Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 32 No. 5s (2025)

Optimized Vehicle Recovery: Leveraging AI and CCTV for Effective Tracking

[1,*] **Priya R. L,** [2]**Sharmila Sengupta,** [3]**Gautam Wadhwani,** [4]**Tanay Phatak,** [5]**Tanmay Thakare,** [6]**Teesha Karotra**

[1,2,3,4,5,6]Department of Computer Engineering, Vivekanand Education Society's Institute of Technology, Chembur, Mumbai, India.

*Email : priya.rl@ves.ac.in

Article History:

Received: *16-10-2024*

Revised: *26-11-2024*

Accepted: *12-12-2024*

Abstract:

Vehicle theft is a rising crime across the globe. GPS trackers are not a standard fit in all vehicles and the lack of internet connectivity makes tracking down stolen vehicles challenging for officials. GPS signals can be affected by different environmental factors such as terrain, or dense foliage, which can cause inaccuracies in location tracking. In order to address these problems, the suggested solution makes use of the CCTV system to follow the route of a stolen car and speed up the search effort. The suggested concept presents a successful method for detecting and monitoring stolen cars through CCTV videos, utilizing advanced computer vision and deep learning methods. Reviewing the CCTV footage from various sources can expedite the reaction and improve the retrieval while upholding data privacy. The accuracy of predicting car color and car model was above 90% in the results. The car routes are forecasted with RMSE score of 0.000171. The suggested model proposes a tracking solution that utilizes the current CCTV infrastructure without requiring any extra hardware to be installed in the vehicles.

Keywords: Vehicle tracking, Computer vision, Deep learning, Artificial Intelligence, Optical Character Recognition.

1. Introduction

Vehicle theft remains a persistent challenge for law enforcement agencies worldwide, demanding innovative strategies to combat this crime effectively. In India, major cities like New Delhi have seen a significant increase in vehicle theft, with reported incidents rising 2.5 times from 2022 to 2023 [1]. Traditionally, law enforcement has relied on labour-intensive fieldwork to gather information from eyewitnesses and local sources, a method often plagued by inefficiencies and inaccuracies.

Recent technological advancements have introduced more sophisticated approaches to tackling vehicle theft. One promising avenue involves harnessing surveillance cameras installed in public spaces, which continuously capture extensive visual data, including criminal activities such as vehicle theft. However, manual analysis of this vast and complex data is impractical. To address this challenge, law enforcement agencies are increasingly turning to smart technologies powered by artificial intelligence (AI) and computer vision.

A promising technique involves using Artificial Intelligence to analyze CCTV footage and automatically detect, identify, and track vehicles. AI algorithms can discern license plates, vehicle make and model, and unique features like dents or decals. By aggregating data from multiple cameras, authorities can map out a comprehensive route of a stolen vehicle's movements, significantly enhancing their ability to apprehend suspects and recover stolen vehicles.

Furthermore, AI-driven systems can analyze vehicle behaviour patterns and anomalies, aiding law enforcement in predicting and preventing future thefts. For instance, deviations from usual routes or suspicious behaviours such as frequent stops in isolated areas can trigger automatic alerts, enabling swift intervention by authorities.

Integrating smart technologies with existing surveillance infrastructure streamlines vehicle theft investigations reduces reliance on manual efforts, and enhances overall effectiveness. These advancements not only bolster law enforcement capabilities but also act as a deterrent to potential offenders, contributing to a safer and more secure society.

The structure of the paper is as follows. Section 2 describes the literature survey performed related to the problem statement. Section 3 describes the proposed system with block diagrams and their detailed description. Section 4 explains the methodology employed. Section 5 discusses the results in detail. Finally, the paper concludes in section 6 followed by section 7 being the references.

2. Literature Survey

K.V. Kadambari, et.al [2], put forth a system that uses a Deep Learning technique to extract number plates from the cars using CCTV footage captured in Arunachal Pradesh. The author employed Tesseract OCR to recognize the characters from the number plates. They effectively addressed error handling by adhering to the standard number plate template. In cases where a number is predicted instead of a character, the system intelligently predicts the character that is most likely to be present, and vice versa. However, a limitation of the proposed system is the inability to handle skewed number plates. Despite having multiple iterations the system struggles to get an accurate straightened image.

Xiying Li et.al [3], aim to quickly search, locate and track the target car using CCTV cameras. Licence plates cannot be identified in every case it could be due to occlusion, bad light or some glaze, hence author plans to focus on feature extraction on the cars captured in the video. To extract features and compare them with the candidate car, they employ the VGG16 model and use Triplet loss and joint multiple loss functions as loss functions. The limitation of VGG16 compared to EfficientNet is that it achieves higher accuracy with significantly fewer parameters, enabling more efficient model deployment on resource-constrained devices. EfficientNet has a balanced design which provides a superior trade-off between computational efficiency and accuracy.

Anil Sharma, et.al [4], adopted state-representation learning using a reinforcement learning-based policy to automate the camera selection process. In comparison with other methods, their strategy produces high-quality results as it employs learned state representations, which helps decrease the time spent on training the Reinforcement Learning policy. They also implemented a reward function into semi-supervised policy training. However, they did not provide the relation between the size of the network and the time required to find the next camera.

M. Seshaiah, et al. [5], proposed an image super-resolution technique to enhance the quality of CCTV surveillance footage. The authors implemented a method called Super Resolution Generative

Adversarial Network (SRGAN) to improve the resolution of images obtained from CCTV cameras. The system employs two networks - a generator and a discriminator - that are trained together using minimax logic. The algorithm utilizes Adam stochastic gradient descent and binary cross-entropy loss function for error calculation and backpropagation during the training process. While achieving an accuracy of 77.235%, the model's limitation lies in its inability to effectively enhance the textual portions of the images.

Youngmin Baek,et.al. [6], the research primarily aims to identify and extract text from images, regardless of the image's orientation, which can range from curved, skewed, or disorganized. The model is trained on the ICDAR dataset. For each text image the author uses region bounding boxes on each text individually comparing its affinity score with the centre of the character. It uses a weekly supervised method that generates pseudo-ground truths from an interim model. Training character-level models may involve vague or noisy guidance, relying on indirect clues for tasks like text generation or recognition.

Qi-Chao Mao et.al. [7], proposed an improved vehicle object detection method based on YOLOv3 to address the challenges of multi-scale vehicle detection and overlapping object detection in traffic surveillance videos. The authors enhanced the YOLOv3 architecture by implementing three key modifications, particularly excelling in detecting small and overlapping vehicles. By incorporating inverted residuals, SPP modules, and Soft-NMS, the system achieves higher detection accuracy and lower missed rates across various scales, making it well-suited for practical applications in diverse traffic scenarios.

Kirill Smelyakov et.al. [8], in this author has done comparative study between two most used OCR models, Tesser OCR and Easy OCR on various forms of texts such as electronic text document, internet resource, and banner. Their experimental analysis focused on both the accuracy and processing time of each OCR tool. Based on their findings, the researchers proposed an effective algorithm for OCR usage and provided specific recommendations for applying these tools to different types of data, categorized as non-distorted, slightly distorted, and highly distorted. The results had concluded Tesser OCR to perform better and faster over Easy OCR in regular book texts, but if natural scenes were taken the EasyOCR had better results.

Delong Cai (2023)[9], introduced the Corner Point Foreground Area Intersection over Union loss function as a new means of enhancing bounding box regression, especially for better localization of small objects. This new loss function overcomes problems associated with the overlap of boxes because it includes multiple objects, which either partially or completely have an overlapping rate higher than 70%, and therefore distinguishing between these two becomes problematic. This method was tested on datasets including vizDrone-DET2019 and SODA-D using both anchor YOLOv5 as well as the non-anchor youlo8. As such, the CFIoU loss function can be considered a major development in improving the bounding box regression accuracy principle, particularly for situations with small-scale object localisation and high overlap.

Shailendra Shende et al. (2023)[10], proposed a CNN-based approach towards missing object detection. This approach uses CNNs that detect missing objects embedded within images. During Image Classification, the model was trained using a dataset of 500 images. Even with the insufficient training dataset, the approach produced good results. Nevertheless, additional advanced training on more intricate images is required to better the model’s capacity and consistency in detecting missing objects under an array of cases.

The authors, 2024 [11,12,13,14] explores the potential of improving the identification of missing objects or persons using deep learning model and data augmentation. The paper performed a detailed survey on the existing systems with its limitations and discussed the techniques to improve the performance using computer vision and deep learning algorithms. Tabatabaei et al. (2024) [15] proposed a neural network model such as Google FaceNet, Ghost FaceNet and quantization methods to reduce heavy processing requirements for real-time facial recognition.

In the proposed solution, licence plate images are accurately aligned by identifying vertical lines within the image using the Hough Transform technique in conjunction with Canny edge detection and image binarization. This process ensures optimal alignment, thereby enhancing the output of optical character recognition (OCR) systems.

The system's functionality is further enhanced by employing vector calculations to determine the next camera selection. By leveraging known camera locations and viewing angles, as well as the positions of other cameras, the system effectively minimises the number of cameras required to locate the target vehicle. This strategic approach streamlines the surveillance process and optimizes resource utilization.

3. Proposed Model

In the context of remote areas, the detection of missing cars presents a unique challenge, particularly when relying completely on GPS technology. The conventional approach involves the car owner filing a missing report at the nearby police station, which subsequently deploys a search team to locate the vehicle. However, this method often relies on a brute-force approach, where authorities use the limited information provided by the car owner to conduct a widespread search.

In the proposed model, the system aims to enhance and optimize the existing process for locating missing vehicles in remote areas, addressing the limitations of the current methodology.

The proposed system relies on CCTV footage as the source to locate the car, where the system initiates the search for the missing car. The basic information for the car will be provided by the car owner with some reference images of the car. Utilizing CCTV footage the processing will be done to track the car’s location and produce a route map of the path it took.

Figure 1 shows the modular diagram of the proposed system that comprises five major components. The applicant of the complaint uses an online form to enter the necessary details to initiate the search process. The form includes fields like Car Model, Car colour, Licence plate number, an optional field to upload the image of the car, the approximate timestamp at which the car has gone missing and the approximate location of the car. This information is stored for further usage in the search process.

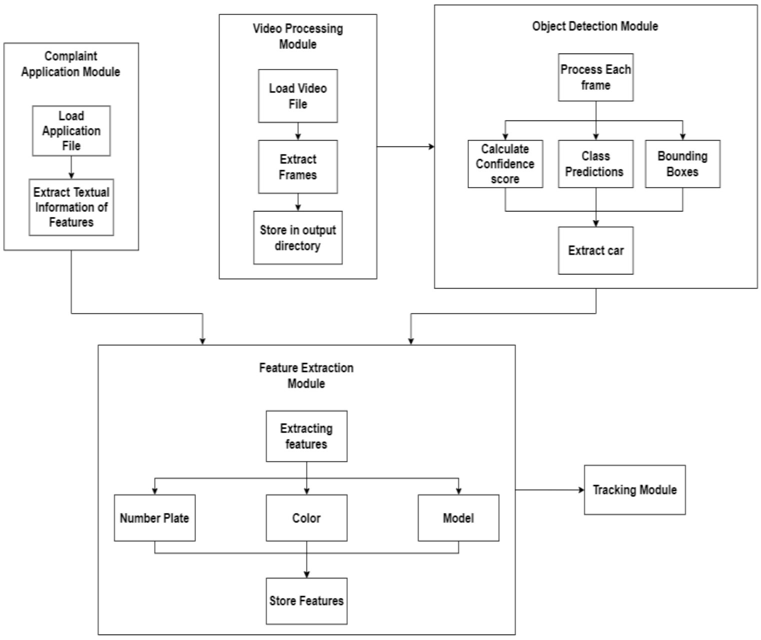

Figure 1: Modular diagram of the missing car detection model

Based on the location and timestamp of theft, the video processing module extracts the videos of appropriate intervals and locations from the CCTV systems, preprocesses them and converts them into frames for further analysis.

Object Detection Module:

The Object Detection Module uses the extracted frames to identify cars using the YOLOv7 pre-trained model [16,19]. The YOLOv7 internally calculates the confidence score, performs class prediction and assigns bounding boxes to identify all the candidate cars in the frame.

- **Confidence Score Calculation:** It calculates a confidence score for each object or potential car within the frame. This score reflects the module's level of certainty that a particular region of the image contains a car. Higher confidence scores indicate a stronger belief that an object is a car.
- **Class Prediction:** The module predicts the class of each object detected in the frame. In this specific case, the class of interest is "car." The module determines whether each object corresponds to a car or not.
- **Bounding Box Generation:** Once an object is identified as a car, the module outlines or highlights it by drawing a bounding box around it. This bounding box visually encloses the detected car within the frame, making it easily distinguishable.

By performing the tasks mentioned above, the Object Detection Module identifies all candidate cars within the video frame. These candidates are regions within the image where the module has detected cars based on its analysis. Further, the identified car images are extracted from the frame and sent for feature matching.

Feature Extraction Module:

The Feature Extraction Module takes the cropped car images and extracts the features of the car. These features include the car model, car colour and the number plate. The extracted features are used to match the car image with the subject car features as provided by the applicant These features are also used when the car has to be reidentified in subsequent videos. It identifies the target car from the set of candidate cars.

Tracking Module:

Figure 2 depicts the tracking module of the system, which serves the purpose of tracking a car within the same video (inter-frame tracking). The target car features from the previous frame are then used for the identification of the same car in the next frame. It goes through the following steps.

- **Initial Identification:** The module starts by identifying and tracking objects in one video frame. This initial identification may involve detecting cars or other objects in the frame and assigning unique identifiers or labels to them.
- **Foreground Detection:** It may perform foreground detection to isolate moving objects from the static background. This step helps in distinguishing the objects of interest (e.g., cars) from the background.
- **Connected Flow Analysis:** The module analyses the flow of connected objects across frames. It tracks how objects move from one frame to the next, allowing it to follow the trajectory of each object.
- **Tracking by Optical Flow:** Optical flow is a technique used to estimate the motion of objects in consecutive video frames. It calculates how pixels move between frames, helping to track the objects' positions accurately over time.
- **Detection of the direction of motion**: The direction of motion of the car helps in identifying the next camera or the set of cameras that might have captured the car. The direction of motion is calculated using vector mathematics.

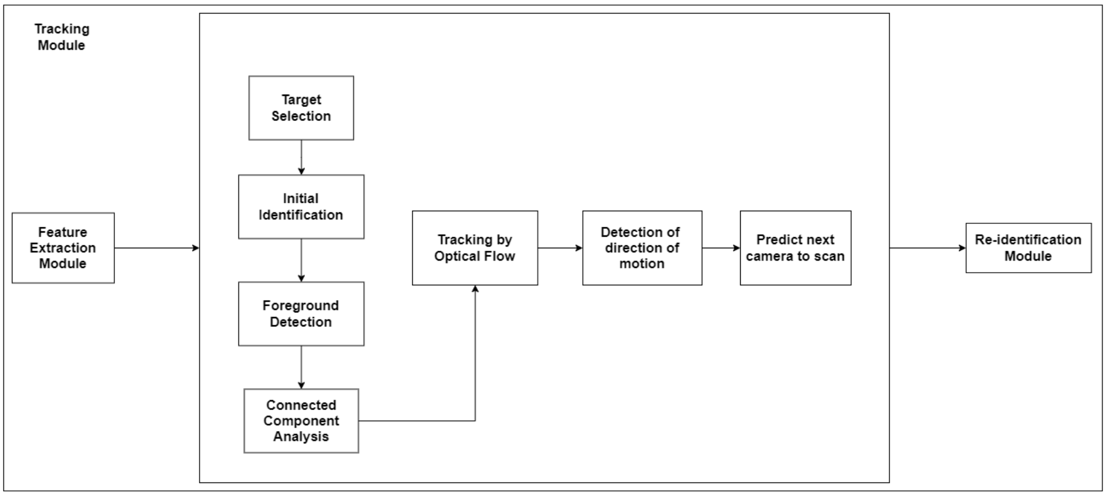

Figure 2: Tracking Module

Re-Identification Module:

Figure 3 shows the re-identification module, which serves the purpose of tracking cars from multiple CCTV footage videos captured at different locations (inter-video tracking). When a car exits one camera coverage it has to be identified in another camera coverage. The identification module calculates similarity scores between the cars identified in the last frame of the previous video and the cars in the first frame of the new video. This score is based on the features of the car extracted from the feature extraction module. Reidentification is the process of associating an object identified in one video source with the same object in another video source. It helps in maintaining tracking across video boundaries.

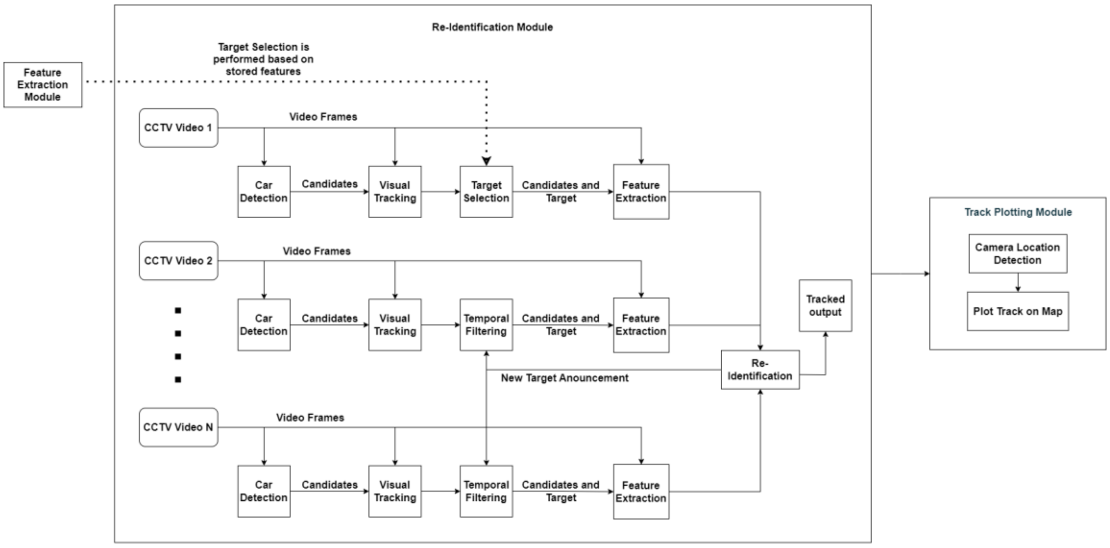

Figure 3: Re-Identification Module

Track Plotting Module:

This module detects the location of the cameras in which the car has been found and then plots the track on a map.

4. Methodology

The methodology of the project is as follows:

4.1 Video Processing Module

In the video processing module video from CCTV is uploaded to the system and then video processing is done, where the frames are converted from the video and stored in one folder.

4.2 Object Detection Module

Each frame undergoes analysis in the object detection module, if any particular frame does not contain objects(cars) then delete the frame. This helps to reduce the complexity in cases of no vehicles in videos. The module uses the YOLO model for vehicle detection. Then after cleaning through the frame data the system starts detecting cars in each frame and creates a bounding box that crops the cars from the frame. The cropped cars are stored in another folder. Additionally, the cropped cars are checked with a number plate in the frame, if a number plate is detected then these number plates are also cropped and stored.

4.3 Feature Extraction Module

The cropped cars and number plates are sent to extract the features of the car such as car model, colour, texture and number plate. This feature extraction is done with the user input image also. Car models and colours are extracted using the CNN model EfficientNet B1. The cropped number plates are passed under OCR.

4.3.1 EfficientNet

EfficientNet is a convolution neural network that works on the concept of compound scaling which is based on the idea of balancing dimensions of width, depth, and resolution by scaling with a constant ratio [17]. This capability allows the system to train the model efficiently, minimising any additional loss during the training process. Our dataset is employed for training models specifically designed for car colour and model classification. The training process involves instructing the model to classify cars into distinct colour classes and model classes.

4.3.2 Car Model Detection

The detection of the model of the subject car is performed using a custom-trained EfficientNet-B1 model. The dataset used for training purposes comprises two popular car models in India, the Maruti Suzuki WagonR and the Maruti Suzuki Swift. The dataset consists of 4000 images and 2000 images per class. To perform optimal training of the model, the dataset is split into train and test sets consisting of 1600 and 400 images respectively, for each of the two classes in the dataset.

4.3.3 Car Color Detection

The color of the subject cars in the CCTV footage is identified by another EfficientNet-B1 model trained on a dataset of 5280 car images, classified into 8 color categories. The color classes present in the dataset are black, blue, brown, green, grey, red, white and yellow. The dataset is split into train and test sets containing 560 and 100 images per class respectively.

4.3.4 Easy OCR

Easy OCR is a font-dependent printed character reader based on a template-matching algorithm.

4.3.5 De-skewing the License plate

In many cases, it was observed that the Easy OCR gave inaccurate results because the licence plate was captured at an angle. To handle such cases, the angle of licence plate boundaries with the horizontal was predicted and then the image was de-skewed by rotating it.

4.4 Tracking Module

Figure 4 is considered to be one of the most important modules as it produces the tracking results that help the officials find the stolen vehicle. The tracking is split into three parts, the car direction detection, the next camera decision-making, and the track plotting.

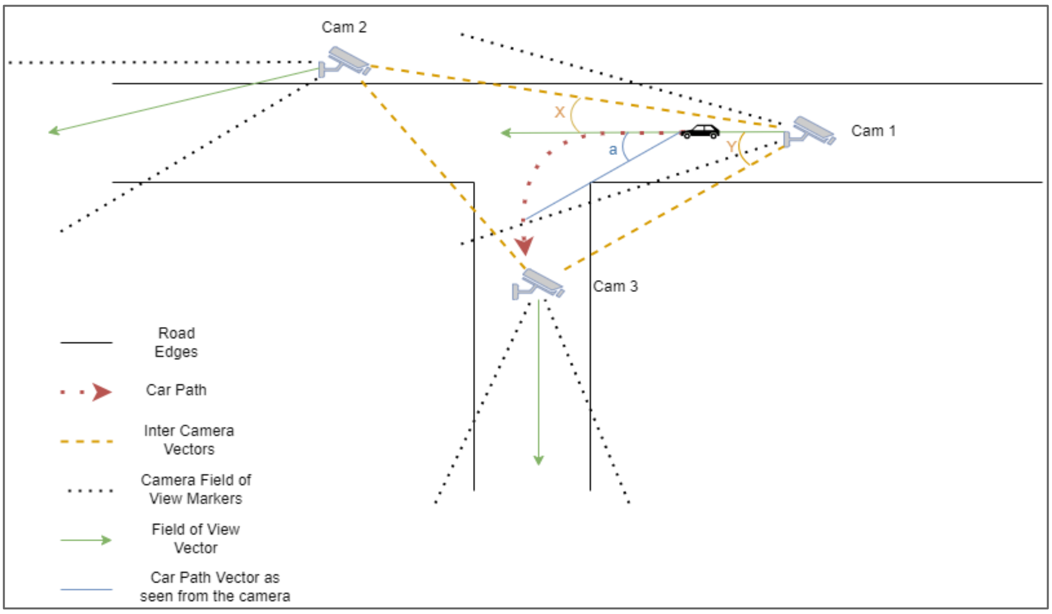

Figure 4: Choosing the next camera to scan

4.4.1 Car Direction Detection

The process starts by scanning through the CCTV footage of the nearest camera from the location of the theft. Once the suspect vehicle is identified in the footage, it is tracked using an object tracking algorithm, 'DeepSort', to produce a track of the path travelled by the vehicle as seen by the camera.

DeepSORT [18] is an advanced algorithm for multi-object tracking. It integrates deep neural networks for object detection and tracking, enhancing traditional SORT (Simple Online and real-time tracking) by associating detected objects over consecutive frames using learned features. DeepSORT excels in crowded scenes and occlusions, vital for surveillance and autonomous systems. The DeepSORT produces a path for every car detected in the video. Using the extracted features in earlier stages, we identify the subject car and extract the tracked path coordinates. These coordinates are used to determine the direction the car moves in the video.

The proposed system uses the first and last coordinates of the centroid of the car bounding box and creates a vector. This vector is compared with the direction vector of the field-of-view of the camera to decide whether the car takes a turn in a specific direction or moves straight.

4.4.2 Next Camera to Choose

Direction vectors between cameras (inter-camera vectors) are created beforehand using the latitude and longitude of each camera. For each inter-camera vector between the current camera and the neighbouring cameras, its angle with the direction vector of the current camera is found. These angles are compared with the angle between the car direction vector and the field-of-view vector of the camera. Next camera is chosen based on the least difference between these two angles in the appropriate direction.

Let us consider the following parameters:

C_i: Position vector of the i-th camera.

F_i: Field of view (FOV) vector of the i-th camera.

$I_{ij} = C_j - C_i$: The inter-camera vector between camera i and camera j.

P_{start} : The initial position of the car.

P_{end}: The final position of the car.

$V_{car} = P_{end} - P_{start}$: Car Movement Vector.

$\alpha = \cos^{-1}\left(\frac{\mathbf{F}_i \cdot \mathbf{V}_{\text{car}}}{\|\mathbf{F}_i\| \|\mathbf{V}_{\text{car}}\|}\right)$: The angle between FoV and car movement vector.

$\beta_{ij} = \cos^{-1}\left(\frac{\mathbf{F}_i \cdot \mathbf{I}_{ij}}{|\mathbf{F}_i\| \|\mathbf{I}_{ij}\|}\right)$: The angle between FoV and inter camera vector for each camera j.

j^*: next camera to scan through.

Therefore the next camera is chosen using the equation:

$j^* = \arg\min(|\alpha - \beta_{ij}|)$

Subject to:

$\mathbf{V}_{car} \cdot I_{ij} > 0$

4.4.3 Track plotting

The above process is applied to subsequent cameras to track the car as far as possible in the CCTV network. Finally, the track is created using the geolocations of each camera that tracked the subject car to produce a complete route map of the path taken by the car. This route is plotted using map APIs like Google Maps.

5. Results and Discussion

The intermediary results of the proposed system are mentioned below in figures (5,6,7,9,11,12).

Figure 5 shows the frame extracted from the CCTV footage and the cropped image of the detected car.

Figure 5: Video Frame and Cropped image of the detected car from the input frame

Figure 6 shows the cropped number plate image from the originally cropped car image.

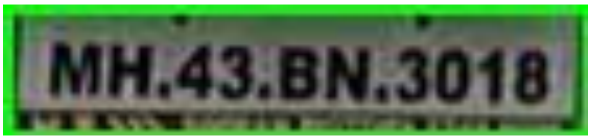

Figure 6: Cropped number plate image from cropped car image

Figure 7 shows the results of OCR performed on the cropped number plate image. The results of OCR show significant improvement after deskewing the image. The skewed images had shown the nature of not detecting the number plate correctly, as in this example the skewed image had output as ‘LcDT’ which is meaningless. Instead just by deskewing the same image, we are able to improve the accuracy up to 95%.

Figure 7: OCR Result on cropped number plate image (before and after deskewing)

Figure 8 shows the training and validation results of the EfficientNet-B1 model trained on a dataset containing 4000 images for 2 cars (Swift and WagnoR). The graphs display training and validation accuracies of above 0.9, with a low training and validation loss of less than 0.2. This ensures good model training for predicting car models.

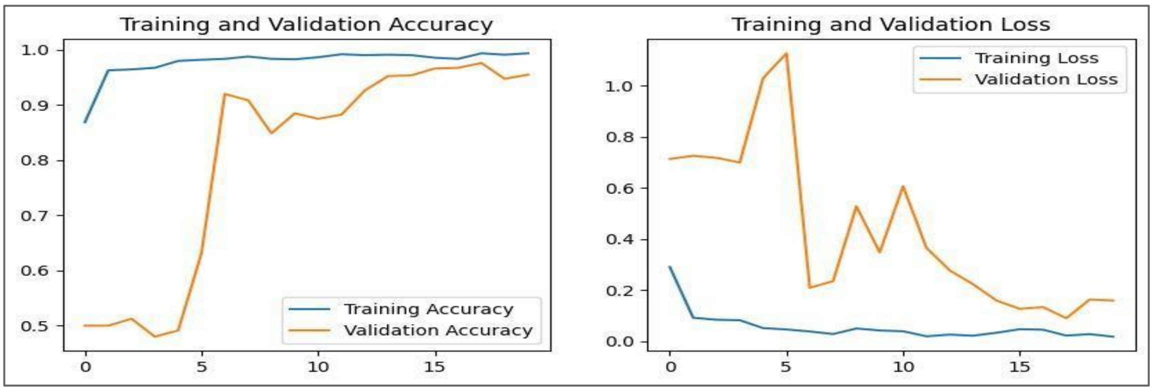

Figure 8: Training and validation results for car model prediction model.

Figure 9 showcases the prediction of the car model given by our model on unseen images.

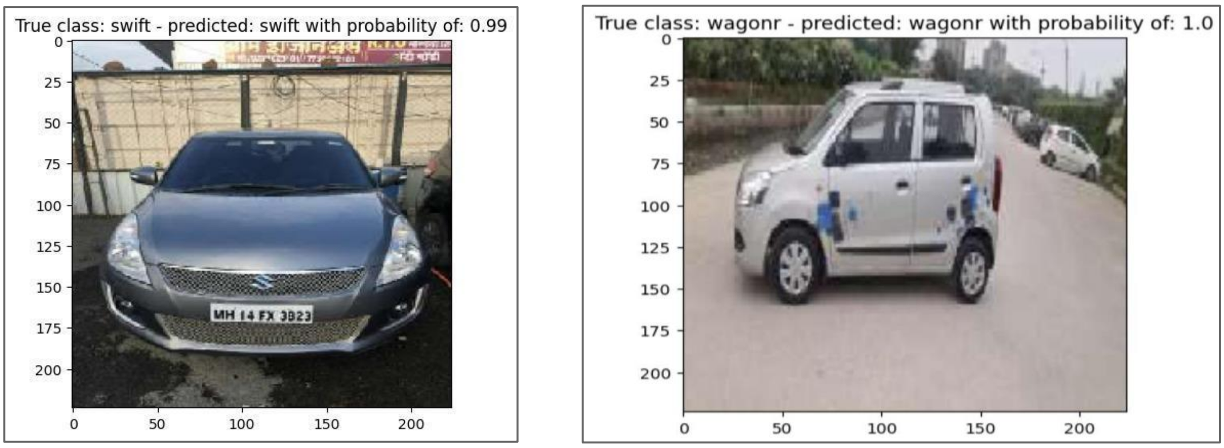

Figure 9: Results of car model recognition

Figure 10 shows the training and validation results of the EfficientNet-B1 model trained on a dataset containing 5280 images for 8 cars' color classes (black, blue, brown, grey, green, red, white, yellow). The graphs display training and validation accuracies of almost 0.9, with a low training and validation loss of less than 0.5. This ensures good model training for predicting car models.

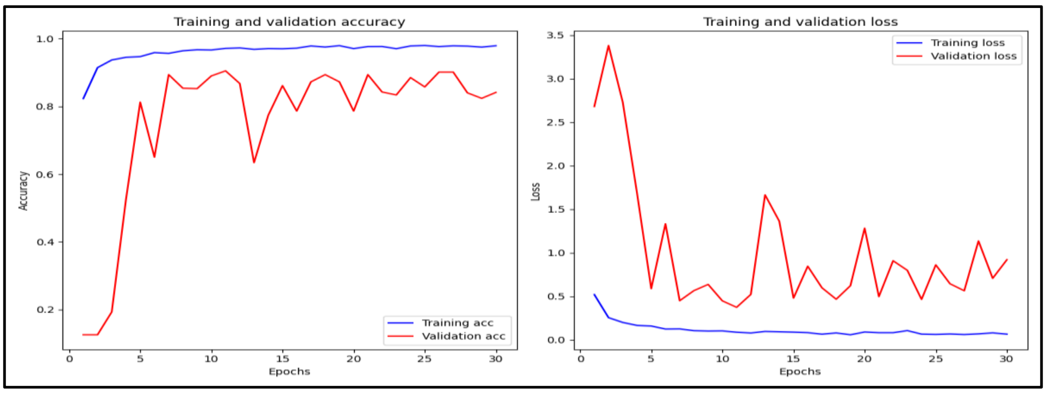

Figure 10: Training and validation results for car color recognition model.

Figure 11 showcases the prediction of the car color given by our model on new images.

Figure 11: Results of car color recognition

Figure 12 shows the application of the DeepSORT algorithm on the video extracted from CCTV footage. The result shows the path, as tracked by the algorithm, taken by the car under observation as seen in the video.

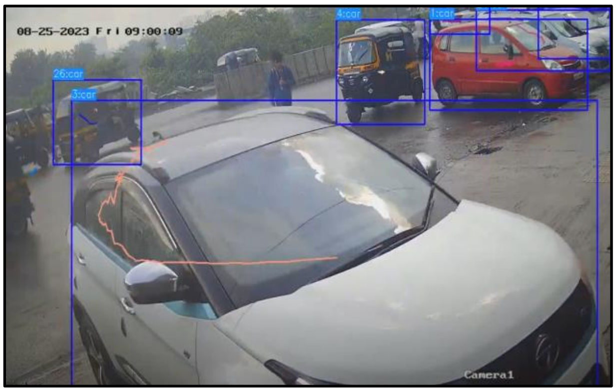

Figure 12: Tracking results of DeepSORT

Figure 13 shows a map of our college campus, illustrating two different paths: the red line represents the predicted path by the system, while the blue line indicates the actual path taken by the car. The difference between the two paths is minimal, even though the car was not detected by one of the cameras on campus. The root mean square error (RMSE) was 0.000171, which concludes that the difference between the actual and predicted path is not too high.

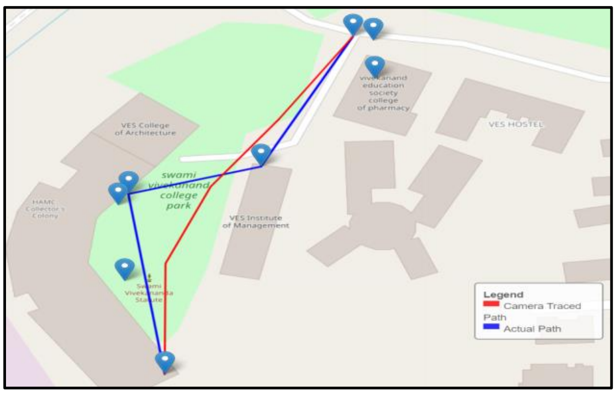

Figure 13: Comparison of camera traced path and actual path

Table 1 shows the execution time comparison of the complete pipeline using CPU and GPU, tested on 3 videos of varying file size and duration.

Table 1: CPU vs GPU execution time comparison

Execution Time Comparison Between CPU and GPU					
Sr No	Video Length(sec)	Video Size(MB)	CPU Time(sec)	GPU Time(sec)	Improvement %
1	4.66	5.8	125.25	21.25	489.41
2	14.32	29.9	160.76	25.49	530.68
3	23	14.5	1010.9	77	1212.86

6. Conclusion

The ongoing challenge of vehicle theft necessitates innovative solutions, especially as criminals adapt to evolving technologies. The fusion of artificial intelligence and computer vision with established surveillance systems offers a promising strategy to combat this crime effectively. By automating the analysis of CCTV footage, law enforcement agencies can swiftly detect, identify, and track stolen vehicles, significantly boosting their capacity to apprehend suspects and recover stolen property. Furthermore, AI-powered systems allow for proactive measures by identifying behavioural patterns and anomalies in vehicle activity, aiding in the prediction and prevention of future thefts. In remote regions where conventional methods may fall short, utilizing CCTV footage as a primary tool for locating missing vehicles presents a practical alternative, streamlining the search process and overcoming the limitations of GPS technology. This technology also aids the crime branch to monitor a particular vehicle for intelligence purposes, additionally traffic police can also detect the speeds of the vehicles without a person to be physically present to check speed. During disaster management, this system can be used for monitoring the evacuation routes. Ultimately, these advancements augment law enforcement capabilities and act as a deterrent to potential offenders, contributing to a safer and more secure society.

References

[1] https://auto.economictimes.indiatimes.com/news/aftermarket/vehicle-theft-increases-by-2-5x-in-india-in-2023-delhi-tops-the-list-acko-report/108454773?utm_source=copy&utm_medium=pshare

[2] K.V. Kadambari, et.al. "Deep Learning Based Traffic Surveillance System for Missing and Suspicious Car Detection", 2020

[3] Xiying Li et.al. "Object Re-Identification Based on Deep Learning",2019

[4] Anil Sharma, et.al. "Intelligent Camera Selection Decisions for Target Tracking in a Camera Network", Indraprastha Institute of Information Technology (IIIT), Delhi,2022

[5] Prof. M. Seshaiah, Abhijith R Nair, Ektha Mallya, Shiv Dev et.al. "CCTV Surveillance Camera's Image Resolution Enhancement using SRGAN",VTU, Belgaum, India,2020

[6] Youngmin Baek, Bado Lee, Dongyoon Han, Sangdoo Yun, and Hwalsuk Lee, et.al., "Character Region Awareness for Text Detection", 2019

[7] Qi-Chao Mao et.al. "Finding every car: a traffic surveillance multi-scale vehicle object detection method", College of Computer Science and Engineering, Shandong University of Science and Technology, 2020

[8] Kirill Smelyakov et.al. "Effectiveness of Modern Text Recognition Solutions and Tools for Common Data Sources",Kharkiv National University of Radio Electronics,2021

[9] Delong Cai et.al. "Corner-Point and Foreground-Area IoU Loss: Better Localization of Small Objects in Bounding Box Regression", School of Electrical Engineering and Intelligentization, DongGuan University of Technology, 2023
[10] Shailendra Shende et.al. "CNN based Missing Object Detection", Government College of Engineering Chandrapur, 2023
[11] H, Dr & J, Archana & N, Gagana. (2024). A Survey of Integrating Deep Learning-Based Missing Person Detection Model into CCTV Systems for Enhanced Identification. International Journal of Advanced Research in Science, Communication and Technology. 259-268. 10.48175/IJARSCT-15336
[12] Abdullahi, Mohammed & Oyelade, Olaide & Kana, Armand & Bagiwa, Mustapha & Abdullahi, Fatimah & Junaidu, Sahalu & Iliyasu, Ibrahim & Ore-ofe, Ajayi & Chiroma, Haruna. (2024). A systematic literature review of visual feature learning: deep learning techniques, applications, challenges and future directions. Multimedia Tools and Applications. 1-58. 10.1007/s11042-024-19823-3
[13] Abdulkareem, Ismael & Al-Shammri, Faris & A. Khalid, Noor & Omran, Natiq. (2024). A Proposed Approach for Object Detection and Recognition by Deep Learning Models Using Data Augmentation. 10.3991/ijoe.v20i05.47171.
[14] Wicaksana, I Putu & Davinsi, Gabriel & Afriyanto, Muhammad & Wibowo, Antoni & Suri, Puti. (2024). Systematic Literature Review: The Influence and Effectiveness of Deep Learning in Image Processing for Emotion Recognition. 10.21203/rs.3.rs-3856084/v1.
[15] Tabatabaei, Seyed Mohammad Taha. (2024). Efficient Neural Network Face Recognition System on a Resource Constrained Device. 10.13140/RG.2.2.34987.66085.
[16] Dkengne Sielenou, Pascal Alain & Girard, Stéphane. (2023). Detection of Traffic Scene Objects using YOLO Algorithm: Theory and Practical Guide. 10.13140/RG.2.2.19162.52163.
[17] Nicolai Wojke et.al. “Simple online and real time tracking with a deep association metric”, University of Koblenz-Landau, 2017
[18] Mingxing Tan et.al. “EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks”, 2020.
[19] Chien-Yao Wang et.al. “YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors”, Institute of Information Science, 2022
[20] Pier Matteo Barone,et.al. "Materials for the study of the locus operandi in the search for missing persons in Italy",American University of Rome, 2021