 85.8% is obtained by Social-Gan+FRCNN, whilst accuracy obtained by former schemes is 78.8%, 79.9%, 80.9%, and 81.7%. The high TPR of 93.7% is generated by Social-Gan+FRCNN, whilst the TPR obtained by former schemes is 84.8%, 85.8%, 86.6%, and 89.8%. The high TNR of 89.8% is generated by Social-Gan+FRCNN, whilst the TNR obtained by former schemes is 81.7%, 83.8%, 84.8%, and 87.7%. The high PPV of 86.5% is generated by Social-Gan+FRCNN, whilst the PPV obtained by former schemes is 77.7%, 79.8%, 81.8%, and 84.8%. The high NPV of 56.8% is generated by Social-Gan+FRCNN, whilst the PPV obtained by former schemes is 78.7%, 73.7%, 68.8%, and 62.8%. Assuming the K-group, the high accuracy of 86.5%, TPR of 94.7%, TNR of 90.9%, NPV of 88.8% and PPV of 55.8%is computed by Social-Gan+FRCNN. The highest values of metrics expose that the proposed Social-Gan+FRCNN is effective in detecting multiple objects from a scene. The improved performance of the proposed Social-GAN-based FRCNN method is due to the integration of Social GAN and FRCNN. Furthermore, the results suggest that integrating Social GAN with FRCNN could be a promising approach for similar real-time object detection tasks in other domains.

Table 2. Comparative estimation

Variation	Metrics	YOLOv2	G-RCNN+MCD-SORT	YOLO-MXANet	MS-CAB	Proposed Social-Gan+FRCNN
Learning Set	Accuracy (%)	78.8	79.9	80.9	81.7	85.8
	TPR (%)	84.8	85.8	86.6	89.8	93.7
	TNR (%)	81.7	83.8	84.8	87.7	89.8
	PPV (%)	77.7	79.8	81.8	84.8	86.5
	NPV (%)	78.7	73.7	68.8	62.8	56.8
K - Group	Accuracy (%)	80.9	81.7	83.8	84.8	86.5
	TPR (%)	86.5	88.6	89.8	91.8	94.7
	TNR (%)	82.8	84.9	85.8	86.8	88.8
	PPV (%)	80	82.7	84.8	86.8	88.8
	NPV (%)	84.7	75.8	65.8	62.9	55.8

4. Conclusion

This paper deliberates a model for multiple object detection in traffic images using the proposed Social-Gan-based FRCNN. The method used is effective in operating the model in challenging weather conditions. The model is helpful in providing real-time detection of objects with enhanced accuracy and speed to help the driver.

The Social-Gan+FRCNN offered enhanced efficacy with enhanced accuracy of 86.5%, TPR of 94.7%, TNR of 90.9%, PPV of 88.8% and NPV of 55.8. Moreover, the proposed model is suitable for integration into vehicles' Electronic Control Unit (ECU). Future work will investigate the model's scalability for broader applications, such as smart city traffic management.

References

[1] Umesh Parameshwar Naik et al., "Implementation of YOLOv4 Algorithm for Multiple Object Detection in Image and Video Dataset using Deep Learning and Artificial Intelligence for Urban Traffic Video Surveillance Application," *2021 Fourth International Conference on Electrical, Computer and Communication Technologies (ICECCT)*, Erode, India, pp. 1-6, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[2] Jiangfan Feng et al., "A Multibranch Object Detection Method for Traffic Scenes," *Computational Intelligence and Neuroscience Computational Intelligence and Neuroscience*, vol. 2019, pp. 1-16, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[3] Barret Zoph et al., "Learning Data Augmentation Strategies for Object Detection," *16th European Conference: Computer Vision – ECCV 2020*, Glasgow, UK, pp. 566-583, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[4] Ratnababu Mamidi, and S.N. Merchant, "Efficient Bayes Saliency-Based Object Detection on Images using Deep Belief Networks," *Turkish Journal of Computer and Mathematics Education*, vol. 10, no. 3, pp. 980-986, 2019. [Publisher Link]

[5] Sk. Arif Ahmed et al., "Trajectory-Based Surveillance Analysis: A Survey," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 29, no. 7, pp. 1985-1997, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[6] Zeinab Ghasemi Darehnaei et al., "Two-Level Ensemble Deep Learning for Traffic Management using Multiple Vehicle Detection in UAV Images," *International Journal of Smart Electrical Engineering*, vol. 10, no. 3, pp. 127-133, 2021. [Google Scholar] [Publisher Link]

[7] Shoaib Azam, Aasim Rafique, and Moongu Jeon, "Vehicle Pose Detection Using Region Based Convolutional Neural Network," *2016 International Conference on Control, Automation and Information Sciences (ICCAIS)*, Ansan, Korea (South), pp. 194-198, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[8] Y. Tian et al., "Pedestrian Detection Aided by Deep Learning Semantic Tasks," *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, USA, pp. 5079-5087, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[9] Xianzhi Du et al., "Fused DNN: A Deep Neural Network Fusion Approach to Fast and Robust Pedestrian Detection," *2017 IEEE Winter Conference on Applications of Computer Vision (WACV)*, Santa Rosa, CA, USA, pp. 953-961, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[10] Jiachen Yang et al., "Aircraft Detection in Remote Sensing Images Based on a Deep Residual Network and Super-Vector Coding," *Remote Sensing Letters*, vol. 9, no. 3, pp. 228-236, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[11] Chenyi Chen et al., "DeepDriving: Learning Affordance for Direct Perception in Autonomous Driving," *2015 IEEE International Conference on Computer Vision (ICCV)*, Santiago, Chile, pp. 2722-2730, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[12] Joseph Redmon et al., "You Only Look Once: Unified, Real-Time Object Detection," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, pp. 779-788, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[13] Shaoqing Ren et al., "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137-1149, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[14] Licheng Jiao et al., "A Survey of Deep Learning-Based Object Detection," *IEEE Access*, vol. 7, pp. 128837-128868, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[15] Ross Girshick et al., "Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation," *2014 IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, OH, USA, pp. 580-587, 2014. [CrossRef] [Google Scholar] [Publisher Link]

[16] Ross Girshick, "Fast R-CNN," *2015 IEEE International Conference on Computer Vision (ICCV)*, Santiago, Chile, pp. 1440-1448, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[17] Perumalsamy Gunasekaran, A. Azhagu Jaisudhan Pazhani, and T.Ajith Bosco Raj, "A Novel Method for Multiple Object Detection on Road Using Improved YOLOv2 Model," *Informatica*, vol. 46, no. 4, pp. 567-574, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[18] Ibrahim Soliman Mohamed, and Lim Kim Chuan, "PAE: Portable Appearance Extension for Multiple Object Detection and Tracking in Traffic Scenes," *IEEE Access*, vol. 10, pp. 37257-37268, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[19] Anima Pramanik et al., "Granulated RCNN and Multi-Class Deep SORT for Multi-Object Detection and Tracking," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 6, no. 1, pp. 171-181, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[20] Jiachen Yang et al., "A RGB-D Based Real-Time Multiple Object Detection and Ranging System for Autonomous Driving," *IEEE Sensors Journal*, vol. 20, no. 20, pp. 11959-11966, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[21] Jing Lian et al., "Small Object Detection in Traffic Scenes Based on AttentionFeature Fusion," *Sensors*, vol. 21, no. 9, pp. 1-16, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[22] Jamuna S. Murthy et al., "ObjectDetect: A Real-Time Object Detection Framework for Advanced Driver Assistant Systems Using YOLOv5," *Wireless Communications and Mobile Computing*, vol. 2022, pp. 1-10, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[23] Chenge Li et al., "TrackNet: Simultaneous Object Detection and Tracking and Its Application in Traffic Video Analysis," *arXiv Preprint*, pp. 1-10, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[24] Xiaowei He et al., "Small Object Detection in Traffic Scenes Based on YOLO-MXANet," *Sensors*, vol. 21, no. 21, pp. 1-16, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[25] Hyochang Ahn, and Han-Jin Cho, "Research of Multi-Object Detection and Tracking using Machine Learning based on Knowledge for Video Surveillance System," *Personal and Ubiquitous Computing*, vol. 26, pp.385-394, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[26] Qibin He et al., "Multi-Object Tracking in Satellite Videos with Graph-based Multitask Modeling," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1-13, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[27] Liang Hou et al., "Multi-Object Detection Method in Construction Machinery Swarm Operations Based on the Improved YOLOv4 Model," *Sensors*, vol. 22, no.19, pp. 1-14, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[28] Xinyue Gao et al., "DetTrack: An Algorithm for Multiple Object Tracking by Improving Occlusion Object Detection," *Electronics*, vol. 13, no.1, pp. 1-16, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[29] Multiple Object Tracking Benchmark Dataset, MOT Challenge. [Online]. Available: https://motchallenge.net/

[30] G. Deng, and L.W. Cahill, "An Adaptive Gaussian Filter for Noise Reduction and Edge Detection," *1993 IEEE Conference Record Nuclear Science Symposium and Medical Imaging Conference*, San Francisco, CA, USA, vol. 3, pp. 1615-1619, 1993. [CrossRef] [Google Scholar] [Publisher Link]

[31] Chandra Has Singh et al, "FRCNN-Based Reinforcement Learning for Real-Time Vehicle Detection, Tracking and Geolocation from UAS," *Drones*, vol. 6, no. 12, pp. 1-22, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[32] Agrim Gupta et al., "Social GAN: Socially Acceptable Trajectories with Generative Adversarial Networks," *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, pp. 2255-2264, 2018. [CrossRef] [Google Scholar] [Publisher Link]