ISSN 2277-2685
IJESR/April-June. 2025/ Vol-15/Issue-2s/1185-1190
Ch.Sri Vidya *et. al.*, / International Journal of Engineering & Science Research

Nocturnal Surveillance Framework For Vehicle Detection

Dr B Raveendranadh Singh[1], Ch.Sri Vidya[2], Paripelli Sreenidhi[3]
[1]HOD, Department of CSE, Bhoj Reddy Engineering College for Women, India
[2,3]B.Tech Students, Department of CSE, Bhoj Reddy Engineering College for Women, India

ABSTRACT:

In the field of traffic surveillance systems, where effective traffic management and safety are the primary concerns, vehicle detection and tracking play an important role. Low brightness, low contrast, and noise are issues with low-light environments that result from poor lighting or insufficient exposure.

In this paper, we proposed a vehicle detection and tracking model based on the aerial image captured during nighttime. Before object detection, we performed fogging and image enhancement using MIRNet architecture. After pre-processing, YOLOv5 was used to locate each vehicle position in the image. Each detected vehicle was subjected to a Scale-Invariant Feature Transform (SIFT) feature extraction algorithm to assign a unique identifier to track multiple vehicles in the image frames.

To get the best possible location of vehicles in the succeeding frames templates were extracted and template matching was performed. The proposed model achieves a precision score of 0.924 for detection and 0.861 for tracking with the Unmanned Aerial Vehicle Benchmark Object Detection and Tracking (UAVDT) dataset, 0.904 for detection, and 0.833 for tracking with the Vision Meets Drone Single Object-Tracking (VisDrone) dataset.

1. INTRODUCTION

Vehicle recognition in aerial images is crucial for both military and civilian applications. Militar y target strikes and traffic control can both benefit from the use of this technology. Researchers have proposed various techniques for object recognition in aerial photos in the daytime with sufficient lightning, producing remarkable results.

However, vehicle detection of objects in low light conditions is a challenging and significant issue with surveillance camera applications. In low-illumination conditions there is less information available and difficulty extracting enough useful features, such as at night time there is background light interference, the object is underexposed, and brightness and contrast are poor which results in low-image quality.

Recent studies on low lighting concentrate on image enhancement to improve basic visual properties in the pre- processing steps. These methods include global and local enhancement techniques. The global enhancement techniques when applied to night-time images may over- expose already bright parts of the images, however, local contrast enhancement methods focus on image details, but it increases noise when contrast gain is high.

2. THE PROPOSED FRAMEWORK

depicts the general architecture of the proposed model. The model mainly consists of five modules: (i) pre-processing steps to enhance the brightness level of nighttime images; (ii) vehicle detection using the deep learning model YOLOv5; (iii) SIFT feature extraction of each detected vehicle and identifier assignment; (iv) vehicle tracking using the template matching algorithm; and

(v) drawing trajectories of each tracked vehicles. Each module of the framework is explained in detail in the following subsections. a machine inspection dataset, the suggested system is tested, and the findings demonstrate that it outperforms a number of cutting-edge object recognition

IMAGE PRE-PROCESSING

1) DEFOGGING

The input images extracted from nighttime traffic videos at the rate of 8 FPS were first resized to 768 × 768 coordinates. To denoise the image, we applied the defogging method which estimated the intensity of noise in each image pixel and then removed as follows: $I(x) = U(x)t(x) + K(1 - t(x))$.

where x represents the location of the pixel, K is the density of the fog, and t(x) is the transmission map. The visualization of the defogging process

LOW-LIGHT ENHANCEMENT USING MIRNet

After denoising the images, the next step is to enhance the brightness level of the images to locate the objects easily. For this purpose, we used the pre-trained model MIRNet. All the images are passed onto the contrast enhancement module. MIRNet is a pre-trained fully convolutional deep learning architecture that retains spatially exact high-resolution representations over the whole network while receiving significant contextual information from the low-resolution representations. The model consists of a feature extraction module that maintains the high- resolution original features to reserve fine spatial details while computing a complementary collection of features at various spatial scales. Also, the characteristics from numerous multi- resolution branches are gradually integrated for better representation learning using a recurring information exchange mechanism. It uses a technique for fusing features from different scales that correctly maintains the original information of the feature at each spatial level while dynamically combining varying receptive fields. To simplify the learning process, the recursive residual gradually decomposes the input image, enabling deep networks to be built. The output of the MIRNet image enhancement is shown in Fig. 2.2 Also, the overall architecture of MIRNet

YOLO v5-BASED VEHICLE DETECTION

Because of its high-performance capabilities, YOLO algorithms are frequently used in object detection systems, especially for vehicle detection tasks. YOLO sees an image as a regression problem with fast speed. While training, YOLO takes the entire image as input training, paying more attention to global information for target detection and returns the position of the object bounding box.

3-EXPERIMENTS AND RESULTS

The experiments were conducted using a laptop with an Intel Core i5-8550U 1.80GHz processor, 6GB of Random Access Memory (RAM), Windows 10 running on the x64 architecture, and the Python tool. Also, to compare the performance of CPU and GPU. We ran the experiment on Tesla K80 GPU which is available free on Google Colab. The training time on the CPU was 1.3 hours whereas it took 0.86 hours to train on the GPU. However, there was no difference in the precision values. The proposed model produces remarkable results when tested on two benchmark datasets: UAVDT and VisDrone datasets.

DATASETs

VisDrone DATASET

The Vision Meets Drone Single Object-Tracking (VisDrone) dataset contains 288 clips of videos with a total of 261,908 frames and 10,209 still photos taken by several drones equipped with cameras and covering a variety of places. We used traffic image sequences taken at nighttime to test our model. Some of the sample images from the VisDrone dataset are displayed in Fig. 3.1

ISSN 2277-2685
IJESR/April-June. 2025/ Vol-15/Issue-2s/1185-1190
Ch.Sri Vidya *et. al.,* / International Journal of Engineering & Science Research

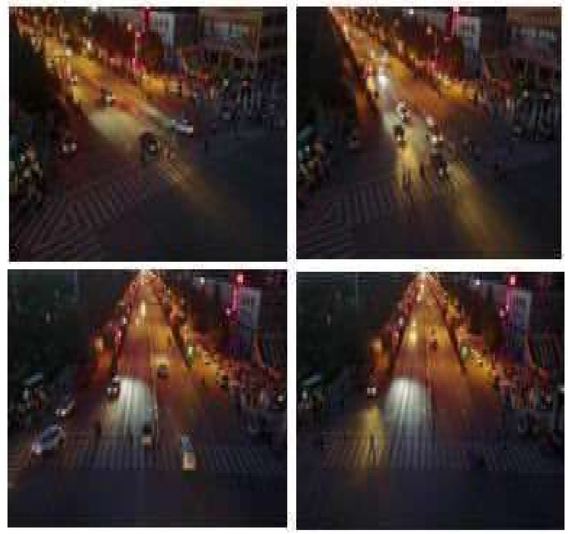

FIG3.1 Sample frames from the VisDrone dataset

TABLE 3.1.1 Precision, Recall, and F1-score for the detection algorithm.

Datasets	Precision	Recall	F1-score
UAVDT	0.924	0.915	0.92
VisDrone	0.904	0.892	0.90

TABLE 3.1.2 Precision, Recall, and F1-score for the tracking algorithm.

Datasets	Precision	Recall	F1-score
UAVDT	0.861	0.881	0.87
VisDrone	0.833	0.830	0.83

EVALUATION OF DETECTION AND TRACKING ALGORITHM

We used three performance metrics to assess our proposed detection and tracking algorithm specially designed for low illumination conditions: Precision, Recall and F1-score. These parameters are calculated

ISSN 2277-2685
IJESR/April-June. 2025/ Vol-15/Issue-2s/1185-1190
Ch.Sri Vidya *et. al.,* / International Journal of Engineering & Science Research

as follows:

Precision = TruePositive (TruePositive + FalsePositive) Recall = TruePositive (TruePositive + FalseNegative) F1 = 2 × Precision × Recall / Precision + Recall

COMPARISION WITH OTHER METHODS

We compared our proposed model with other methods in terms of precision score. Our model outperforms other techniques for both vehicle detection and tracking. Table 3.3.1 demonstrates the contrast of our proposed detection model with other methodologies. Table 3.3.2 shows the comparison of our proposed tracking algorithm with other methodologies. A comparison of detection and tracking techniques with state-of-the-art techniques has been demonstrated in the Tables 3.3.3 and 3.3.4.

Datasets	Methods	Precision	Tracking speed(seconds/frame)
UAVDT	HOG+Template Matching	0.68	160
	SURF+Template Matching	0.75	
	SIFT+Template matching	**0.861**	**214**
VisDrone	HOG+Template Matching	0.70	173
	SURF+Template Matching	0.77	226
	SIFT+Template matching	**0.833**	**231**

TABLE 3.3.1 Comparison of detection algorithm with other methods.

Datasets	Methods	Precision	Tracking speed(seconds/frame)
UAVDT	HOG+Template Matching	0.68	160
	SURF+Template Matching	0.75	
	SIFT+Template matching	**0.861**	**214**
VisDrone	HOG+Template Matching	0.70	173
	SURF+Template Matching	0.77	226
	SIFT+Template matching	**0.833**	**231**

TABLE 3.3.2 Comparison of tracking algorithm with other methods.

TABLE 3.3.3 Comparison of detection algorithm with state-of-the-art techniques.

Datasets	Methods	Precision
UAVDT	NDFT [52]	0.520
	FPN [52]	0.490
	Our Model [nighttime sequences]	**0.924**
VisDrone	GA-FPN [53]	0.82
	MRCN [54]	0.81
	Our Model [nighttime sequences]	**0.904**

TABLE 3.3.4 Comparison of tracking algorithm with state-of-the-art techniques.

Datasets	Methods	Precision
UAVDT	ASRDCT [55]	0.762
	Our Model [nighttime sequences]	**0.861**
VisDrone	Affinity Network [56]	0.74
	Our Model [nighttime sequences]	**0.833**

4-CONCLUSION

In this study, we propose a lightweight and efficient vehicle detection and tracking algorithm specially designed for low illumination conditions. First of all , we pre-processed the nighttime traffic scenes to adjust the brightness level of the image. Then, we applied semantic segmentation based on FCM clustering to segment the image into multiple uniform regions to reduce the overall complexity. For detection, we used YOLOv5 which can detect small objects precisely. We assign identifiers based on SIFT features to track multiple vehicles within a single image frame. Then, template matching was employed to get each vehicle's possible location and its corresponding identifier was retrieved by SIFT feature matching. The evaluation experimentation on public datasets demonstrates that our proposed framework can efficiently detect and track automobiles and outperforms other methods. In the future, we aim to enhance vehicle monitoring techniques to adapt to more complex traffic scenarios.

5-REFERENCES

[1] Y.Chen, N. Li, D. Zhu, C. C. Zhou, Z. Hu, Y. Bai, and J. Yan, ''BEVSOC: Self-supervised contrastive learning for calibration-free BEV 3-D object detection,'' IEEE Internet Things J., vol. 11, no. 12, pp. 22167–22182, Jun. 2024, doi: 10.1109/jiot.2024.3379471.

[2] J. Li, C. Zhang, Z. Liu, R. Hong, and H. Hu, ''Optimal volumetric video streaming with hybrid

saliency based tiling,'' IEEE Trans. Multimedia, vol. 25, pp. 2939–2953, 2022, doi: 10.1109/TMM.2022.3153208.

[3] L. Toloşi and T. Lengauer, ''Classification with correlated features: Unreliability of feature ranking and solutions,'' Bioinformatics, vol. 27, no. 14, pp. 1986–1994, Jul. 2011.

[4] J. Li, L. Han, C. Zhang, Q. Li, and Z. Liu, ''Spherical convolution empowered viewport prediction in 360 video multicast with limited FoV feedback,'' ACM Trans. Multimedia Compute., Commun., Appl., vol. 19, no. 1, pp. 1–23, Jan. 2023, doi: 10.1145/3511603.

[5] R. Zhang, L. Li, Q. Zhang, J. Zhang, L. Xu, B. Zhang, and B. Wang, ''Differential feature awareness network within antagonistic learning for infrared-visible object detection,'' IEEE Trans. Circuits Syst. Video Technol., p. 1, 2023, doi: 10.1109/TCSVT.2023.3289142.