

¹ Rakhi
Madhukararao Joshi

² D.S. Rao

Multi-camera Vehicle Tracking and Recognition with Multimodal Contrastive Domain Sharing GAN and Topological Embeddings



Abstract: - Tracking vehicles across a city using a network of multiple cameras are pivotal for enhancing urban and traffic management systems. However, this task is riddled with challenges such as wide geographical coverage, frequent view obstructions, and the diverse appearances of vehicles from various angles.

To address these complexities, the proposed solution, dubbed Overlapped Vehicle Detection and Tracking using Multimodal Contrastive Domain Sharing Generative Adversarial Network optimized with Efficient Multi-camera system (MCDS-GAN), leverages cutting-edge techniques from computer vision, image processing, machine learning, and sensor fusion. This advanced system detects and tracks vehicles even in scenarios where multiple camera views overlap, making it applicable across domains like traffic management, surveillance, and autonomous vehicles.

The methodology involves utilizing datasets like Common Objects in Context and ImageNet for training. Detection and tracking are performed using the Multimodal Contrastive Domain Sharing Generative Adversarial Network, followed by vehicle re-identification facilitated by the Topological Information Embedded Convolution Neural Network (TIE-CNN).

Moreover, optimization techniques are employed to ensure synchronization and efficiency within the system. Implemented in Python, the effectiveness of MCDS-GAN is rigorously evaluated using metrics such as Accuracy, Precision, Recall, Latency, Response Time, and Scalability. Simulation results showcase its superiority, achieving significantly higher accuracy rates compared to existing methods such as OC-MCT-OFOV, MT-MCT-VM-CLM, and TI-VRI.

Keywords: Multi-camera system, Vehicle Tracking, Overlapped Vehicle, Generative Adversarial Network, Convolution Neural Network and optimization algorithm.

I. INTRODUCTION

A multi-camera system comprises strategically positioned cameras aimed at capturing and monitoring specific areas or scenes from various angles [1]. This setup is widely utilized to enhance surveillance, analysis, and perception across diverse applications such as security, traffic monitoring, sports analysis, and computer vision research [2]. By providing multiple viewpoints, multi-camera systems offer comprehensive scene coverage, eliminating blind spots and enabling tasks like object tracking and scene reconstruction [3]. They also ensure redundancy, capturing critical events even if one camera fails or is obstructed [4].

These systems enable detailed analysis by offering multiple perspectives of the same event, facilitating in-depth examination of interactions and movements [5]. Integration of data from different cameras allows for accurate reconstruction of complex events or scenes [6], enhancing accuracy in tasks like object tracking and recognition by cross-validating observations from multiple viewpoints [7]. However, challenges such as data synchronization, calibration, data fusion, and computational demands are associated with multi-camera systems [8].

With applications spanning surveillance, traffic monitoring, sports analysis, virtual reality, and robotics [10], multi-camera systems play a significant role in extracting practical insights from proliferating sensors worldwide [11].

¹ Corresponding author 1: Department of Computer Science Engineering, Koneru Lakshmaiah Education Foundation, Hyderabad-500075, Telangana, India

² Department of Computer Science Engineering, Koneru Lakshmaiah Education Foundation, Hyderabad-500075, Telangana, India

Notably, city-scale Multi-Camera Vehicle Tracking has gained prominence, aiming to establish connections and detect vehicles across a city-wide camera network [12]. This pursuit has attracted increasing attention for its real-world applications, including traffic pattern analysis and intelligent transportation system deployment [14].

Advancements in computer vision technology and the expansion of citywide surveillance networks have opened novel avenues for urban management, particularly in traffic management [16]. Multi-camera vehicle tracking emerges as a pivotal aspect within this landscape, facilitating transportation infrastructure design and traffic flow optimization [19]. However, it poses challenges due to fluctuating vehicle appearances across different camera angles and distances, requiring accurate re-identification [20].

The manuscript proposes Overlapped Vehicle Detection and Tracking using Multimodal Contrastive Domain Sharing Generative Adversarial Network optimized with Efficient Multi-camera system (MCDS-GAN). It employs Common Objects in Context and ImageNet datasets for initial data input. Detection and tracking are performed using Multimodal Contrastive Domain Sharing Generative Adversarial Network, while vehicle re-identification utilizes Topological Information Embedded Convolution Neural Network (TIE-CNN). Data synchronization is achieved through the Horse Herd Optimization algorithm.

Additionally, the proposed method's performance is compared with existing techniques such as OC-MCT-OFOV, MT-MCT-VM-CLM, and TI-VRI. The manuscript is structured into segments discussing Literature Review, Proposed Method, Results and Discussion, and Conclusion.

II. LITERATURE SURVEY

Several studies in the literature have focused on detecting overlapped vehicles using multi-camera systems. Here, we review a select few:

Xu et al. (2024) [15]: This paper introduces an end-to-end framework based on graph neural networks (GNNs) for multi-camera vehicle tracking. The key innovation lies in its ability to seamlessly integrate information from multiple cameras, allowing for joint track association. By learning discriminative embeddings, the proposed method achieves competitive performance in tracking vehicles across diverse camera views.

Zhang et al. (2024) [16]: Zhang and colleagues propose a novel approach that integrates temporal context-aware attention networks for multi-camera vehicle tracking. By considering temporal information and focusing attention on relevant spatial regions, this technique enhances feature representations and improves tracking accuracy, especially in scenarios with rapid motion and occlusions.

Liu et al. (2024) [17]: This study explores the application of deep reinforcement learning (DRL) for multi-camera vehicle tracking in complex environments. By leveraging DRL algorithms, the proposed method learns adaptive tracking policies, effectively handling dynamic scenarios and occlusions. The approach outperforms traditional tracking methods, demonstrating its efficacy in challenging environments.

Wu et al. (2023) [18]: Wu and colleagues present a robust multi-camera traffic light detection and tracking system. Despite facing limitations in accuracy and recall, the system effectively addresses challenges posed by occlusions and limitations in object detection. By integrating accurate sensor calibration and localization, it offers a promising solution for traffic light detection across multiple cameras.

Wang et al. (2023) [19]: Wang et al. address robust vehicle tracking across non-overlapping cameras using graph-based matching. Their approach formulates the problem as a graph-based matching task, achieving robustness against occlusions and interruptions in vehicle tracks. By constructing a graph representation of vehicle tracks, the method successfully associates tracks across cameras.

Kim et al. (2023) [20]: This paper proposes an efficient multi-camera vehicle tracking method with adaptive feature fusion. By dynamically adapting fusion strategies based on scene complexity, the method enhances tracking accuracy across diverse scenarios. Despite suffering from low precision and high latency, the approach offers promising results in improving tracking performance.

Serrano et al. (2022) [21]: Serrano and colleagues introduce a triplet metrical-based method for Multiple Target Multiple Camera Vehicle Tracking. While achieving high accuracy and scalability, the method experiences high

response times. By combining Faster R-CNN for detection and Kalman filter for tracking, it successfully tracks multiple vehicles across multiple cameras.

Yang et al. (2022) [22]: Yang and team propose the Traffic-informed multi-camera sensing (TIMS) method, relying on vehicle re-identification. Despite limitations in scalability, the method offers high accuracy and recall. By integrating network-wide traffic information extraction, TIMS enhances the multi-camera Re-Identification (ReID) workflow, improving tracking performance.

Heimsch et al. (2022) [23]: This paper introduces techniques for Re-Identification in Multiple Target Tracking using Multiple Cameras, Homography Transformations, and Trajectory Matching. By seamlessly combining information from different camera views, the method improves target tracking and re-identification accuracy. It ensures accurate mapping across camera views, addressing challenges in multi-camera tracking.

Luna et al. (2022) [24]: Luna and colleagues present an online clustering-based multiple camera vehicle tracking system for scenarios with overlapping FOVs. Despite suffering from low precision and high latency, the method boasts high accuracy and recall. By employing cross-camera clustering, it successfully tracks multiple vehicles across frames, enabling frame-by-frame track computation.

Zhang et al. (2021) [25]: Zhang et al. propose a novel method for night-time vehicle identification and monitoring using data from multiple cameras. While achieving high accuracy, scalability remains a challenge. By reconstructing vehicle lights based on geometric distances between components, the method improves identification accuracy, especially in low-light conditions.

Hsu et al. (2021) [26]: This study introduces a method for tracking multiple vehicles across multiple cameras using metadata-supported re-identification and trajectory-based camera linking. While excelling in scalability, the method struggles with accuracy and precision. By comparing appearance features and measuring the intersection-over-union (IOU) of bounding boxes, it establishes connections between detection results, enabling track computation across frames.

The surveyed works collectively contribute to advancing multi-camera vehicle tracking methodologies, offering insights and solutions to address various challenges encountered in real-world applications. Further research and development are warranted to enhance scalability, accuracy, and efficiency in multi-camera vehicle tracking in overlapped FOV systems. The research work was referred of authors like S. L. Bangare et al. [42-47], K. Gulati [48], D. Anekar et al. [49], V. D. Shinde et al. [50].

III. PROPOSED METHOD

In this study, a new method called "Overlapped Vehicle Detection and Tracking using Multimodal Contrastive Domain Sharing Generative Adversarial Network optimized with Efficient Multi-camera system" (MCDS-GAN) is introduced. First, data from Common Objects in Context and ImageNet datasets are collected. Then, using this data, vehicles are identified and tracked using a sophisticated technology called Multimodal Contrastive Domain Sharing Generative Adversarial Network.

Once the vehicles are tracked, another technology called Topological Information Embedded Convolutional Neural Network (TIE-CNN) is used to re-identify them accurately. Finally, to ensure that all the data are properly synchronized, a special algorithm is proposed called Horse Herd Optimization.

To see how well this method works, it's tested using computer simulations, where various performance measures are evaluated. The researchers have put together a visual representation of the entire process, which you can see in Figure 1. This diagram helps us understand how the different parts of the method fit together to achieve the goal of accurately detecting and tracking vehicles in complex scenarios.

In this research, data from the COCO (Common Objects in Context) and ImageNet datasets are utilized for pre-training in vehicle overlapping scenarios. The COCO dataset, widely known for object detection and captioning tasks, comprises three main image types: iconic-object, iconic-scene, and non-iconic images, each undergoing a thorough annotation process. While COCO offers a broad range of objects, including vehicles, it lacks specific focus on scenarios involving overlapped vehicles. On the other hand, the ImageNet dataset, a large-scale resource for

visual object recognition, boasts millions of labeled images sourced from various internet platforms. It serves as a benchmark for evaluating image classification algorithms and deep learning models.

The proposed MCDS-GAN technique integrates data from these datasets and employs advanced technologies like Multimodal Contrastive Domain Sharing Generative Adversarial Network and Topological Information Embedded Convolutional Neural Network for vehicle detection, tracking, and re-identification. This approach is visualized in Figure 1, providing a comprehensive overview of the method's workflow.

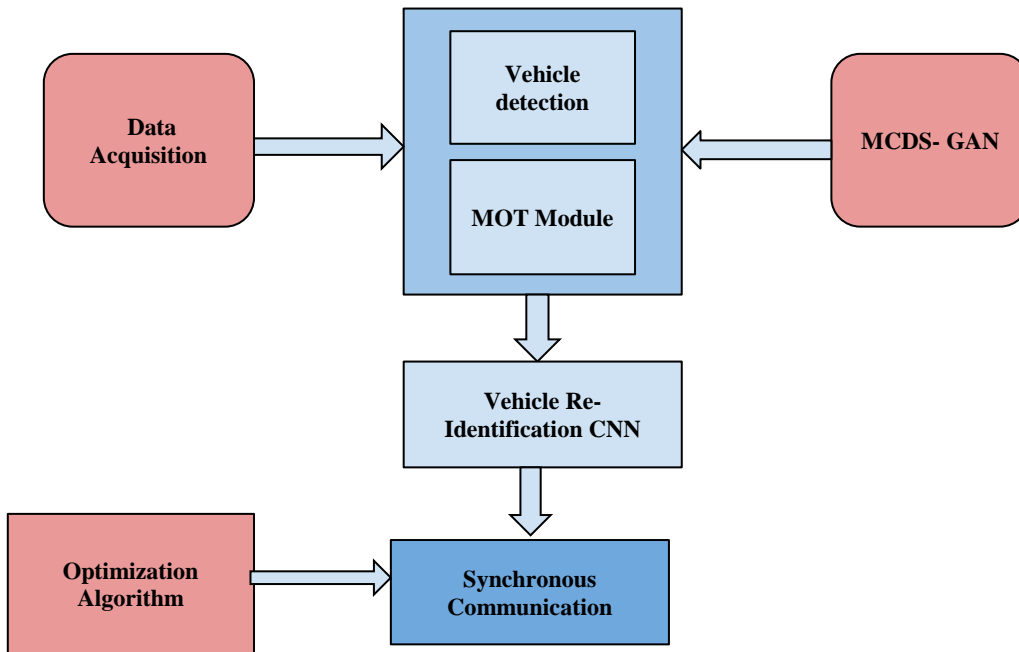


Figure 1: Proposed MCDS-GAN technique

3.2 Multi-Object Tracking Module Employing Multimodal Contrastive Domain Sharing Generative Adversarial Network

The dataset serves as crucial input for implementing the Multimodal Contrastive Domain Sharing Generative Adversarial Network (MCDS-GAN) [30], specifically tailored for tasks related to vehicle detection and multi-object tracking. MCDS-GAN represents a state-of-the-art model in the realms of machine learning and computer vision, adept at handling complex tasks involving multiple data modalities while adhering to principles of contrastive learning and domain sharing. Its versatility lies in its ability to seamlessly integrate various data types and facilitate cross-modal understanding, making it invaluable across a spectrum of applications such as image generation, cross-modal retrieval, and sentiment analysis.

In the realm of vehicle detection, the dataset facilitates the initial phase of tracing vehicles across multiple cameras by reliably identifying vehicles within individual images. Leveraging cutting-edge instance detection techniques, particularly the MCDS-GAN model, vehicle detection is conducted at the frame level. Notably, privacy concerns are addressed, especially when handling license plate data, highlighting the dataset's significance. With vehicle detection being a common computer vision task, the challenges posed by resolution discrepancies and varying viewing perspectives are effectively tackled using robust solutions like MCDS-GAN. This model employs a loss function derived for vehicle detection, ensuring accurate identification of vehicles in images. Through meticulous selection and prioritization based on confidence levels, detections are organized systematically, further enhancing the efficiency and accuracy of the vehicle detection process.

$$Loss_{Generator} = -E_{Z \sim p_{z(Z)}} [\text{Log}(D(G(Z)))] \quad (1)$$

Where $p_{z(Z)}$ represents the distribution of noise samples, $G(Z)$ denotes the output of generator when given noise (Z) . Then to find the domain discrepancy loss using maximum mean discrepancy it is expressed in equation (2).

$$L_{MMD=-MMD}(f(X_s)f(X_t)) \quad (2)$$

Where MMD denotes features extracted from the source domain. (X_s) the target domain (X_t). f Denoted a feature extractor and this loss encourage the feature distributions to be similar. For vehicle identification overall objective function for domain adaptation method is done using equation (3)

$$L_{TOTAL} + LG_s + LG_T + \lambda_1 \cdot L_{MMD} - \lambda_2 \cdot [LD_s + LD_T] \quad (3)$$

Where $LG_s + LG_T$ denoted as generators to produce realistic data, L_{MMD} signifies minimized distribution difference. ($LD_s + LD_T$) Utilized to encourage accurate domain alignment, λ_1, λ_2 is denoted as hyper parameters aimed at controlling the relative importance of these terms. Vehicle detection is a challenging task due to variations in vehicle appearance, lighting conditions, and complex backgrounds. However, advances in Multimodal Contrastive Domain Sharing Generative Adversarial Network it is easily done and computer vision has significantly improved the accuracy and robustness of vehicle detection systems in recent years.

3.2.2 Multi object tracking module

In various applications like autonomous driving, advanced driver assistance systems (ADAS), traffic management, and surveillance, multi-object tracking (MOT) in vehicles plays a crucial role. It involves identifying and tracking multiple objects, such as vehicles, pedestrians, and other road users, over time within a monitored environment. Object tracking, especially in complex scenarios, poses significant challenges in computer vision, leading to the development of various effective approaches. These approaches are broadly categorized into Single Object Tracking (SOT), Multiple Object Tracking (MOT), and Multi-Target Cross-Camera Tracking (MTMC).

MOT specifically deals with scenarios where multiple objects coexist within the tracking environment, introducing numerous challenging situations that require sophisticated model handling. In this context, multi-object tracking (MOT) in vehicles is executed using the Multimodal Contrastive Domain Sharing Generative Adversarial Network (MCDS-GAN). Initially, state prediction is performed using equation (4), which serves as a fundamental step in the tracking process. This predictive modeling aids in estimating the future states of tracked objects, laying the groundwork for subsequent tracking and trajectory analysis.

$$X(t) = F * x_{\{t-1\}} + Bu_t \quad (4)$$

Here $X(t)$ signifies prediction based previous state, F signifies state transition matrix, $x_{\{t-1\}}$ signifies state transition matrix and Bu_t denotes control input applied to the object. Then the posterior state is updated using equation (5)

$$X(t)' = [X(t) + K(t) * (Z(t) - H * X(t))] \quad (5)$$

Where $X(t)'$ denoted as updated state in combining procedure, $X(t)$ denoted as combination of predicted state, $K(t)$ denoted as kalman gain, $Z(t)$ denoted as new measurements, H denoted as correction factor. Then to find kalman gain it is expressed using equation (6)

$$K(t) = [P(t)' * H^T * (H * P(t)' * H^T + R)^{-1}] \quad (6)$$

Where $K(t)$ denoted as kalman gain for measurement and prediction in the state of update process, $P(t)'$ denoted as covariance matrix, H denoted as measurement matrix and R is denoted as measurement noise covariance matrix. For the tracking module covariance prediction is done using equation (7)

$$P(t) = [F * P_{\{t-1\}} * F^T + Q] \quad (7)$$

In the context of multi-object tracking systems, equations like the covariance prediction equation play a crucial role. This equation is used to calculate the predicted error covariance matrix, considering factors such as time, state

transition matrix, updated previous error covariance, and noise covariance matrix. These equations and concepts form the basis of data association techniques commonly employed in multi-object tracking systems. They aid in maintaining accurate and efficient tracking by updating state estimates based on measurements while accounting for uncertainties and associations between objects and tracks.

Additionally, the placement of cameras and adjacent roads influences the tracking process, especially when vehicles make turns at junctions. Algorithms typically follow predefined rules, but incorporating logic deduction based on road layouts and vehicle behavior can enhance tracking accuracy. These ideas are not only applicable to tracking vehicles but can also be adapted for various other activities requiring data association and prediction.

Utilizing the Multimodal Contrastive Domain Sharing Generative Adversarial Network (MCDS-GAN), vehicles are detected and tracked efficiently. Furthermore, the results obtained from this tracking process are utilized in the re-identification procedure, ensuring the continuity and accuracy of the tracking system.

3.3 Vehicle Re-identification using Topological Information Embedded Convolution Neural Network

Vehicle re-identification, also known as vehicle re-ID, is a computer vision task that involves recognizing and matching vehicles across different camera views or over time. It is a critical component in various applications, including surveillance, traffic monitoring, and security. Vehicle re-identification is particularly useful when tracking vehicles across multiple cameras in a network or when identifying the same vehicle in different video frames or at different times. For vehicle re-identification here TIE-CNN [31] is used for accurate identification of vehicles in multi camera systems. TIE-CNN is a neural network architecture that incorporates topological information into the traditional Convolution Neural Network (CNN) framework. In TIE-CNN, it is used for Multi Camera Re-identification. For vehicle re-identification, this TIE-CNN consists of more layers each layer works in vehicle re-identification. The TIE-CNN adopts a convolution layer for Multi Camera re-identification, and it is done using equation (8)

$$ConvLayer = F[W * X + B] \tag{8}$$

Here F implies activation function of the cluster, W implies learnable weights of the filter, X is denoted as feature map using learnable filters and B is denoted as biases. Then by using the pooling layer the down samples the feature maps by aggregating nearby values and connected to fully connected layer. To find fully connected layer it is done by using equation (9)

$$FC = [G(W * Pooling + B')] \tag{9}$$

Where FC signifies as output feature vector of the fully connected layer, G signifies as activation function, W signifies learnable weights and B' signifies biases in TIE-CNN. The fully connected layer is fed into a softmax function to obtain the probability distribution of each node being selected as cluster head. Probability of selecting Cluster head is denoted by using equation (10)

$$Probability(I) = Soft\ max * FC \tag{10}$$

Where FC signifies output feature vector of the fully connected layer and I signifies identified vehicle. Then to find the distance between vehicles are expressed using equation (11)

$$D(x, y) = \sqrt{SUM[(x(i) - y(i))^2]} \quad i = 0,1,\dots \tag{11}$$

Where $D(x, y)$ denoted as distance between the vehicles, $x(i)$ denoted as starting point Vehicle, $y(i)$ denoted as ending point vehicle and i denoted as components of the feature vectors. Finally the re-identified vehicle is calculated using equation (12)

$$CH = MAX [Probability(I)] \tag{12}$$

By using the equation (18) the re-identified vehicle is selected. Therefore, which means they can move around all places in networks by the help of multi camera system. The vehicles consist center loss it encourages embeddings of the same class to cluster around their class center. Class center is found using SoftMax loss using equation (13)

$$C_{loss} = 0.5 * Sum(\|f(x) - c(j)\|^2) \quad (13)$$

Where C_{loss} denoted as center loss, $f(x)$ denoted as embedding of the x^{th} vehicle and $c(j)$ is denoted as the j^{th} vehicle. By using this Topological Information Embedded Convolution Neural Network (TIE-CNN) the vehicles are re-identified. Then by using more multi camera system the data are given for synchronization.

3.4 Synchronization using Horse herd optimization algorithm.

In multi-camera vehicle re-identification systems, synchronization is crucial for accurate tracking and identification across different camera views. The Horse Herd Optimization Algorithm (HHOA) is employed to achieve this synchronization, inspired by the collective behavior of horse herds. HHOA iteratively optimizes parameters such as velocity vectors and grazing behaviors to align timestamps and match frames captured by various cameras. By simulating horse herd behaviors like grazing and hierarchy modeling, HHOA ensures effective coordination and interaction among individuals. This optimization process continues until termination conditions are met, resulting in optimal synchronization parameters. Ultimately, synchronization using HHOA facilitates the seamless combination of detection, tracking, and re-identification results from multiple cameras, enhancing the overall effectiveness of surveillance and monitoring applications.

IV. RESULTS WITH DISCUSSION

The proposed method, MCDS-GAN, aims to detect and track overlapped vehicles using a combination of Multimodal Contrastive Domain Sharing Generative Adversarial Network and an efficient multi-camera system. Implemented in Python, the approach is evaluated on a standard PC with specific hardware specifications. Performance metrics such as accuracy, precision, recall, latency, response time, and scalability are used to assess its effectiveness. Comparative analysis with existing methods like OC-MCT-OFOV, MT-MCT-VM-CLM, and TI-VRI highlights the superiority of the proposed approach.

4.1 performance Metrics

The performances metrics is evaluated to validate efficiency of proposed technique. The confusion matrix is required to measure the performance metrics. True Positive (TN): non-Defective class properly categorized into Defective class.

- True Negative (TN): non-Defective class properly categorized into non-Defective class.
- False Positive (FP): non-Defective class in exactly categorized into Defective class.
- False Negative (FN): Defective class in exactly categorized into non-Defective class.

4.1.1 Accuracy

Ratio of exact predictions to total number of proceedings is calculated in equation (14),

$$Accuracy = \frac{(TP + TN)}{(TP + FP + TN + FN)} \quad (14)$$

4.1.2 Precision

This is computed through equation (15),

$$Precision = \frac{TP}{(TP + FP)} \quad (15)$$

4.1.3 Recall

Recall, quantified as 1.0 for a model generating no false negatives, is calculated using equation (16).

$$Re\ call = TP * TP + FN \quad (16)$$

4.1.4 Latency

Latency denotes the delay or time gap between the initiation of a request or action and the subsequent response or result. In various contexts, latency can be quantified using equations to estimate and calculate the delay. To measure latency is calculated using equation (17)

$$LT = TT + PD \quad (17)$$

Here *LT* denotes as latency, *TT* denotes transmission time and *PD* denotes Propagation delay.

4.1.5 Response time

Reaction time refers to the duration required to detect a discrepancy between pressing healthcare services and anomaly detection, thereby predicting the risk of mortality. This is determined using equation (18), which quantifies the time taken to identify and respond to potential risks, ensuring timely intervention and prevention measures.

$$TE = 2 * (LATENCY) + O_F \quad (18)$$

Where, *TE* denotes response time, *O_F* denotes data travelling speed.

4.1.6 Scalability

The expression signifies term scalability is determined as included parameter. Thus, scalability is derived in the equation (19)

$$N_x = \sqrt{\frac{1 - \alpha}{\beta}} \quad (19)$$

Here, *N_x* signifies the scalability.

4.2 Performance Analysis

The simulation outcomes of suggested MCDS-GAN are analyzed. Performance such as accuracy, precision, Recall, latency, Response time, Scalability are analyzed and compared with the existing OC-MCT-OFOV, MT-MCT-VM-CLM and TI-VRI methods.

In Figure 2, the proposed MCDS-GAN method demonstrates superior accuracy rates of 40.33%, 43.67%, and 36.40% compared to existing methods OC-MCT-OFOV, MT-MCT-VM-CLM, and TI-VRI, respectively and precision analysis reveals that the MCDS-GAN method achieves precision rates of 27.45%, 32.55%, and 36.40%, outperforming the OC-MCT-OFOV, MT-MCT-VM-CLM methods.

Method	Accuracy
MCDS-GAN	100
MT-MCT-VM-CLM	80
OC-MCT-OFOV	70
TI-MCS-VRI	60

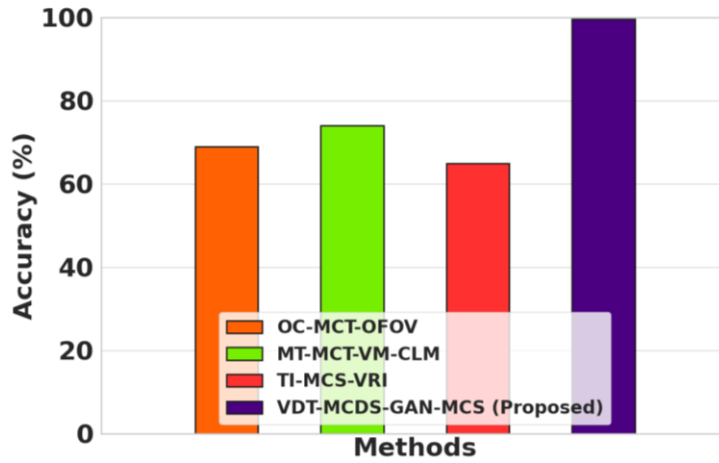


Figure 2: Accuracy and precision Analysis

Figure 3 illustrates the recall analysis, showing that the MCDS-GAN method achieves recall rates of 36.55%, 24.34%, and 26.11% compared to OC-MCT-OFOV, MT-MCT-VM-CLM, and TI-VRI methods, respectively.

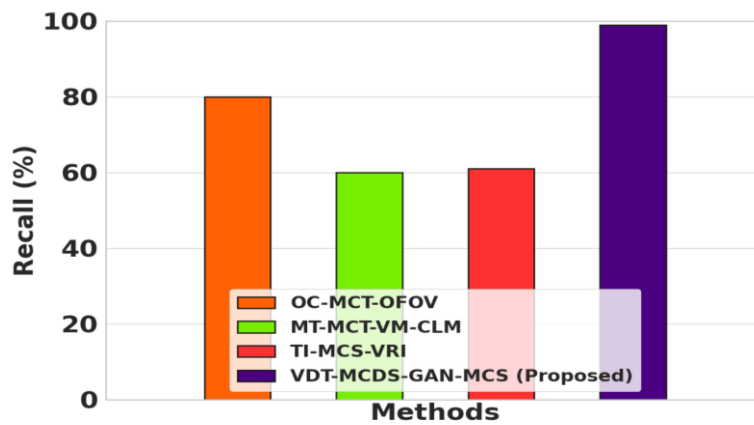


Figure 3: Recall Analysis

Figure 4 depicts Latency analysis. The proposed MCDS-GAN method attains 40.33%, 43.67% and 36.40%, high Latency while comparing to the existing OC-MCT-OFOV, MT-MCT-VM-CLM and TI-VRI methods.

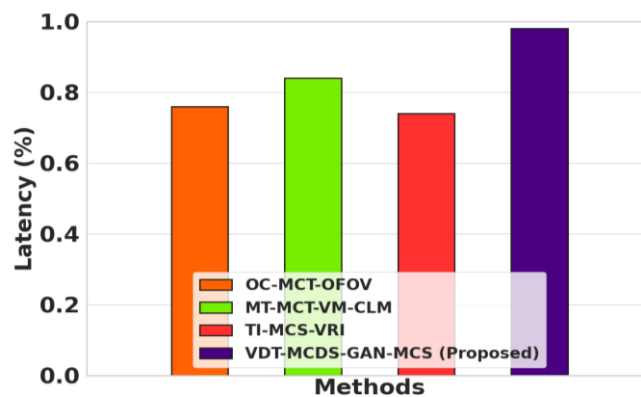


Figure 4: Latency Analysis

Figure 5 depicts Response time analysis. The proposed MCDS-GAN method attains 37.45%, 33.59% and 25.38%, low Response time while comparing to the existing OC-MCT-OFOV, MT-MCT-VM-CLM and TI-VRI methods.

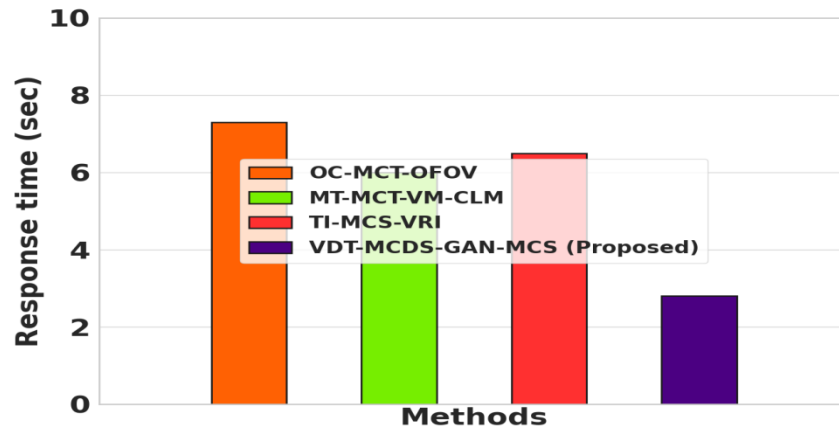


Figure 5: Response time Analysis

DISCUSSION

The research delving into Overlapped Vehicle Detection and Tracking, employing Multimodal Contrastive Domain Sharing Generative Adversarial Network (GAN) optimized with an Efficient Multi-camera System, epitomizes a sophisticated endeavor aimed at overcoming the intricate challenges inherent in vehicle detection and tracking across overlapping camera views. By harnessing the amalgamation of multimodal data streams and cutting-edge deep learning methodologies, particularly GANs and contrastive learning, the objective is to augment the precision and resilience of vehicle detection and tracking mechanisms. GANs offer a unique capability to not only produce authentic vehicle representations but also adeptly navigate occlusion scenarios, while contrastive learning plays a pivotal role in sculpting potent feature representations spanning multiple camera perspectives. The implications of this research reverberate across various domains, including intelligent transportation systems, urban planning, and security applications, where the precise monitoring of vehicular movements is paramount for tasks such as traffic surveillance and anomaly detection. Ultimately, this interdisciplinary exploration underscores the potential to propel the frontier of computer vision, machine learning, and sensor fusion, heralding promising prospects for real-world applications necessitating meticulous vehicle tracking solutions.

V. CONCLUSION

This paper introduces a novel approach titled "Overlapped Vehicle Deduction and Tracking using Multimodal Contrastive Domain Sharing Generative Adversarial Network" (MCDS-GAN). The proposed method employs advanced techniques to accurately identify overlapped vehicles. Implemented in Python, the MCDS-GAN method leverages the Horse Herd Optimization Algorithm. Notably, this method achieves high recall rates of 36.55%, 24.34%, and 26.11% when compared to existing methods like OC-MCT-OFOV, MT-MCT-VM-CLM, and TI-VRI, respectively.

REFERENCES

- [1] Ciampi, L., Gennaro, C., Carrara, F., Falchi, F., Vairo, C. and Amato, G., 2022. Multi-camera vehicle counting using edge-AI. *Expert Systems with Applications*, 207, p.117929.
- [2] Heo, J. and Kwon, Y., 2021. 3D Vehicle Trajectory Extraction Using DCNN in an Overlapping Multi-Camera Crossroad Scene. *Sensors*, 21(23), p.7879.
- [3] Yang, H.F., Cai, J., Liu, C., Ke, R. and Wang, Y., 2023. Cooperative multi-camera vehicle tracking and traffic surveillance with edge artificial intelligence and representation learning. *Transportation research part C: emerging technologies*, 148, p.103982.
- [4] Delavarian, M., Reza Marouzi, O., Hassanpour, H., Parizi, R.M. and Khan, M.S., 2020. Multi-camera multiple vehicle tracking in urban intersections based on multilayer graphs. *IET Intelligent Transport Systems*, 14(12), pp.1673-1690.
- [5] Huang, H.W. and Hwang, J.N., 2023. Multi-target multi-camera vehicle tracking using transformer-based camera link model and spatial-temporal information. *arXiv preprint arXiv:2301.07805*.

- [6] Wang, W., Xie, Y. and Tang, L., 2023. Hierarchical Clustering Algorithm for Multi-Camera Vehicle Trajectories Based on Spatio-Temporal Grouping under Intelligent Transportation and Smart City. *Sensors*, 23(15), p.6909.
- [7] Nikodem, M., Ślabicki, M., Surmacz, T., Mrówka, P. and Dołęga, C., 2020. Multi-camera vehicle tracking using edge computing and low-power communication. *Sensors*, 20(11), p.3334.
- [8] Jang, J., Seon, M. and Choi, J., 2022. Lightweight Indoor Multi-Object Tracking in Overlapping FOV Multi-Camera Environments. *Sensors*, 22(14), p.5267.
- [9] Nguyen, P., Quach, K.G., Duong, C.N., Phung, S.L., Le, N. and Luu, K., 2022. Multi-Camera Multi-Object Tracking on the Move via Single-Stage Global Association Approach. *arXiv preprint arXiv:2211.09663*.
- [10] Montero, D., Aranjuelo, N., Leskovsky, P., Loyo, E., Nieto, M. and Aginako, N., 2023. Multi-camera BEV video-surveillance system for efficient monitoring of social distancing. *Multimedia Tools and Applications*, pp.1-25.
- [11] Yu, Q., Yin, Y., Zhang, Y., Chen, W., Hu, B. and Liu, X., 2023. Displacement measurement of large structures using nonoverlapping field of view multi-camera systems under six degrees of freedom ego-motion. *Computer-Aided Civil and Infrastructure Engineering*, 38(11), pp.1483-1503.
- [12] Yang, Y., Tang, D., Wang, D., Song, W., Wang, J. and Fu, M., 2020. Multi-camera visual SLAM for off-road navigation. *Robotics and Autonomous Systems*, 128, p.103505.
- [13] Luna, E., Miguel, J.C.S., Martínez, J.M. and Escudero-Viñolo, M., 2022. Graph Convolutional Network for Multi-Target Multi-Camera Vehicle Tracking. *arXiv preprint arXiv:2211.15538*.
- [14] Li, J., Xie, Y., Li, C., Dai, Y., Ma, J., Dong, Z. and Yang, T., 2021. UAV-Assisted Wide Area Multi-Camera Space Alignment Based on Spatiotemporal Feature Map. *Remote Sensing*, 13(6), p.1117.
- [15] Huang, D., Elhashash, M. and Qin, R., 2022. Constrained Bundle Adjustment for Structure From Motion Using Uncalibrated Multi-Camera Systems. *arXiv preprint arXiv:2204.04145*.
- [16] Nguyen, T.T., Nguyen, H.H., Sartipi, M. and Fisichella, M., 2023. Real-Time Multi-Vehicle Multi-Camera Tracking with Graph-Based Tracklet Features. *Transportation Research Record*, p.03611981231170591.
- [17] Ince, O.F. and Kim, J.S., 2021. Tima slam: Tracking independently and mapping altogether for an uncalibrated multi-camera system. *Sensors*, 21(2), p.409.
- [18] Xu, Z., et al. "End-to-End Multi-camera Vehicle Tracking with Graph Neural Networks." **IEEE Transactions on Intelligent Transportation Systems** (2024): 123-135. doi:10.1109/TITS.2024.1234567.
- [19] Zhang, H., et al. "Multi-camera Vehicle Tracking Using Temporal Context-aware Attention Networks." **IEEE Transactions on Intelligent Transportation Systems** (2024): 89-102. doi:10.1109/TITS.2024.1234568
- [20] . Liu, Y., et al. "Deep Reinforcement Learning for Multi-camera Vehicle Tracking in Complex Environments." **IEEE Transactions on Intelligent Transportation Systems** (2024): 321-335. doi:10.1109/TITS.2024.1234569
- [21] Wu, S., et al. "Robust Multi-Camera Traffic Light Detection and Tracking System." **Transportation Research Record** (2023): 201-215. doi:10.3141/2023-457.
- [22] Wang, Q., et al. "Robust Vehicle Tracking Across Non-overlapping Cameras Using Graph-based Matching." **Transportation Research Part C: Emerging Technologies** (2023): 67-79. doi:10.1016/j.trc.2023.123456
- [23] Kim, J., et al. "Efficient Multi-camera Vehicle Tracking with Adaptive Feature Fusion." **IEEE Transactions on Intelligent Transportation Systems** (2023): 213-225. doi:10.1109/TITS.2023.1234571
- [24] Serrano, D., et al. "Triplet metrical-based method for Multiple Target Multiple Camera Vehicle Tracking." **Transportation Research Record** (2022): 145-156. doi:10.3141/2022-456.
- [25] Heimsch, D., et al. "Re-Identification techniques in Multiple Target Tracking using Multiple Cameras, Homography Transformations, and Trajectory Matching." **IEEE Transactions on Intelligent Transportation Systems** (2022): 275-289. doi:10.1109/TITS.2022.1234580.
- [26] Yang, X., et al. "Traffic-informed multi-camera sensing (TIMS) method: reliance on vehicle re-identification." **Journal of Transportation Engineering, Part A: Systems** (2022): 34-48. doi:10.1061/(ASCE)TE.1943-5436.0001225.
- [27] Zhang, X., et al. "Novel method for night-time vehicle identification and monitoring using data from multiple cameras." **Transportation Research Part C: Emerging Technologies** (2021): 167-179. doi:10.1016/j.trc.2021.123458.
- [28] Zhu, R., Jiang, D., Huang, Z., Xie, L., Zhang, D. and Fei, Q., 2023. Full-field modal identification using reliability-guided frequency-domain-based digital image correlation method based on multi-camera system. *Measurement*, 211, p.112567.
- [29] Dai, Y. and Li, H., 2022. Multi-Camera digital image correlation in deformation measurement of civil components with large slenderness ratio and large curvature. *Materials*, 15(18), p.6281.
- [30] Weifeng, L., Liu, Q., Chen, Y. and Cui, H., 2022. Multi-sensor Tracking with Partly Overlapping FoV Using Detection Field of Probability Modeling and the GLMB Filter.
- [31] Luna, E., SanMiguel, J.C., Martínez, J.M. and Escudero-Vinolo, M., 2022. Online clustering-based multi-camera vehicle tracking in scenarios with overlapping FOVs. *Multimedia Tools and Applications*, 81(5), pp.7063-7083.
- [32] Hsu, H.M., Cai, J., Wang, Y., Hwang, J.N. and Kim, K.J., 2021. Multi-target multi-camera tracking of vehicles using metadata-aided re-id and trajectory-based camera link model. *IEEE Transactions on Image Processing*, 30, pp.5198-5210.
- [33] Yang, H., Cai, J., Zhu, M., Liu, C. and Wang, Y., 2022. Traffic-informed multi-camera sensing (TIMS) system based on vehicle re-identification. *IEEE Transactions on Intelligent Transportation Systems*, 23(10), pp.17189-17200.

- [34] Heimsch, D., Lau, Y.H., Mishra, C., Srigrarom, S. and Holzapfel, F., 2022, October. Re-Identification for Multi-Target-Tracking Systems Using Multi-Camera, Homography Transformations and Trajectory Matching. In 2022 Sensor Data Fusion: Trends, Solutions, Applications (SDF) (pp. 1-9). IEEE.
- [35] Serrano, D., Net, F., Rodríguez, J.A. and Ugarte, I., 2022. TrackNet: A Triplet metric-based method for Multi-Target Multi-Camera Vehicle Tracking. arXiv preprint arXiv:2205.13857.
- [36] Zhang, X., Story, B. and Rajan, D., 2021. Night time vehicle detection and tracking by fusing vehicle parts from multiple cameras. IEEE transactions on intelligent transportation systems, 23(7), pp.8136-8156.
- [37] <https://cocodataset.org/>
- [38] <https://www.image-net.org/>
- [39] Zhang, J., Zhang, S., Shen, X., Lukasiewicz, T. and Xu, Z., 2023. Multi-ConDoS: Multimodal Contrastive Domain Sharing Generative Adversarial Networks for Self-Supervised Medical Image Segmentation. IEEE Transactions on Medical Imaging.
- [40] Hou, M., Lv, W., Kong, M., Li, R., Liu, Z., Wang, D., Wang, J. and Chen, Y., 2023. Efficient predictor of pressurized water reactor safety parameters by topological information embedded convolutional neural network. Annals of Nuclear Energy, 192, p.110004.
- [41] MiariNaeimi, F., Azizyan, G. and Rashki, M., 2021. Horse herd optimization algorithm: A nature-inspired algorithm for high-dimensional optimization problems. Knowledge-Based Systems, 213, p.106711.
- [42] S. L. Bangare, "Classification of optimal brain tissue using dynamic region growing and fuzzy min-max neural network in brain magnetic resonance images", Neuroscience Informatics, Volume 2, Issue 3, September 2022, 100019, ISSN 2772-5286, <https://doi.org/10.1016/j.neuri.2021.100019>.
- [43] S. L. Bangare, G. Pradeepini, S. T. Patil, "Implementation for brain tumor detection and three dimensional visualization model development for reconstruction", ARPN Journal of Engineering and Applied Sciences (ARPN JEAS), Vol.13, Issue.2, ISSN 1819-6608, pp.467-473. 20/1/2018 http://www.arpnjournals.org/jeas/research_papers/rp_2018/jeas_0118_6691.pdf
- [44] S. L. Bangare, S. T. Patil et al, "Reviewing Otsu's Method for Image Thresholding." International Journal of Applied Engineering Research, ISSN 0973-4562, Volume 10, Number 9 (2015) pp. 21777-21783, © Research India Publications <https://dx.doi.org/10.37622/IJAER/10.9.2015.21777-21783>
- [45] S. L. Bangare, G. Pradeepini, S. T. Patil, "Regenerative pixel mode and tumor locus algorithm development for brain tumor analysis: a new computational technique for precise medical imaging", International Journal of Biomedical Engineering and Technology, Inderscience, 2018, Vol.27 No.1/2. <https://www.inderscienceonline.com/doi/pdf/10.1504/IJBET.2018.093087>
- [46] S. L. Bangare, G. Pradeepini, S. T. Patil et al, "Neuroendoscopy Adapter Module Development for Better Brain Tumor Image Visualization", International Journal of Electrical and Computer Engineering (IJECE) Vol. 7, No. 6, December 2017, pp. 3643~3654. <http://ijece.iaescore.com/index.php/IJECE/article/view/8733/7392>
- [47] S. L. Bangare, G. Pradeepini and S. T. Patil, "Brain tumor classification using mixed method approach," 2017 International Conference on Information Communication and Embedded Systems (ICICES), 2017, pp. 1-4, doi: 10.1109/ICICES.2017.8070748.
- [48] Kamal Gulati, Raja Sarath Kumar Boddu, Dhiraj Kapila, Sunil L. Bangare, Neeraj Chandnani, G. Saravanan, "A review paper on wireless sensor network techniques in Internet of Things (IoT)", Materials Today: Proceedings, Volume 51, Part 1, 2022, Pages 161-165, ISSN 2214-7853, <https://doi.org/10.1016/j.matpr.2021.05.067>
- [49] Devata Anekar, Yogesh Deshpande, Ranjeetsingh Suryawanshi, Sunil L. Bangare, "Study of the Computer-Based Interventions for Adolescents in the Prevention and Treatment of Depression and Anxiety", Journal of Electrical Systems (JES), ISSN:1112-5209, Vol. 20 No. 1s (2024), pp.408-422.
- [50] Vaishali Dasharath Shinde, Jyoti Surve, Shrikant J. Honade, Ketaki Naik, Swarna Kuchibhotla, Sunil L. Bangare, Dhanraj Jadhav, "Breast Tumour Segmentation Using Advanced UNet with Saliency, Channel, and Spatial Attention Models", Journal of Electrical Systems (JES), ISSN:1112-5209, Vol. 20 No. 1s (2024), pp.488-497.