Reliable Vehicle Detection and Tracking on Highways using Geo-Heinz Corner Response, DBSCAN, and CSRT

Mustafa Sadeghi[1], **Mohamad Mehdi Khorsand Aalam**[1], and **Hadi Sadoghi Yazdi**[*1,2]

[1]Department of Computer Engineering, Ferdowsi University of Mashhad, Mashhad, Iran
[2]Center of Excellence in Soft Computing and Intelligent Information Processing, Ferdowsi University of Mashhad, Mashhad, Iran

Abstract

We propose a reliable vehicle detection and tracking framework designed for highway camera surveillance, integrating multi-scale *Geo-Heinz Mean* corner response, DBSCAN clustering, and CSRT-based tracking. The method begins with a Gaussian Mixture-based Background Subtractor (MOG2) to detect moving objects, followed by morphological operations to refine the foreground mask.

A novel multi-scale *Geo-Heinz Mean* corner response algorithm is then applied to extract prominent corners by computing gradients, constructing a structure tensor, and utilizing a Heinz mean-based quality criterion for precise corner detection. These detected corners undergo motion analysis and foreground masking before being grouped into potential vehicle regions using density-based spatial clustering (DBSCAN).

To enhance tracking stability, a temporal smoothing mechanism minimizes bounding box fluctuations across consecutive frames, effectively reducing motion jitter and transient noise. CSRT trackers are then initialized on the refined candidate regions, with reinitialization guided by an Intersection-over-Union (IoU) criterion to maintain consistent vehicle identities despite occlusions or rapid motion.

Extensive experiments on highway traffic video datasets validate the high detection accuracy and robust tracking performance of the proposed approach under real-world conditions. The modular framework ensures real-time operation and provides a scalable foundation for future advancements, such as adaptive parameter tuning and deep learning-based classification. This method is particularly well-suited for applications in traffic monitoring, intelligent transportation systems, and advanced surveillance networks.

Keywords: Vehicle Detection, Vehicle Tracking, Highway Surveillance, Geo-Heinz Mean Corner Response, Corner Detection, DBSCAN Clustering, CSRT Tracking, Real-Time Tracking, Intelligent Transportation Systems

[*]Corresponding author: h-sadoghi@um.ac.ir

1 Introduction

Highway camera surveillance is a crucial component of modern traffic monitoring and intelligent transportation systems. As vehicle traffic continues to grow and road networks become increasingly complex, the need for a reliable and real-time vehicle detection and tracking system becomes more pressing. This paper presents a robust and efficient highway surveillance framework, integrating multi-scale Geo-Heinz Mean corner response, DBSCAN clustering, and CSRT-based tracking to achieve high detection accuracy and stable tracking performance in real-world conditions.

The proposed method begins by detecting moving objects using a Gaussian Mixture-based Background Subtractor (MOG2) and refining the foreground mask with morphological operations. Next, a novel multi-scale Geo-Heinz Mean corner response algorithm is applied to detect prominent corners by computing image gradients, constructing a structure tensor, and applying a Heinz mean-based quality criterion. These detected corners undergo motion analysis and foreground masking before being grouped into potential vehicle regions using DBSCAN clustering. To enhance tracking stability, a temporal smoothing mechanism minimizes bounding box fluctuations across consecutive frames, thereby reducing motion jitter and transient noise. Finally, CSRT trackers are deployed in optimized candidate regions, with reinitialization guided by an Intersection-over-Union (IoU) criterion, ensuring that vehicle identities remain consistent even during occlusion or rapid motion.

In this paper, we thoroughly describe the proposed framework, detailing each step—from foreground segmentation to feature extraction and tracking. We then present experimental evaluations on highway traffic video datasets, demonstrating the system's high detection accuracy and reliable tracking performance. A discussion follows, exploring potential improvements, future research directions, and real-world applications for highway surveillance. The paper concludes with a summary of key findings and insights into how this work contributes to the advancement of intelligent transportation systems and automated traffic monitoring.

1.1 Motivation

The motivation behind this framework stems from the unique challenges of highway surveillance. Highways involve high-speed vehicles, rapidly changing environmental conditions, and frequent occlusions, all of which pose significant difficulties for vehicle detection and tracking. These factors can lead to missed detections, unstable tracking, and inconsistencies in vehicle identity over time.

To overcome these challenges, our approach utilizes multi-scale analysis, allowing it to capture both fine and coarse features, ensuring robust corner detection across various vehicle sizes and distances. Additionally, unlike data-intensive deep learning models that require extensive labeled datasets, our method operates effectively without the need for large-scale training data, making it a practical and scalable alternative. By integrating temporal smoothing and CSRT-based tracking, the system further reduces motion jitter and ensures continuous and stable tracking, making it highly suitable for real-time highway surveillance applications.

Overall, our method provides a scalable, efficient, and practical solution for vehicle detection and tracking in highway environments. It lays a strong foundation for future advancements in traffic monitoring, intelligent transportation systems, and automated surveillance networks.

2 Related Works

Li et al. (2009) [1] proposed a vehicle type recognition method based on Harris corner detection, leveraging feature point extraction for classification. Their approach utilized the Hausdorff distance to compare vehicle shapes, achieving high accuracy while eliminating the need for complex training models. However, their study highlighted challenges related to noise sensitivity and scale variation, which remain areas for further refinement in vehicle recognition systems.

Anandhalli and Baligar [2] proposed a corner point-based vehicle detection method using the Harris corner detector and Lucas-Kanade tracking, demonstrating robustness across varying climatic conditions. Tsai et al. explored color-based classification with Bayesian models, while Lin et al. utilized road segmentation for airborne vehicle detection. However, traditional methods often struggle with false positives in complex environments. Feature-based tracking has proven effective in dynamic scenes, with potential improvements through deep learning integration.

Munajat et al. [3] introduced a vehicle detection and tracking method combining Kanade-Lucas-Tomasi (KLT) corner detection with line-adjacent features, improving accuracy in dynamic environments. Beymer et al. utilized Gaussian Mixture Models (GMM) for real-time traffic monitoring, while Kamijo et al. proposed a Spatio-Temporal Markov Random Field (ST-MRF) model for movement analysis. Feature-based tracking has shown promise in overcoming occlusions and lighting variations, with future improvements expected through deep learning integration.

Kiran Kumar et al. [4] proposed a vision-based vehicle speed detection method using frame subtraction, edge detection, and Harris corner detection, achieving high accuracy in speed estimation. Rad et al. introduced the Correlation Velocity Sensor (CVS) technique, while Zheng et al. explored gray-level corner detection for motion tracking. These studies demonstrate the effectiveness of computer vision in transportation monitoring, with future advancements expected through deep learning integration.

Yuan et al. [5] proposed an adaptive Harris corner detection algorithm that improves vehicle recognition by dynamically adjusting thresholds and incorporating Hausdorff distance-based matching. While deep learning methods offer higher accuracy, they require extensive datasets and computational power. Adaptive corner detection provides an efficient alternative, highlighting its potential for robust vehicle identification in intelligent transportation systems.

3 Perposed Method

This section outlines the key components of our vehicle detection and tracking system, designed for highway surveillance. It uses traditional computer vision methods to achieve real-time performance without needing large amounts of training data.

The process begins with detecting moving objects through a Gaussian Mixture-based Background Subtractor (MOG2), which helps separate vehicles from the static background. We then refine this detection using morphological operations to remove noise and enhance the visibility of moving vehicles. This refined mask ensures that only relevant moving areas are processed further.

Next, we use a multi-scale Geo-Heinz Mean corner response algorithm to detect strong corner points within the moving regions. The algorithm targets the moving areas specifically, identifying corners that exhibit significant movement above a set threshold. This step filters out less relevant features, ensuring that only features tied to actual vehicle motion are processed. By calculating image gradients and constructing a structure tensor, we apply a hybrid geometric and Heinz mean quality measure to accurately identify important corners, capturing both fine details and larger structures.

Once the corners are extracted, we analyze motion and apply further masking to eliminate false detections. The remaining reliable features are grouped into potential vehicle regions using DBSCAN clustering. To improve tracking stability, we apply temporal smoothing to reduce jitter and fluctuations in bounding boxes between frames. CSRT trackers are then initialized on these regions, and we reinitialize trackers based on an Intersection-over-Union (IoU) threshold to maintain consistent vehicle identities, even during occlusions or fast movements.

This streamlined approach offers a practical and efficient solution for real-time highway surveillance, ensuring accurate detection and robust tracking performance for reliable traffic monitoring.

3.1 Foreground Segmentation and Mask Refinement

Our vehicle detection pipeline begins by separating moving objects from the static background using a Gaussian Mixture-based Background Subtractor (MOG2). This method models each pixel as a mixture of Gaussian distributions, allowing it to effectively differentiate between background and moving objects in dynamic environments. In highway surveillance scenarios, where vehicles move at high speeds and lighting conditions vary, MOG2 efficiently highlights regions corresponding to moving vehicles, making it well-suited for real-time detection.

After background subtraction, the resulting foreground mask may contain noise and small artifacts due to environmental variations or sensor limitations. To address this, we apply morphological operations, specifically a combination of opening and closing, to refine the mask. The opening operation removes small isolated noise points by first eroding the mask, shrinking objects to eliminate minor noise, and then restoring its structure through dilation, expanding objects back to their proper shape. Conversely, the closing operation fills in small gaps within detected vehicle regions, ensuring that moving objects appear as continuous and well-defined areas.

By applying these refinements, the foreground mask becomes clearer, improving the accuracy of feature extraction and the formation of candidate vehicle regions in the later stages of the detection pipeline. This preprocessing step ensures that only meaningful features are passed forward, enhancing the robustness of the overall detection and tracking system in highway surveillance applications.

3.2 Multi-Scale Geo-Heinz Mean Corner Response

This subsection introduces the Multi-Scale Geo-Heinz Mean Corner Response, an adaptive corner detection method derived from the structure tensor, which plays a key role in our vehicle detection framework. At each pixel x, a 2×2 structure tensor S is computed, producing eigenvalues λ_1 and λ_2 that represent local intensity variations. Unlike traditional methods that rely on a fixed determinant-trace formulation, our approach employs an adaptive measure based on the Heinz mean, providing greater flexibility and robustness.

The Heinz mean $H_\nu(\lambda_1, \lambda_2)$ is defined as:

$$H_\nu(\lambda_1, \lambda_2) = \frac{1}{2}\left(\lambda_1^{1-\nu}\lambda_2^{\nu} + \lambda_1^{\nu}\lambda_2^{1-\nu}\right), \quad 0 \le \nu \le 1.$$

Here, the tunable parameter ν allows for interpolation between minimum, geometric, and arithmetic means, enabling the response to adapt to different types of corner structures.

Based on this formulation, two variants of the Geo-Heinz Mean operator are derived:

$$Q_H(\text{geom})(x) = \frac{H_\nu(\lambda_1, \lambda_2)}{\lambda_1 \, \lambda_2},$$

$$Q_H(\text{arith})(x) = \frac{H_\nu(\lambda_1, \lambda_2)}{\frac{1}{2}(\lambda_1 + \lambda_2)}.$$

Both variants generate normalized scores within the range $[0, 1]$. Higher scores (close to 1) indicate strong, well-balanced eigenvalues, characteristic of distinct corner-like structures such as vehicle edges, headlights, and key contours. Lower scores (close to 0) correspond to flat or unstructured regions, where strong corner features are absent.

For our method, we utilize the geometric mean variant $Q_H(\text{geom})$, which enforces a stricter measure of isotropy for detecting corners. Importantly, we compute the Geo-Heinz Mean Corner Response exclusively in moving regions identified through the refined foreground mask. This approach ensures that only vehicle-related features are processed, thereby reducing computational overhead and minimizing false detections from static background elements.

To enhance robustness across various vehicle sizes and distances, the Geo-Heinz Mean operator is applied at multiple scales. This multi-scale analysis enables the system to capture both fine details and larger vehicle structures, ensuring consistent detection across different traffic conditions. Additionally, the operator's intrinsic smoothing properties help to reduce sensor noise, mitigate low-contrast regions, and compensate for uneven illumination common challenges in real-world highway surveillance.

By integrating the geometric mean variant of the Geo-Heinz Mean Corner Response—computed only on moving regions into our detection pipeline, we achieve a highly reliable and noise-resistant feature extraction mechanism. This is a crucial step for subsequent clustering and tracking, ultimately improving the accuracy and stability of vehicle detection in highway surveillance applications.

3.3 Feature Filtering and Motion Analysis

After detecting robust corner points using the multi-scale Geo-Heinz Mean corner response algorithm, the next step is to filter and validate these features through motion analysis. The primary goal of this stage is to ensure that only meaningful features corresponding to vehicle movement are retained, while eliminating irrelevant or false detections caused by noise or static background elements.

We begin by leveraging the refined foreground mask, which highlights the moving regions in the scene. Only corners that lie within these moving regions are kept, as the background or stationary objects do not contribute to vehicle detection. Afterward, we analyze the movement of the detected corners across consecutive frames. A corner movement threshold is applied to filter out corners that do not exhibit sufficient displacement between frames, as these are likely noise or background features rather than meaningful vehicle features.

Corners with displacement exceeding the predefined threshold are considered significant and are retained for further processing. This threshold ensures that only corners associated with actual vehicle motion are tracked, improving the efficiency and accuracy of the detection process.

Once the reliable corners are filtered based on their movement, additional foreground masking is applied to further refine the feature set and remove any remaining false positives. The resulting validated features are then ready to be grouped into potential vehicle regions in the subsequent steps, such as DBSCAN clustering. This feature filtering and motion analysis process ensures that only the most relevant and stable features are used for vehicle detection and tracking, thus enhancing the robustness of the overall system.

3.4 Candidate Region Formation via DBSCAN Clustering

The next step in our vehicle detection pipeline, after filtering and validating the detected features through motion analysis, is to group the reliable features into potential vehicle regions. We do this using DBSCAN (Density-Based Spatial Clustering of Applications with Noise), a clustering algorithm that excels at finding dense spatial areas and ignoring isolated outliers.

DBSCAN works by grouping points that are close to each other based on a distance threshold, while treating points that don't fit into these groups as noise. In the vehicle detection process, the filtered reliable corners are input into DBSCAN. These corners are then clustered into regions that likely correspond to individual vehicles.

There are two main parameters in DBSCAN: epsilon (ε), which defines the maximum distance for two points to be part of the same cluster, and min_samples, which specifies the minimum number of points needed to form a dense cluster. By adjusting these parameters, we can control the sensitivity of the clustering, ensuring that only meaningful clusters, which represent actual vehicles, are formed.

Once DBSCAN groups the valid features into clusters, these clusters become potential vehicle regions. These regions are then processed in later stages, like temporal smoothing and tracking, to confirm the presence of vehicles and maintain their identity over time.

In short, DBSCAN clustering is key to transforming individual feature points into meaningful vehicle regions. It groups nearby features into candidate regions, helping with the next steps of vehicle tracking and identification.

3.5 Temporal Smoothing and CSRT-Based Tracking

After identifying potential vehicle regions using DBSCAN clustering, the next priority is to stabilize vehicle detection and tracking over time. To achieve this, we use temporal smoothing and CSRT-based tracking to ensure consistent vehicle identities and reduce fluctuations in tracking, even in challenging situations like occlusions or fast movement.

Temporal smoothing is applied to minimize fluctuations in the bounding boxes between consecutive frames. When tracking vehicles, particularly in fast-moving highway scenes, slight jittering in the bounding boxes can happen due to noise, sensor issues, or small tracking errors. To counter this, temporal smoothing averages the position of the bounding boxes across several frames, smoothing out minor movements and reducing the chances of false positives caused by these small fluctuations.

Once temporal smoothing is applied, we move to CSRT (Channel and Spatial Reliability Tracker) for tracking the candidate regions. CSRT is a robust and efficient tracking algorithm that combines channel and spatial reliability to track objects accurately over time. It's especially useful for handling occlusions, rapid motion, and changing object appearances in dynamic environments. CSRT trackers are initialized on the vehicle regions, and they adapt to changes in the vehicle's position, size, and appearance across frames.

To ensure accurate tracking, the CSRT tracker is reinitialized when necessary. This is done based on an Intersection-over-Union (IoU) criterion, which measures the overlap between the predicted bounding box and the actual region in the current frame. If the overlap falls below a certain threshold—indicating the tracker may have lost the vehicle—a reinitialization is triggered to re-establish correct tracking.

In conclusion, the combination of temporal smoothing and CSRT-based tracking is a reliable method for maintaining vehicle identities across frames, even in the face of occlusions, rapid movement, or tough environmental conditions. It ensures continuous, stable tracking of vehicles throughout the entire video sequence, which is critical for highway surveillance.

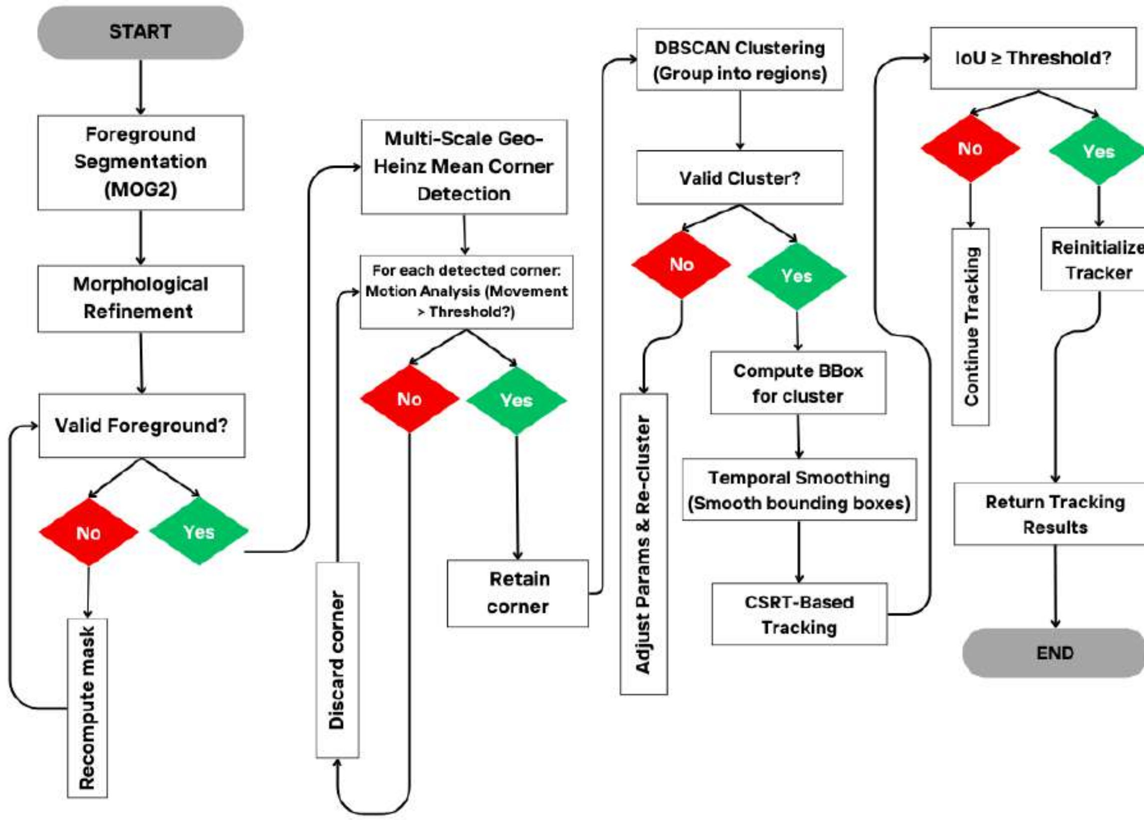

Figure 1: The process begins with foreground segmentation using MOG2, followed by morphological refinement to enhance the foreground mask. Multi-scale Geo-Heinz Mean corner detection is applied to extract robust corner features, which are filtered based on motion analysis. The validated features are grouped into potential vehicle regions using DBSCAN clustering. Temporal smoothing is applied to reduce fluctuations in bounding boxes, and CSRT-based tracking ensures stable vehicle tracking, with reinitialization based on the Intersection-over-Union (IoU) criterion.

Algorithm 1 Vehicle Detection and Tracking for Highway Surveillance

```
Input: Video frames, initial background model
Output: Vehicle bounding boxes and tracking results for each frame
Set Geo-Heinz Mean corner response parameters
Set corner movement threshold
Initialize Background Subtractor MOG2
Initialize Temporal Smoothing Mechanism
Initialize CSRT Trackers
for each frame in video do
    Step 1: Foreground Segmentation
    Apply MOG2 to extract foreground mask
    Refine foreground mask using morphological operations (Opening + Closing)
    Step 2: Multi-Scale Geo-Heinz Mean Corner Detection
    for each moving region in foreground mask do
        Detect corners using multi-scale Geo-Heinz Mean corner response
        Store detected corners for motion analysis
    end for
    Step 3: Motion Analysis and Feature Filtering
    for each corner in detected corners do
        Compute displacement of corner between current and previous frame
        if displacement > corner movement threshold then
            Retain corner for further processing
        else
            Discard corner
        end if
    end for
    Step 4: DBSCAN Clustering
    Apply DBSCAN clustering on retained corners
    for each cluster in DBSCAN results do
        Compute bounding box around cluster
    end for
    Step 5: Temporal Smoothing
    for each vehicle bounding box in detected clusters do
        Apply temporal smoothing to reduce bounding box fluctuations across frames
    end for
    Step 6: CSRT-Based Tracking
    for each vehicle region in vehicle regions do
        Initialize CSRT tracker with bounding box
        if IoU between predicted and actual bounding box < threshold then
            Reinitialize CSRT tracker using Intersection-over-Union (IoU) criterion
        end if
    end for
end for
Return: Vehicle bounding boxes and tracking results for each frame
```

4 Experimental Results and Evaluation

In this section, we present the experiments and evaluate the results. First, we will describe the Experimental setup in the first subsection, followed by a discussion of the evaluation metrics used to assess the performance in the second subsection. Next, we will present the Qualitative results, which include images of frames where vehicles were successfully detected. Finally, we will explore and analyze the Quantitative results of our experiments.

4.1 Datasets and Experimental Setup

We used three videos to run our experiments, all captured by cameras on highways, focusing on real and natural traffic scenarios. The first video was recorded on a sunny day on a one-way highway, with vehicles moving in a single direction. The second video was taken at night on a two-way highway, with vehicles traveling in both directions. This video, sourced from a live camera on French roads, has relatively low quality, and the headlights of the vehicles create numerous corners, making it particularly challenging. The third video was recorded on a two-way highway during a rainy day, sourced from YouTube. In this scenario, the presence of rain reduced visibility, and vehicles often moved in close proximity to each other, making detection and tracking even more difficult. These videos were selected to provide a variety of conditions, including different traffic flow directions and challenging weather and lighting effects, to comprehensively test the system's robustness.

4.2 Evaluation Metrics

To evaluate the performance of our algorithm on the experiments, we use two kinds of metrics. The first type is Qualitative metrics, which includes images of frames where vehicles are detected and tracked. These images are presented in a composite view consisting of four panels: the Original frame, showing the original video frame; Moved corners, which highlight the corners that are candidates for clustering and vehicle detection;DBSCAN clusters, illustrating the clusters formed by the detected corners, with blue bounding boxes indicating potential vehicles; and finally, CSRT, which displays stable green bounding boxes around the detected vehicles with their respective IDs. These qualitative results allow us to visually assess the performance and quality of the algorithm in real-world scenarios.

The second type is Quantitative metrics, which are used to assess the algorithm's performance on our dataset. These include the ID switch count, which measures the number of times a vehicle's ID changes during tracking, indicating potential instability in tracking; Inter-frame IoU consistency, which evaluates the consistency of the Intersection-over-Union (IoU) between predicted and ground truth bounding boxes across frames; and Bounding box variance, which examines the variation in bounding box coordinates (position and size) across frames to assess tracking stability. Additionally, we measure the Number of truly detected vehicles, which represents the number of vehicles correctly detected by the algorithm, the Number of missed vehicles, which tracks how many vehicles were not detected, and the Number of wrongly detected vehicles, which counts false positives. These quantitative metrics provide a detailed and numerical assessment of the algorithm's detection and tracking accuracy, helping to measure its effectiveness under various conditions.

4.3 Qualitative Evaluation of the Daytime Video Results

In this video, there are three vehicles, all of which are detected and tracked very well. At some frames like 11, the DBSCAN bounding boxes are missed, but the CSRT tracker stabilizes the tracking and ensures that the vehicles are not lost. As seen in the detected corners view, corners only appear for the moving parts of the video. Static parts, like the road, or areas with low movement, such as trees swaying gently in the wind, do not show any corner points, which is necessary for detecting vehicles accurately. The vehicle IDs do not change significantly after detecting the vehicle and remain consistent until the vehicle exits the frame. The bounding boxes span the vehicles' shapes properly and remain stable and consistent, even when they move. Overall, the performance in this scenario is stable, and the tracking remains reliable throughout the video. These results demonstrate that this method is well-suited for vehicle tracking, highlighting its potential for real-time highway surveillance applications.

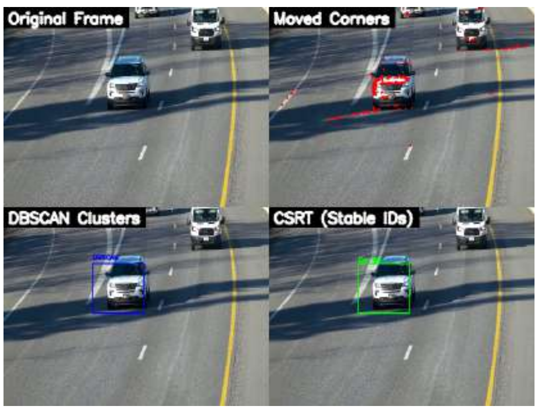

Figure 2: frame00024

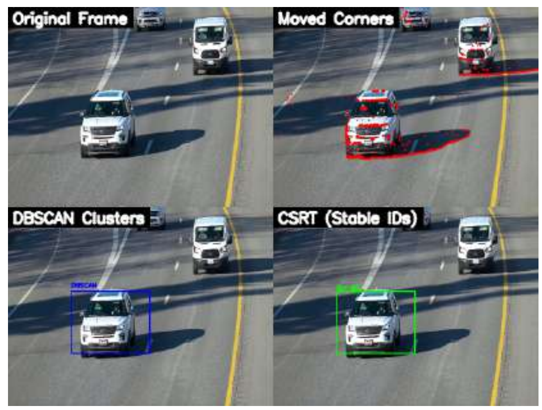

Figure 3: frame00040

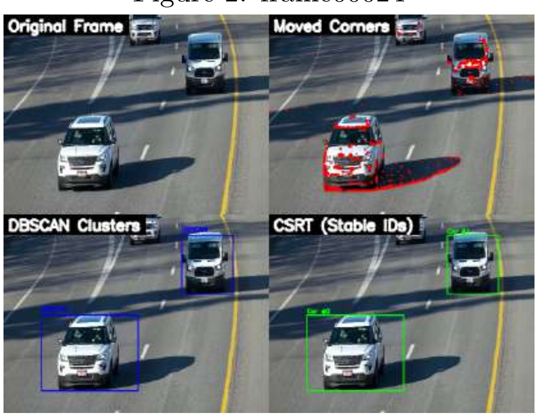

Figure 4: frame00046

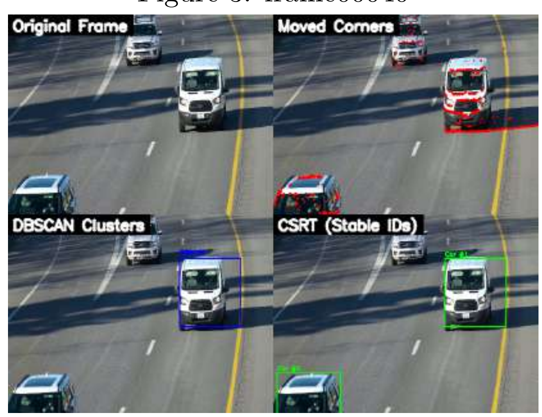

Figure 5: frame00060

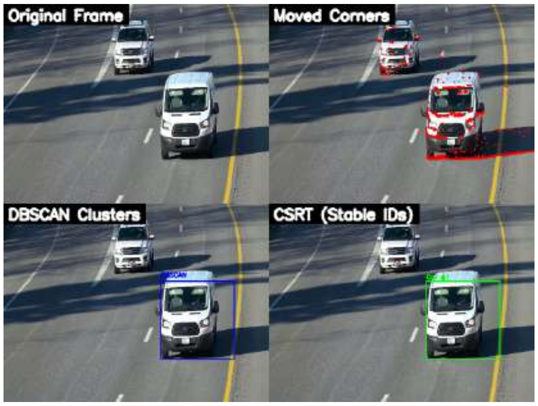

Figure 6: frame00070

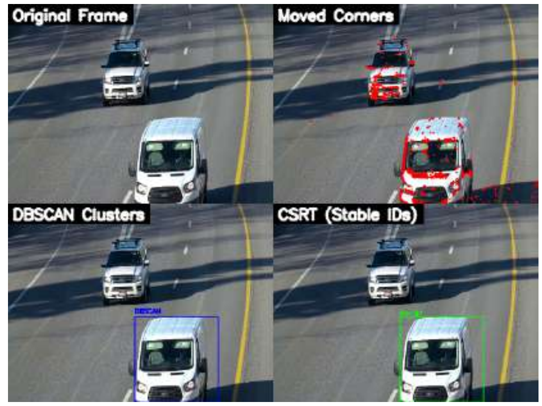

Figure 7: frame00083

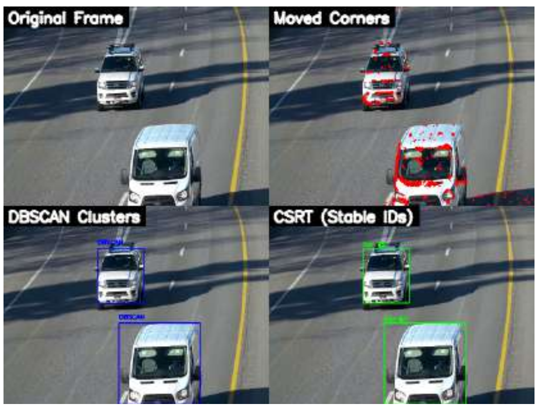

Figure 8: frame00084

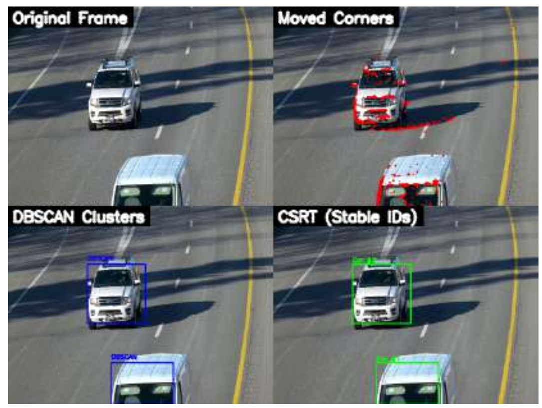

Figure 9: frame00090

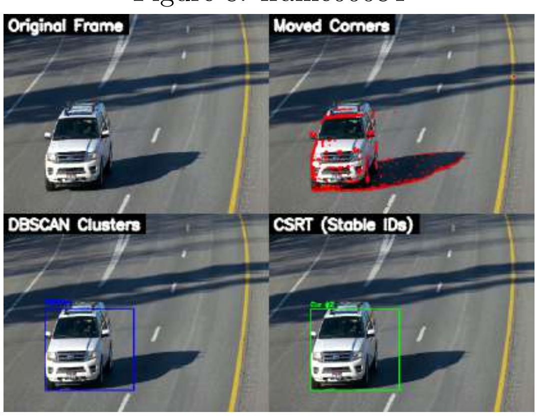

Figure 10: frame00102

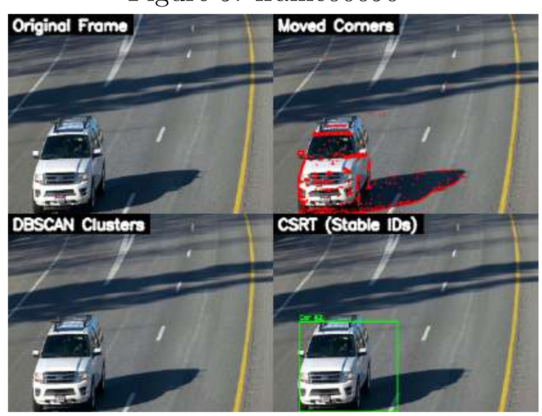

Figure 11: frame00106

Figure 12: Qualitative Evaluation of the Daytime Video Results

4.4 Qualitative Evaluation of the Nighttime Video Results

In this video, recorded at night from a road surveillance camera in France, vehicles move in two directions: towards the camera and away from it. A total of 8 vehicles appear in the video. Among the vehicles moving towards the camera, 4 were present, of which 3 were successfully detected, and 1 was not detected at all. For the vehicles moving away from the camera, there were also 4, of which 1 was detected correctly, while 3 were missed. The cause of these missed detections is primarily the vehicles' distance from the camera and the limited number of corners detected. The bounding boxes covered the vehicles well, but one challenge was that the vehicle headlights created many corner points, which caused the bounding boxes to occasionally appear larger than necessary. This issue improved after the vehicles got closer to the camera. Despite these challenges, our performance in this scenario remained good and reliable.

The vehicles that were detected remained consistently tracked throughout the video, and their bounding boxes did not disappear. The vehicle IDs for the detected vehicles remained stable, and the algorithm effectively identified the correctly detected vehicles throughout the video.

Figure 13: frame00060

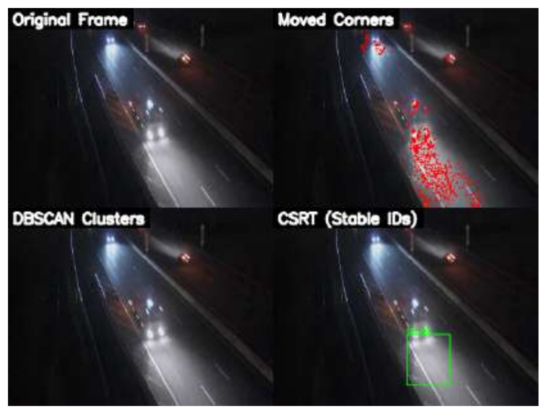

Figure 14: frame00080

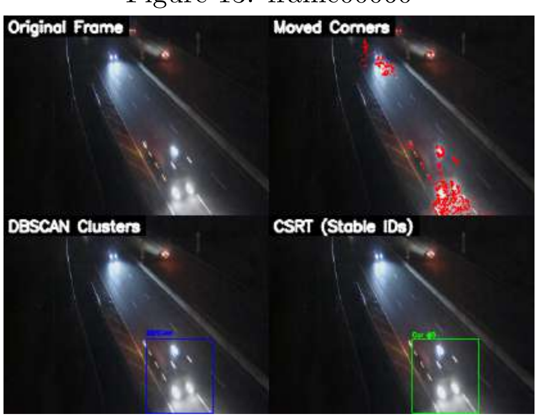

Figure 15: frame00090

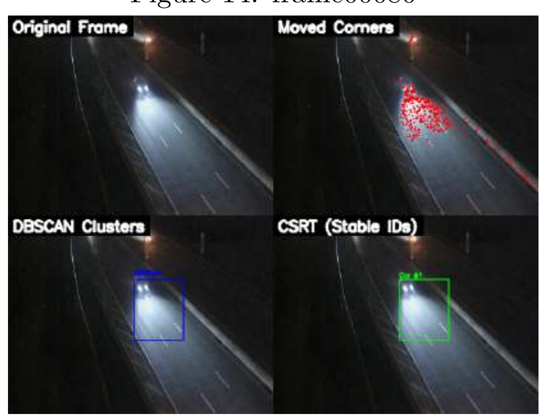

Figure 16: frame00110

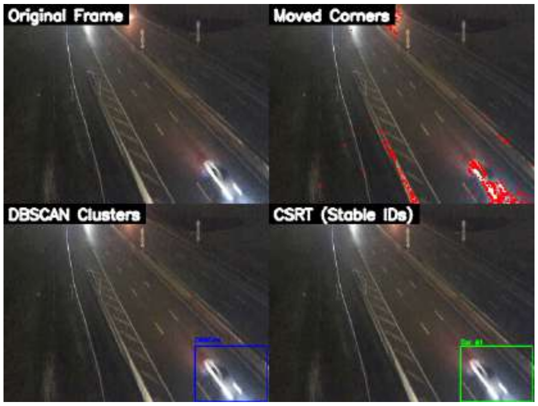

Figure 17: frame00135

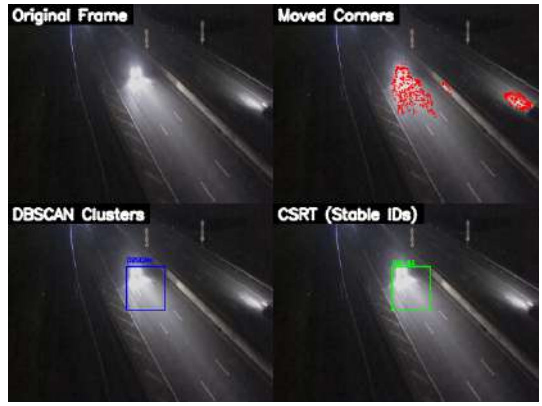

Figure 18: frame00238

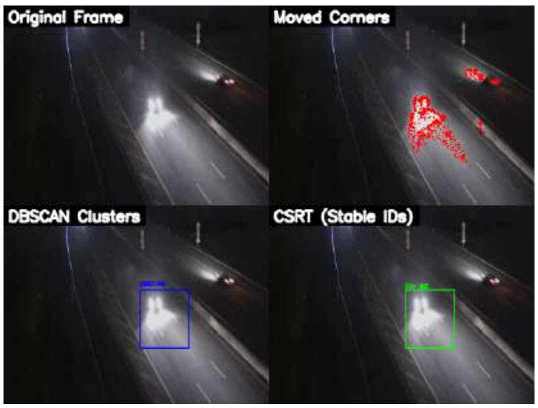

Figure 19: frame00250

Figure 20: frame00267

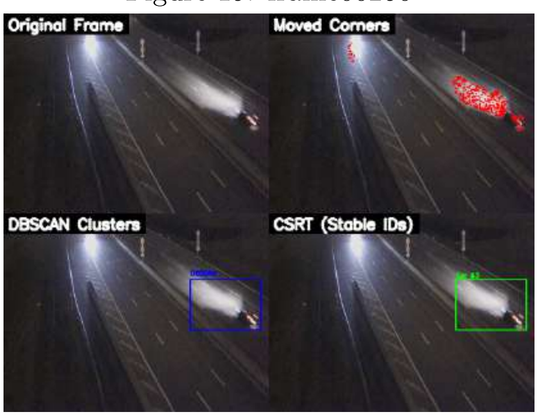

Figure 21: frame00330

Figure 22: frame00349

Figure 23: Qualitative Evaluation of the Nighttime Video Results

4.5 Qualitative Evaluation of the Rainy Day Video Results

This video was recorded on a rainy day, where the ground was wet and small waves of water were created as the vehicles moved, which generated corners and posed a challenge in accurately detecting the vehicles. This factor is similar to the effect of headlights at night. In this video, 9 vehicles passed by, 6 of which were moving away from the camera and 2 were moving towards it. Notably, one of the vehicles moving towards the camera was too far away and can be disregarded. Interestingly, one of the vehicles was a large truck that was very far from the camera and was not visible on the road, but our algorithm was still able to detect it.

Of the vehicles moving away from the camera, only one was not detected; this occurred at the beginning of the video when the vehicle was too far from the camera and there were fewer detected corners. Among the vehicles moving towards the camera, one was a truck that was easily detected with an appropriately sized bounding box. Another vehicle was very far from the camera, and once it entered the scene, the video ended, causing the vehicle to be missed, but this can be disregarded.

In most cases, the bounding boxes and vehicle IDs showed good stability, although there were a few instances where they disappeared. Despite this, the total number of vehicles detected remained accurate, and there were no incorrect detections or instances where a vehicle was detected multiple times with different IDs.

Figure 24: frame00005

Figure 25: frame00020

Figure 26: frame00030

Figure 27: frame00049

Figure 28: frame00060

Figure 29: frame00073

Figure 30: frame00079

Figure 31: frame00100

Figure 32: frame00150

Figure 33: frame00158

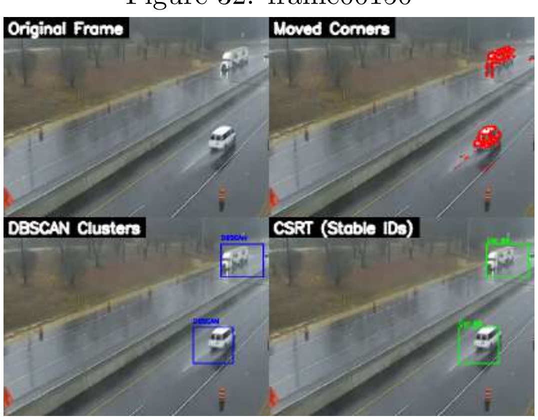

Figure 34: frame00175

Figure 35: frame00185

4.6 Quantitative Evaluation of the Experiments Results

In this section, we present the quantitative evaluation of the proposed vehicle detection and tracking algorithm across three different video scenarios: Daytime, Nighttime, and Rainy Day. The evaluation is based on several important metrics, including precision, recall, ID switch count, IoU consistency, bounding box variance, and the number of missed or wrongly detected vehicles.

For the Daytime Video, the algorithm performed exceptionally well, achieving a high precision of 100% and a recall of 75%. The ID consistency was maintained without significant ID switches, indicating stable tracking throughout the video. The bounding boxes were stable, with minimal variance, showing that the algorithm could accurately track the vehicles. The number of missed vehicles was low, with only one vehicle not detected, while there were no wrong detections.

In the Nighttime Video, the detection accuracy was slightly impacted due to challenging lighting conditions, with precision remaining at 100% but recall dropping to 50%. The ID switch count was minimal, and the bounding box variance increased slightly, especially around the vehicle headlights. Despite the challenges, the algorithm successfully tracked most vehicles and handled the headlights' influence on corner detection. The number of missed vehicles was higher, with 4 vehicles missed in total.

For the Rainy Day Video, the algorithm showed reasonable performance despite the wet conditions and reduced visibility. The precision remained 100%, and the recall was 67%. The ID switch count was low, and the bounding box variance was slightly higher due to the reflections and wet ground. There were 3 missed vehicles, but the algorithm still detected the majority of the vehicles. The number of wrongly detected vehicles was zero, indicating the algorithm's reliability in a dynamic and challenging environment.

Overall, the proposed algorithm demonstrated strong performance in all three scenarios, with the highest accuracy in the daytime video. The rainy and nighttime conditions introduced challenges, but the algorithm adapted well, maintaining high precision and stability. These results indicate that the method is robust and well-suited for real-time vehicle detection and tracking in varying environmental conditions.

Metric	**Daytime Video**	**Nighttime Video**	**Rainy Day Video**
True Positives (TP)	3	4	6
False Positives (FP)	0	0	0
False Negatives (FN)	1	4	3
Precision	100%	100%	100%
Recall	75%	50%	67%

Table 1: Detailed Quantitative Evaluation Metrics

5 Conclusion and Future Works

In this paper, we presented a robust vehicle detection and tracking framework designed for highway camera surveillance, leveraging multi-scale Geo-Heinz Mean corner response, DBSCAN clustering, and CSRT-based tracking. Our proposed method demonstrated strong performance in detecting and tracking vehicles in real-world highway environments, showcasing its effectiveness across various challenging conditions, including daytime, nighttime, and rainy weather scenarios.

The qualitative results illustrated that our approach is capable of handling diverse environmental conditions, maintaining stable tracking, and minimizing detection errors. The quantitative evaluation further validated the algorithm's accuracy, achieving high precision in all video scenarios. However, challenges were observed under low-visibility conditions, such as at night or during rain, where the performance slightly dropped due to factors like vehicle headlights and wet surfaces. Despite these challenges, the algorithm successfully tracked the majority of vehicles, demonstrating its robustness and potential for real-time application in highway surveillance systems.

Our method does have limitations, such as the need for proper parameter tuning and occasional difficulties with vehicle overlap, especially when tracking larger or unusually shaped vehicles like trucks. These vehicles, being longer than regular cars, pose additional challenges. However, despite these limitations, the method still provides satisfactory and reliable outputs. In future works, these limitations can be addressed by refining parameter tuning, enhancing the detection of overlapped vehicles, and improving the algorithm's ability to track large or non-standard vehicle shapes more effectively.

For future work, additional features beyond corner points, such as edges, can be integrated into the method to improve detection accuracy. However, in this study, relying solely on corners provided reliable and satisfactory results. Further improvements can include adaptive parameter tuning for the Geo-Heinz Mean corner response and DBSCAN clustering to enhance performance in dynamic environments. Further experiments with larger and more diverse datasets could also help assess the algorithm's robustness across different traffic types and camera setups.

By addressing these future directions, we believe the proposed method can be further optimized to provide highly reliable, scalable, and real-time vehicle detection and tracking solutions for highway surveillance and intelligent transportation systems.

References

[1] L. Jian, Z. Wangzi, and G. Hui, "Vehicle type recognition based on harris corner detector," in *International Conference on Transportation Engineering 2009 (ICTE 2009)*. American Society of Civil Engineers (ASCE), 2009, pp. 3320–3325.

[2] M. Anandhalli and V. P. Baligar, "An approach to detect vehicles in multiple climatic conditions using the corner point approach," *Journal of Intelligent Systems*, vol. 27, no. 3, pp. 363–376, 2018.

[3] M. E. Munajat, D. H. Widyantoro, and R. Munir, "Vehicle detection and tracking based on corner and lines adjacent detection features," in *Proceedings of the 2nd International Conference on Science in Information Technology (ICSITech)*. IEEE, 2016.

[4] K. K. KV, P. Chandrakant, S. Kumar, and K. KJ, "Vehicle speed detection in video frames using corner detection," in *Proceedings of the International Conference on Signal and Image Processing (ICSIP)*. IEEE, 2013.

[5] W. Yuan, Y. Jiang, and F. Cai, "Research of vehicle identification based on adaptive corner detection," in *Proceedings of International Conference on Computer Science and Information Technology*, ser. Advances in Intelligent Systems and Computing, vol. 255. Springer India, 2014, pp. 75–82.