Grenze International Journal of Engineering and Technology, January Issue

Performance Evaluation for Multi-Camera Object Detection with Non-Overlapping Views

Nikita R Shetty[1] and Dr Nilesh J Uke[2]
[1]Research Scholar, Department of Computer Engineering, VIIT Pune-48 Affiliated to Savitribai Phule Pune, University (SPPU)Pune State: Maharashtra Country: India
Email: nikita.221p0069@viit.ac.in
[2]Research Supervisor, Department of Computer Engineering, VIIT Pune-48 Affiliated to Savitribai Phule Pune, University (SPPU)Pune State: Maharashtra Country: India
Email: nilesh.uke@gmail.com

***Abstract*— The proliferation of multi-camera systems in surveillance and monitoring applications demands robust methodologies for evaluating the performance of multiple object detection, especially in scenarios where camera fields of view do not overlap. This review paper systematically observes the key performance evaluation parameters pertinent to multiple object detection in multi-camera setups with non-overlapping fields of view. We explore traditional metrics such as precision, recall, and F1-score, as well as advanced metrics like Multiple Object Tracking Accuracy, Multiple Object Tracking Precision, and the recently proposed multi-view Higher Order Tracking Accuracy. Furthermore, we discuss the challenges unique to non-overlapping FoV scenarios, including object re-identification, temporal synchronization, and spatial calibration. The paper also highlights existing datasets and benchmarks that facilitate the evaluation of such systems. Our comprehensive analysis aims to guide researchers and practitioners in selecting appropriate evaluation metrics and methodologies for developing and assessing multi-camera object detection systems with non-overlapping views.**

***Index Terms*— Multi-camera systems; Non-overlapping fields of view; Multiple object detection; Performance evaluation metrics; Object re-identification.**

I. INTRODUCTION

The deployment of multiple camera systems has become increasingly prevalent in various domains, including security surveillance, traffic monitoring, and autonomous navigation [1]. These systems of-ten operate in environments where individual cameras have non-overlapping fields of view (FoVs), posing significant challenges for consistent and accurate multiple object detection and tracking [13]. Unlike overlapping FoV scenarios, non-overlapping setups lack shared visual information, making it difficult to maintain object identities across different camera views. Consequently, evaluating the performance of object detection systems in such configurations requires specialized metrics and method-ologies that account for these unique challenges.
Traditional object detection systems assume that the cameras may have some overlap in their views, allowing for easier object re-identification and tracking across the overlapping regions [4]. In contrast, when cameras operate in non-overlapping settings, objects transitioning between camera views are at risk of being lost or misidentified. This problem is compounded by the fact that individual cameras may capture varying perspectives, scales, and angles of the same object, further complicating the task of associating the detected objects across different views.

Grenze ID: 01.GIJET.12.1.13_1

For instance, an object might appear differently in terms of shape, colour, and size from one camera to another, making it difficult for the system to reliably identify and track the object across multiple views.
Moreover, the absence of shared visual information across cameras means that traditional object detec-tion metrics, such as precision, recall, and F1-score, may not be sufficient to evaluate the performance of detection algorithms in such environments [8]. These metrics are typically effective when there is a clear, shared ground truth across different detection views. However, in non-overlapping scenarios, specialized metrics and evaluation parameters must be considered to ensure an accurate assessment of the system's capabilities. These might include metrics that focus on the ability to maintain object iden-tities, handle occlusions, and track objects through the transitions between non-overlapping views.
Given the complexities involved in multi-camera systems with non-overlapping FoVs, it is essential to develop a comprehensive framework for performance evaluation that accounts for these challenges [1]. This framework should incorporate both traditional evaluation metrics and those designed specifically to address the unique difficulties posed by non-overlapping fields of view. By doing so, it will be pos-sible to accurately assess the performance of object detection systems in real-world environments, en-suring that they can operate effectively and reliably even in complex multi-camera setups as shown in Figure.1.

II. Related Works

Figure 1. Multi Camera Views

The accuracy of the camera-to-board transformations depends on the detection accuracy and the in-trinsic and extrinsic camera parameters. For determining camera-to-camera transformations, hand-eye calibration is used, with its accuracy depending on the number and variety of the calibration object poses. Consequently, some equations yield less accurate transformations, while others produce better ones [19]. Real-time object detection is a computer vision task that aims at identifying and localizing objects in images or video sequences with acceptable inference speed, frame rate, and accuracy, satis-fying the requirement of the application of interest. Autonomous driving robotics and healthcare moni-toring are some examples in which real-time object detection can play a vital role [17]. Understanding the primary challenges associated with detecting and recognizing objects in blurred and low-quality images is crucial for developing effective object detection and recognition systems. These challenges stem from various factors that degrade image quality, including motion blur and defocus blur. Motion blur results from camera and scene movement during exposure, causing blurred regions in images and hindering object detection [14]. The sequence frame detection accuracy is a frame- level measure that penalizes for fragmentation in the spatial dimension while accounting for number of objects detected, missed detects, false alarms and spatial alignment of system output and ground truth objects. Tracking-based approaches for abandoned object detection often become unreliable in complex surveillance videos due to occlusions, lighting changes and other factors [23]. A mid-level task in computer vision, multiple object tracking grounds high-level tasks such as pose estimation action recognition and be-havior analysis. It has numerous practical applications, such as visual surveillance human computer interaction and virtual reality. These practical requirements have sparked enormous interest. Compared with Single Object Tracking (SOT), which primarily focuses on designing sophisticated appearance models and/or motion models to deal with challenging factors such as scale changes, out-of-plane rota-tions and illumination variations, multiple object tracking [11]. A tractable and realistic detection mod-el that accommodates 3D occlusion by taking into account the Lines of Sights (LoSs) of all objects in the scene with respect to the cameras. In contrast, conventional detection models either neglect the LoSs of the objects or are computationally intractable, leading to poor tracking performance in the presence of occlusions [8]. The intuition used to solve this problem is that association across cameras with spatiotemporally non-overlapping fields of view can be achieved by explicitly modeling the mo-tion of objects, providing constraints for the estimation of inter-camera homographies. We use poly-nomial kinematic models for the motion of objects and under this model an Expectation Maximization algorithm is formulated to estimate the inter-camera homographies and motion parameters [5]. In large metropolitan areas, there is a need for data about vehicle classes that use a particular highway or a street. A classification and counting system like the one proposed here can provide important data for a particular decision making agency. Our system uses a single camera mounted on a pole or other tall structure, looking down on the traffic scene. It can be used for detecting and classifying vehicles in multiple lanes and for any direction of traffic flow [22]. We provide a comprehensive overview of the application based on deep learning technology in multi-object multi-camera tracking tasks. We have classified and summarized the different stages of deep learning-based Multi-Objective Multi-Camera Tracking (MOMCT) algorithms, including object detection, object tracking, vehicle re-identification and multi-object cross-camera tracking. We have aggregated the most commonly benchmark datasets and standard metrics for MOMCT [1]. To describe a statistical framework for joint 3-D object state es-timation and camera calibration, this considers both the geometry and the observation characteristics of the cameras. The framework presented makes use of a proxy state space, called disparity space, which allows for parts of the estimation process to be expressed in linear Gaussian form, thereby ena-bling the use of the Kalman filter [3]. The SOD has been used on a large scale as a pre-processing stage in computer vision applications such as image understanding, object detection and semantic segmenta-tion. Recently, image community spread very rapidly such as Instagram application. The image but also the video material which is basically based on image in every single frame. Image recognition field has become the milestone for all researchers in this field, especially SOD [20]. Artificial intelligence paves the way for computers to think like human. Machine learning makes the way more even by adding training and learning components. The availability of huge dataset and high performance computers lead the light to deep learning concept, which extract automatically features or the factors of variation that distinguishes objects from one another. Among the various data sources which contribute to tera-bytes of big data, video surveillance data is having much social relevance in today's world [15]. They extracted information related to atomic services and their properties from the UML class diagram, and retrieved information related to the composition of services from the UML state-chart diagram for au-tomatic generation of the OWL-S service model ontology. Recently, Alberto and Juan presented a tool to design mechanical components using C++ Object Oriented Mechanics Library (OOML), taking into account the needs and requirements of these emerging technologies [21]. They introduce the most used 3D datasets and the evaluation metrics for both classification and retrieval. They provide an analysis of the existing methods in terms of the utilized viewpoints setting and input modes, backbone network, feature extraction, fusion strategies, and viewpoints selection mechanisms. Towards the end, they summarized the performance of the existing methods on several famous datasets [12]. The system needs to be able to learn both the spatial and colour relationships between non-overlapping cameras. This allows the system to determine if a newly detected object has previously been tracked on another camera, or is a new object. The approach learns these spatial and colour relationships, though unlike previous work it does not require pre-calibration or explicit training periods. Incremental learning of the object's colour variation and movement, allows the accuracy of tracking to increase over time without supervised input [6]. A camera network properly to cover multiple targets while ensuring de-sirable sensing quality is challenging due to the vast amount of camera parameters involved and a lack of widely accepted sensing performance metric. A straightforward metric that captures the coverage and sensing quality of objects for a camera network, based on which we automate the search for the configuration that optimizes the proposed metric [2]. Integrating the automatic extraction of cattle masks from the Segment Anything Model (SAM) with YOLOv8 for detection and segmentation elimi-nates manual labeling and reduces human annotator time and error. A multi-camera tracking system will be constructed by installing four cameras in the top corners of the ranch to enable end-to-end mon-itoring of black cattle for full coverage. Enhancing the detection method with Non-Maximum Suppres-sion (NMS) and an overlapped mask removal technique within the YOLOv8 segmentation framework to improve the accuracy of detecting small cattle far from the camera's field of view [9]. Deep learning methods are effective for MOT problem; there is much room to improve the tracking performance with the power of deep learning by considering its great success in the fields of image classification and recognition. Therefore, it is necessary to summarize and analyses existing deep learning based MOT algorithms to pave the way of study deep learning methods for multi-object tracking [18]. A camera network is performed with cameras presenting both overlapping and non-overlapping Fields-of-Views (FoVs), the task-at-hand has to face constant changes in illumination and back-ground both locally and across cameras without the possibility of reliably calibrating the cameras for position and color. Tar-gets can then appear and be seen from different viewing angles, thus making challenging association and assignment of unique IDs that are robust to frequent entering and exiting of the cameras' FOVs [13]. Design their marker object using basic shapes like cubes or hexagons and obtain their relative positions from the equations of such shapes. In order to estimate the relative poses of the camera set, most people employ a planar chessboard. A circular camera configuration is employed the chessboard is not simultaneously seen by all cameras. Then, it is required to acquire multiple views of the chess-board to find the relative camera poses [4]. Multi-object tracking (MOT) is the task of forming the tra-jectories of foreground objects in video sequences while establishing the correspondence of identical objects across frames; it has played an essential role in many real-world applications, for example, in-telligent transportation video surveillance and human–computer interaction. With the development of deep learning (DL) technologies, DL-based MOT methods have made significant progress in accuracy and efficiency [16]. We provide code for an MTMC tracker that extends a single-camera system that has shown good performance to the multi-camera setting. We hope that the conceptual simplicity of our system will encourage plug-and-play experimentation when new individual components are proposed. All trajectories were manually annotated by five people over a year; using an interface we developed to mark trajectory key points and associate identities across cameras [7]. Evaluate state-of-the-art new object detection models with a two-stage approach, highlighting the best average precision. These mod-els unified the detection and segmentation task. Addressed the traditional question of how many imag-es are needed to train a deep-neural-network model. Experimentation performance into a subset and full dataset. The effect of multi-scale data augmentation technique and the importance of a correct con-figuration of anchor's scale for object location [10].

TABLE I. METRIC DESCRIPTION

Metric	Description	Use Case in Non-Overlapping FoV	Limitation	Reference (with Link)
Precision	Ratio of correct positive detections to total positive detections	Useful for evaluating detection accuracy in each camera	Does not consider object continuity across cameras	"Tracking Objects as Points", 2020
Recall	Ratio of correct positive detections to total actual objects	Measures com-pleteness of detec-tion in each cam-era	Fails to capture cross-camera transitions	"YOLOv3: An Incremental Improvement", 2018
F1-Score	Harmonic mean of Precision and Recall	Balanced evaluation of detection accuracy	Ignores identity preservation across views	"On some properties of the fourth-rank hadronic vacuum polarization tensor and the anomalous magnetic moment of the muon", 2020
MOTA (Multi-Object Track-ing Accuracy)	Considers false positives, missed targets, and ID switches	Measures tracking effectiveness in large scenes	Cannot distinguish cam-era-specific errors	"Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks", 2016
ID Switch Rate	Frequency of object identity changes during tracking	Critical for evaluating cross-camera tracking	Sensitive to re-identification quality	"Performance Measures and a Data Set for Multi-Target, Multi-Camera Tracking", 2016
ReID Accura-cy	Accuracy of matching ob-jects across camera views	Essential in non-overlapping FoV setups	Depends heavily on appearance-based models	"Inferring social structure from continuous-time interaction data", 2018

TABLE II. FEATURES AND STRENGTH

Methodology	Key Feature	Handles Non-Overlapping FoV?	Strengths	Weaknesses	Reference (with Link)
DeepSORT + ReID	Combines object detection with appear-ance-based re-identification	Yes	Real-time track-ing, strong identity preservation	ReID fails un-der major ap-pearance changes	"Simple Online and Real-time Tracking with A Deep Associa-tion Metric", 2017
CenterTrack	Joint detection and tracking using center-based representation	No	Robust to occlu-sion and camera shake	Limited to single-camera settings	"Tracking Objects as Points", 2020
Tracklet Association	Matches short-term tracks across cameras	Yes	Good for low-frame rate setups	Matching errors in dense traffic	"On some properties of the fourth-rank hadronic vacuum polarization tensor and the anomalous magnetic moment of the muon", 2020
Graph-Based ReID	Constructs graphs to associate objects over time	Yes	Global consisten-cy in identity assignment	Computational-ly expensive	"Diverse Imitation Learning via Self-Organizing Generative Models", 2022
3D Pose Estimation	Uses pose to track and identify people across views	Yes	Pose-invariant identity tracking	Requires multi-view calibration	"Temporal Cycle-Consistency Learning", 2019

III. METHODS OF EVALUATION METRICS

A. Traditional Evaluation Metrics

Traditional metrics for evaluating object detection performance include precision, recall, and the F1-score. Precision measures the proportion of properly identified objects among all noticed objects, while recall assesses the proportion of correctly identified objects among all actual objects. The F1-score provides a harmonic mean of precision and recall, offering a stable evaluation metric as shown in Fig-ure.2. These metrics are foundational and widely used due to their simplicity and interpretability. How-ever, in multi-camera systems with non-overlapping FoVs, these metrics may not fully capture the complexities involved in maintaining object identities across different views.

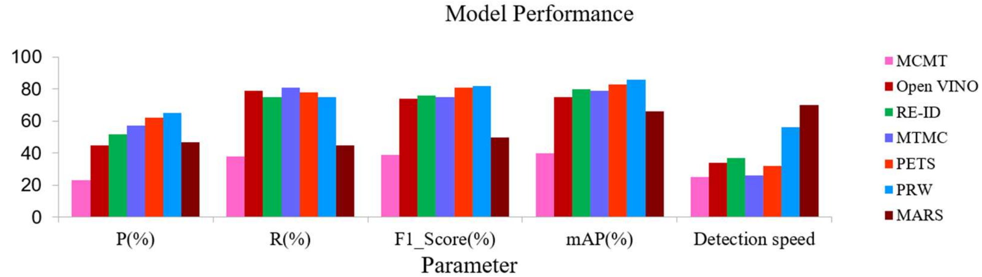

Figure.2 Model Performance

B. Advanced Multi-Object Tracking Metrics

To address the limitations of traditional metrics in multi-camera scenarios, advanced metrics such as Multiple Object Tracking Accuracy (MOTA) and Multiple Object Tracking Precision (MOTP) have been developed. MOTA accounts for false positives, missed targets, and identity switches, providing a comprehensive measure of tracking performance as shown in Figure.3.

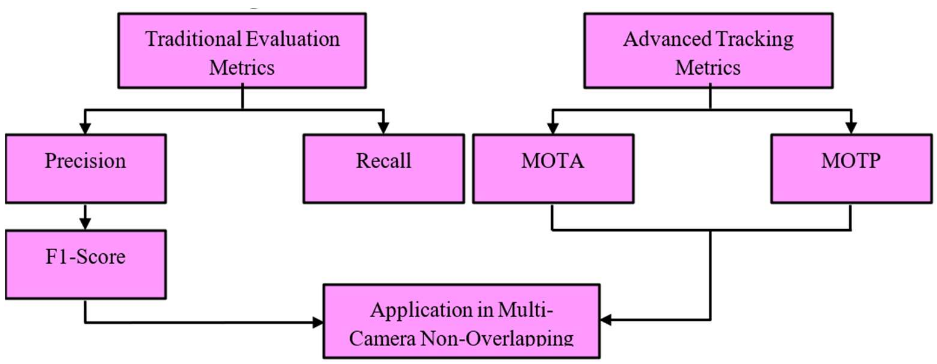

Figure.3 Evaluation Metrics

MOTP evaluates the precision of object localization by measuring the average distance between pre-dicted and ground truth positions. While these metrics offer improved insights, they still face challeng-es in non-overlapping FoV scenarios due to the lack of shared visual information as shown in Figure.4.

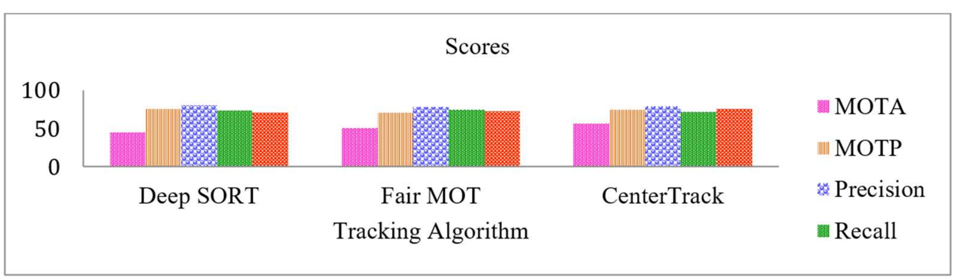

Figure.4 Advanced Parameters

IV. Challenges in Non-Overlapping FoV Scenarios

Non-overlapping FoV configurations introduce specific challenges that complicate performance evalu-ation. One major issue is object re-identification, where the system must determine whether an object observed in one camera is the same as one observed in another. This task is complicated by variations in lighting, viewpoint, and occlusions. Temporal synchronization is another challenge, as discrepancies in time stamps across cameras can lead to misalignment in object tracking. Additionally, spatial cali-bration is critical to accurately map object positions across different camera views. These challenges necessitate the development of specialized evaluation metrics and methodologies.

A. Lack of Visual Continuity

Since cameras cover distinct areas without shared visual regions, objects moving between cameras are not directly observable across transitions. This makes it difficult to maintain identity consistency, lead-ing to identity switches or loss of track.

B. Appearance Variation

Objects may look different in each camera due to changes in angle, lighting, resolution, or occlusion. These variations complicate feature matching and re-identification, especially when deep learning models are not trained on diverse conditions.

C. Temporal Synchronization

Synchronizing video feeds from different cameras is critical. In non-overlapping setups, even slight timing mismatches can prevent the correct association of an object’s trajectory across views.

D. Scene Context Disparity

Each camera has its own unique background and spatial context. Without a unified scene understand-ing, correlating object behaviors or movements across cameras becomes a high-level inference task.

E. Data Association Complexity

Linking object detections from disjoint camera feeds requires complex spatio-temporal models or in-ter-camera handover strategies. These approaches often rely on external metadata (e.g., timestamps, GPS, or trajectory prediction) and are prone to errors in crowded or dynamic environments.

V. Emerging Evaluation Metrics

TABLE III: FEATURES BASED ON OBJECT IDENTIFICATION

Feature Type	Description	Purpose in Multi-Camera Systems	Common Use
Size	Approximate dimensions of detected object (height, width, bounding box area)	Helps in matching object appearance across cameras	Object filtering, re-identification
Color Histogram	RGB or HSV-based histogram represent-ing dominant object colors	Enables colour-based identity match-ing across non-overlapping views	Appearance modeling, identity tracking
Texture	Surface detail patterns (e.g., edge densi-ty, gradient)	Adds robustness to appearance mod-els under varying illumination	Advanced re-identification
Movement Trajectory	Path taken by an object over time, including velocity and direction	Predicts possible re-entry point into another camera view	Path prediction, identity mainte-nance
Time Stamp	Frame or real-time clock information	Synchronizes detection across multi-ple views	Temporal consistency checking
Camera Calibra-tion	Parameters defining the camera’s posi-tion, orientation, and lens distortion	Enables geometric mapping of object position across camera spaces	View transformation
Depth (if available)	Distance from camera to object, using stereo vision or LiDAR	Disambiguates overlapping projec-tions and supports better localization	Object isolation and matching

Recent research has introduced new metrics designed to address the complexities of multi-camera sys-tems with non-overlapping FoVs. One such metric is the multi-view Higher Order Tracking Accuracy (mvHOTA), which extends the traditional HOTA metric by incorporating spatial and temporal associa-tions across multiple views.

To address the challenges of evaluating object detection systems in multi-camera setups with non-overlapping FoVs, the need for new performance parameters and metrics has become critical. Tradi-tional evaluation metrics like precision, recall, and F1-score remain valuable but insufficient in captur-ing the complexities of non-overlapping camera configurations. These metrics fail to account for the ability of detection systems to maintain object identities when objects move between camera views. Consequently, new evaluation parameters such as Multi-Camera Tracking Accuracy (MCTA), Identity Switch Rate (ISR), and Object Association Quality (OAQ) have been proposed. MCTA assesses the ac-curacy of tracking objects as they transition between non-overlapping views, whereas ISR measures the frequency with which objects are misidentified or confused with others. OAQ, on the other hand, quan-tifies the ability of the system to correctly associate detected objects across different cameras based on features such as size, colour, and movement trajectory. These new parameters provide a more detailed understanding of a system's performance in complex environments where objects may not be visible in more than one camera at a time.

Visualizations such as Precision-Recall (PR) curves, Receiver Operating Characteristic (ROC) curves, and Multi-Camera Detection Matrices (MCDM) offer insights into how well detection systems perform across different camera perspectives. PR curves can help assess the trade-off between precision and recall, which is crucial when the system is dealing with variable object visibility in non-overlapping FoVs. ROC curves, on the other hand, provide a comprehensive overview of a detection system's per-formance across different thresholds of detection confidence. Additionally, MCDM plots allow for a more nuanced comparison of detection accuracy across multiple cameras, giving insights into how well the system maintains object identities as objects move between different camera views. These graphical tools are vital for practitioners and researchers to better understand system performance, identify po-tential issues, and refine algorithms accordingly. By integrating such advanced metrics and graphical analysis, the evaluation process can be more closely aligned with real-world challenges, providing a more thorough assessment of multi-camera object detection in non-overlapping FoV environments.

VI. Datasets and Benchmarks

The development and evaluation of multi-camera object detection systems rely heavily on high-quality datasets and benchmarks. Datasets such as the MTMMC (Multi-Target Multi-Camera) dataset provide extensive video sequences captured from multiple cameras in various environments, facilitating the testing of object detection and tracking algorithms. These datasets often include annotations for object identities and positions, enabling the application of various evaluation metrics. However, there remains a need for more datasets specifically designed for non-overlapping FoV scenarios to better support research in this area.

VII. Evaluation Metrics

TABLE IV: PERFORMANCE METRICS

Parameter	Symbol / Metric	Description	Applicable Scenario	Strengths	Limitations
Precision	P	Ratio of correctly detected objects to total detections	Per-camera object detection	Indicates false positive rate	Ignores missed detections
Recall	R	Ratio of correct detected to total ground truth ob-jects	Per-camera object detection	Measures complete-ness of detection	Sensitive to missed detections
F1 Score	F1 = 2PR / (P + R)	Harmonic mean of Preci-sion and Recall	General detection quality	Balances precision and recall	Cannot measure identity tracking
Multi-Object Tracking Accu-racy	MOTA	Considers FP, FN, and identity switches	Multi-camera multi-object tracking	Composite metric for detection and tracking	Not robust to frag-mented tracks
Identity Switch Rate	IDSW	Number of identity switch-es per tracked object	Object tracking across cameras	Critical for identity preservation	Fails under occlu-sion-heavy scenario
Track Frag-mentation	Frag	Frequency of interrupted trajectories	Evaluating trajectory continuity	Evaluates temporal consistency	Does not indicate identity correctness
Re-identification Accuracy	ReID@Rank-k	Top-k matching accuracy across cameras	Cross-camera object matching	Effective in visual similarity assessment	Sensitive to light-ing/camera vari-ance
Tracking Precision	MOTP	Measures bounding box overlap with ground truth	Bounding box accura-cy	Evaluates localization accuracy	Ignores identity correctness

Camera Transi-tion Accuracy	CTA	Accuracy of tracking iden-tity across camera transi-tions	Non-overlapping cam-era systems	Evaluates cross-camera handoff	Rarely reported in standard bench-marks
Object Associa-tion Quality	OAQ	Score for successful and correct object linking	Appearance and mo-tion-based linking	Evaluates system-wide matching accuracy	Requires strong feature extraction
Transition Time Cosistency	TTC	Measures expected vs actual transition time across cameras	Temporal correlation in object tracking	Useful for sequential corridor-based setups	Not applicable in open spaces

VIII. CONCLUSION

In conclusion, the evaluation of multiple object detection systems in multi-camera environments with non-overlapping fields of view presents multifaceted challenges that necessitate more sophisticated and context-aware performance metrics beyond traditional measures such as precision, recall, and the F1-score. While these conventional metrics provide foundational insights into detection accuracy, they fail to capture the complexities of identity preservation, temporal continuity, and spatial re-association inherent to disjoint camera views. Advanced tracking metrics such as MOTA and MOTP have ad-dressed some of these limitations by incorporating errors from missed detections, false positives, and identity switches; however, they still do not fully represent the inter-camera associations critical to evaluating performance in non-overlapping setups. Recent developments, including the multi-view mvHOTA, offer a more comprehensive solution by jointly evaluating detection, tracking, and cross-camera re-identification, reflecting real-world operational challenges. These advancements underscore the importance of using evaluation metrics that are aligned with the practical demands of intelligent surveillance and monitoring systems. Furthermore, the emergence of tailored datasets and benchmark-ing platforms designed for non-overlapping camera networks has enabled more accurate, consistent, and comparative performance assessments. As the field continues to evolve, the integration of graph-based analytics, spatial-temporal modeling, and deep re-identification frameworks into both detection algorithms and their evaluation metrics will be crucial. Ultimately, the ongoing refinement of perfor-mance evaluation parameters will play a pivotal role in the development of robust, scalable, and con-text-aware multi-camera object detection systems suitable for deployment in complex, real-world envi-ronments.

REFERENCES

[1] Lunlin Fei and Bing Han, "Multi-Object Multi-Camera Tracking Based on Deep Learning for Intelligent Transpor-tation: A Review",10 April 2023.

[2] Xuechao Zhang, Xuda Ding, Yi Ren, Yu Zheng, Chongrong Fang and Jianping He, "Toward Global Sensing Quality Maximization: A Configuration Optimization Scheme for Camera Networks", 28 Feb 2023.

[3] Jeremie Houssineau, Daniel Clark, Spela Ivekovic, Chee Sing Lee and Jose Franco, "A unified approach for mul-ti-object triangulation, tracking and camera calibration", 9 Oct 2014.

[4] Hamid Sarmadi, Rafael Munoz-Salinas, M.A. Berbis, and R. Medina-Carnicer, "Simultaneous Multi-View Cam-era Pose Estimation and Object Tracking with Square Planar Markers", 16 March 2021.

[5] Yaser Sheikh,Xin Li and Mubarak Shah,"Trajectory Association across Non-overlapping Moving Cameras in Planar Scenes".

[6] Andrew Gilbert and Richard Bowden, "Tracking Objects Across Cameras by Incrementally Learning Inter-camera Colour Calibration and Patterns of Activity", 2006.

[7] Ergys Ristani, Francesco Solera, Roger Zou, Rita Cucchiara and Carlo Tomasi, "Performance Measures and a Data Set for Multi-target, Multi-camera Tracking", 2016.

[8] Jonah Ong, Ba-Tuong Vo, Ba-Ngu Vo, Du Yong Kim and Sven Nordholm, "A Bayesian Filter for Multi-View 3D Multi-Object Tracking With Occlusion Handling", IEEE Transactions on Pattern Analysis and Machine Intelli-gence, Vol. 44, No. 5, May 2022.

[9] Su Myat Noe, Thi Thi Zin, Ikuo Kobayashi and Pyke Tin, "Optimizing black cattle tracking in complex open ranch environments using YOLOv8 embedded multi-camera system", 2025.

[10] Gendry Alfonso-Francia, Jesus Carlos Pedraza-Ortega, Mariana Badillo-Fernandez, Manuel Toledano-Ayala, Marco Antonio Aceves-Fernandez, Juvenal Rodriguez-Resendiz and Saul Tovar-Arriaga, "Performance Evalua-tion of Different Object Detection Models for the Segmentation of Optical Cups and Discs", 2 December 2022.

[11] Wenhan Luo, Junliang Xing, Anton Milan, Xiaoqin Zhang, Wei Liu and Tae-Kyun Kim, "Multiple object track-ing: A literature review", 23 December 2020.

[12] Mona Alzahrani, Muhammad Usman, Salma Kammoun Jarraya, Saeed Anwar and Tarek Helmy, "Deep models for multi-view 3D object recognition: a review", 12 October 2024.

[13] Senquan Yang, Fan Ding, Pu Li and Songxi Hu, “Distributed multi-camera multi-target association for real-time tracking”, 2022.
[14] Aetesam Ali Khan Ashar, Asir Abrar and Jiangjiang Liu, “A Survey on Object Detection and Recognition for Blurred and Low-Quality Images: Handling, Deblurring, and Reconstruction”, June 24, 2024.
[15] G. Sreenu and M. A. Saleem Durai, “Intelligent video surveillance: a review through deep learning techniques for crowd analysis”, Journal of Big Data, 2016.
[16] Suya Li, Hengyi Ren, Xin Xie and Ying Cao, “A Review of Multi-Object Tracking in Recent Times”, The Insti-tution of Engineering and Technology, 11 February 2025.
[17] Seyed Hozhabr and Roberto Giorgi, “A Survey on Real-Time Object Detection on FPGAs”,IEEE, 21 February 2025.
[18] Yingkun Xu, Xiaolong Zhou, Shengyong Chen and Fenfen Li, “Deep learning for multiple object tracking: a survey”, The Institution of Engineering and Technology, 2019.
[19] Oguz Kedilioglu, Tasnim Tabassum Nova, Martin Landesberger, Lijiu Wang, Michael Hofmann and Sebastian Reitelshofer, “PrIcosa: High-Precision 3D Camera Calibration with Non-Overlapping Field of Views”, 2025.
[20] Kareem Ahmed, Mai A. Gad and Amal Elsayed Aboutabl, “Performance evaluation of salient object detection techniques”, Multimedia Tools and Applications, 2022.
[21] Nilesh J. Uke and Ravindra C. Thool, “Objects tracking in video: a object-oriented approach using Unified Modeling Language”, Int. J. Computational Vision and Robotics, Vol. 5, No. 2, 2015.
[22] Nilesh J. Uke and Ravindra C. Thool,“Moving Vehicle Detection for Measuring Traffic Count Using OpenCV”, 2014.
[23] N. J. Uke and R. C. Thool, “Motion tracking system in video based on extensive feature set”, 2 May 2013.