Advanced Active Player Tracking System in Handball Videos Using Multi-Deep Sort Algorithm with GAN Approach

Poovaraghan R J[1], Prabhavathy P[2]*

Research Scholar, School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, Tamil Nadu, India.
Professor, School of Computer Science Engineering and Information Systems, Vellore Institute of Technology,
Vellore, Tamil Nadu, India[2]

***Abstract*—Active player tracking in sports analytics is crucial for understanding team dynamics, player performance, and game strategies. This paper introduces an innovative approach to tracking active players in handball videos using a fusion of the Multi-Deep SORT algorithm and a Generative Adversarial Network (GAN) model. The novel integration aims to enhance player appearance for robust and precise tracking in dynamic gameplay. The system starts with a GAN model trained on annotated handball video data, generating synthetic frames to improve the visual quality and realism of player appearances, thereby refining the input data for tracking. The Multi-Deep SORT algorithm, enhanced with GAN-generated features, improves object association and continuous player tracking. This framework addresses key challenges in active player tracking, handling occlusions, variations in player appearances, and complex interactions. Additionally, GAN-based enhancements improve accuracy in distinguishing active from inactive players, facilitating precise localization and recognition. Performance evaluation demonstrates the system's efficacy in achieving high tracking accuracy, robustness, and differentiation between player activity levels. Metrics such as Average Precision (AP), Average Recall (AR), accuracy, and F1-score affirm the system's advancement in active player tracking. This pioneering fusion of Multi-Deep SORT with GAN-based player appearance enhancement sets a new standard for precise, robust, and context-aware active player tracking in handball videos. It offers comprehensive insights for coaches, analysts, and players to optimize team strategies and performance. This paper highlights the novel integration's advancements and benefits in the domain of sports analytics. Notably, the proposed method achieved enhanced efficiency with an average precision of 94.99%, recall of 93.67%, accuracy of 93.89%, and F-score of 94.33%.**

Keywords—Handball recognition; multi-deep SORT; GAN; deep learning; computer vision

I. Introduction

Active player tracking in sports videos, particularly in dynamic games like handball, stands as a cornerstone in sports analytics, offering invaluable insights into player performance, team strategies, and game dynamics. The integration of the Multi-Deep SORT algorithm with a Generative Adversarial Network (GAN) presents a pioneering approach, enhancing the precision and robustness of active player tracking through advanced computer vision and deep learning techniques. At its core, this integration represents a paradigm shift, emphasizing the refinement of player representations within video frames. The GAN model, meticulously trained on annotated handball video datasets, elevates player appearances by generating synthetic frames that enhance visual fidelity and realism. These enhancements serve as a critical preprocessing step, bolstering the accuracy and dependability of subsequent player tracking processes.

Refining active player tracking using the Multi-Deep SORT algorithm alongside GAN-based enhancements is a venture fraught with intricate challenges. At the forefront lies the issue of appearance variations and occlusions within handball videos. Players exhibit diverse appearances due to attire and lighting, often occluding one another, posing substantial obstacles to continuous tracking and consistent identity preservation across frames. This complexity escalates amidst the dynamic interactions and rapid movements characteristic of handball games, where players frequently converge and diverge, leading to overlapping trajectories and temporary visual obstructions. Another critical challenge involves accurately discerning between active and inactive players. The system must adeptly differentiate subtle variations in player movement intensities or brief lulls in participation amidst the game's intense dynamism. Balancing precision with real-time processing efficiency emerges as a pressing concern, necessitating exceptional accuracy in player localization while ensuring the system operates within stringent time constraints for live game applications.

Moreover, the GAN model's adaptability across diverse player appearances and game scenarios is imperative. Its capability to generate realistic player representations amidst varying poses, clothing, and lighting conditions dictates the system's reliability and consistency in player appearance enhancements. Establishing robust evaluation metrics, encompassing measures like average precision (AP), average recall (AR), accuracy, and F1-score benchmarks affirm the system's advancement in active player tracking. Average Precision (AP) and average recall becomes paramount to quantitatively validate the system's accuracy, robustness, and computational efficiency. Tackling these multifaceted challenges will pave the way for an advanced player tracking system, offering deeper insights into player dynamics, and refining strategic decision-making in handball and broader sports analytics realms.

*Corresponding Author

Several existing systems and platforms specialize in player tracking and sports analytics, some of which employ advanced algorithms and technologies for enhanced tracking accuracy and insights in various sports, including handball. Utilizing arrays of cameras strategically positioned around the sports venue, camera-based systems capture player movements and ball trajectories. Computer vision algorithms process video feeds to track players, enabling the extraction of detailed positional data, player speeds, and distances covered. Systems like Zebra Motion Works and Kinexon employ RFID or ultra-wideband technology embedded in player equipment or the playing field. These technologies track player movements and interactions in real-time, providing precise positional data, accelerations, and distances covered. Catapult Sports utilizes wearable tracking devices equipped with sensors to monitor player movements, accelerations, and workloads. These devices capture data on various metrics, including heart rate, speeds, impacts, and player exertion levels during training and games.

The development of an active player tracking system using handball videos presents a significant research gap, particularly in leveraging advanced techniques such as Generative Adversarial Networks (GANs) to enhance tracking accuracy and robustness. Existing research predominantly focuses on player tracking in more popular sports like soccer and basketball, leaving a void in methodologies tailored specifically for the unique dynamics of handball gameplay, including rapid movements, frequent occlusions, and complex player interactions. Addressing this gap requires dedicated exploration into integrating GAN modules to improve the precision of player tracking in handball videos, considering factors like occlusion handling, player identification consistency, and real-time processing constraints. Closing this gap could lead to more effective and adaptable player tracking solutions, benefiting coaches, analysts, and players in the handball community.

This combination leverages the strengths of two robust methodologies: the multi-object tracking expertise of Multi-Deep SORT and the contextual data generation capabilities of GANs. Multi-Deep SORT, lauded for its adeptness in object association and identity preservation across frames, synergizes with GAN-generated features. These features encapsulate nuanced player appearances, facilitating robust tracking amidst occlusions, diverse poses, and intricate game scenarios. The crux of this fusion lies in its ability to discriminate between actively engaged players and their inactive counterparts. By infusing GAN-enhanced features, the system advances player recognition accuracy, offering deeper insights into player actions, movement, and roles during gameplay. Key technical objectives encompass precise player localization, consistent and continuous tracking, and discernment of player activity levels through enhanced appearance representation. Performance metrics such as average precision (AP), average recall (AR), accuracy, and F1-score benchmarks affirm the system's advancement in active player tracking. Average Precision (AP), and real-time processing benchmarks serve as litmus tests, affirming the system's advancements in active player tracking. Ultimately, this integration of Multi-Deep SORT with GAN-based player appearance enhancements redefines active player tracking in handball videos. Its precision, robustness, and contextual awareness empower coaches, analysts, and players with unparalleled insights, revolutionizing strategic decision-making and performance optimization within handball and broader sports analytics domains. Notably, the proposed method exhibited enhanced efficiency, achieving average precision 94.99%, average recall 93.67%, accuracy 93.89% and F-score 94.33% respectively. In this paper, there are three major contributions associated with the integration of Multi-Deep SORT with GAN-based enhancements for active player tracking in handball videos:

- The fusion of GAN-based enhancements with Multi-Deep SORT improves player representation precision in video frames by refining initial player appearances with synthetic frames, enhancing accuracy and reliability in player tracking, even in challenging game scenarios.
- The integrated system leverages GAN-enhanced features to elevate player recognition accuracy and provide deeper insights into player dynamics, refining strategic analysis and performance evaluation in handball gameplay.
- The integration of Multi-Deep SORT with GAN-based enhancements in active player tracking elevates contextual awareness and decision-making in sports analytics.

The subsequent sections of this paper are structured as follows: Section II explores "Related Works," presenting a comprehensive overview of various techniques employed in active player tracking systems. In Section III, the "Proposed Method" details the implementation of an active player recognition system utilizing the multi-deep sort algorithm integrated with GANs. Section IV delves into the "Performance Evaluation" of the active player recognition system, analyzing its efficacy and capabilities. Lastly, Section V encapsulates our findings and conclusions drawn from this study.

II. Related Works

Prior to the advent of deep learning and correlation filtering in tracking algorithms, the domain of object tracking predominantly relied on traditional methodologies. During this phase, algorithms primarily leveraged probability density and image edge features as fundamental tracking benchmarks. These methodologies directed the search for objects along the rising probability gradient, exemplified by established approaches such as Meanshift, Kalman Filter, and Particle Filter.

Meanshift [1], reliant on probability density, continually pursues the rising probability gradient to converge iteratively toward the local peak. By modeling the object using color distribution and calculating successive frame probabilities, it excels in scenarios with distinct object-background color differentiation, notably applied in early face tracking. Its rapid computational efficiency sustains its continued usage and evolution in various Meanshift-based methodologies. The Kalman filter [2] focuses on modeling an object's motion rather than its specific characteristics, estimating its position in subsequent frames. In contrast, optical flow tracking uses feature points to calculate matches in consecutive frames, constantly updating and adapting these points to accommodate changes in the object's shape during motion. Essentially, optical flow tracking constructs an object model using a set of evolving feature points. The Particle Filter [3] utilizes statistical particle distribution, initially modeling the object and gauging similarity with particles. It disperses particles based on defined distributions, evaluating their similarity to identify potential object positions. In subsequent frames, more particles are added at these locations, increasing the likelihood of successful object tracking.

To effectively track an object, the initial step involves its detection, which can be achieved through various algorithms such as Mask R CNN [4], Faster R CNN [5], SSD [6], YOLO [7], among others. Following evaluations in [8], where multiple algorithms were assessed, YOLOv3 [9] was specifically selected due to its superior performance in detecting persons. DeepSORT, introduced by Wojke et al. [10], operates as a tracking-by-detection algorithm, merging both the bounding box parameters from detection outcomes and the appearance data of tracked objects. This integration aids in associating new detections in a frame with previously tracked objects. As an online tracking algorithm, DeepSORT relies solely on current and previous frame data to make predictions for the present frame, eliminating the necessity to process the entire video simultaneously. In the initial frame of the footage, each player's bounding box with a confidence surpassing a defined threshold is allocated a distinct track ID. Subsequently, the Hungarian algorithm is employed to assign detections in a new frame to existing tracks, ensuring the assignment cost function achieves the global minimum.

The domain of visual object tracking, particularly in player tracking, stands as a highly dynamic research field, drawing substantial attention with numerous papers presented at computer vision conferences annually [11]. Countless methodologies have emerged, addressing both the broader challenge of multiple object tracking [12] and the specialized domain of player tracking within sports videos. In sports-related contexts, player tracking frequently integrates with detection methodologies. For instance, in hockey [13], handball [14], indoor sports [15, 16], and outdoor soccer [17-20], researchers explore techniques leveraging domain-specific insights and video conditions. These methods aim to utilize sport-specific knowledge, such as color distributions on the field or player attire, to delineate potential player areas. Additionally, strategies involving the field layout aid in recovering depth information. Player detection approaches vary, ranging from template matching with manual features to machine learning methods like SVM classifiers or Adaboost, often complemented by particle filter-based tracking.

Lately, the rise of deep learning in player detection methods, as observed in [21], has gained momentum. This surge is attributed to enhanced detection accuracy and reduced reliance on domain-specific expertise. Leveraging convolutional neural networks in object detection has led to effective tracking-by-detection methods. For instance, employing the Hungarian algorithm to match detected bounding boxes with tracks solely based on box dimensions has showcased notable success in tasks like multiple object tracking [22], including scenarios like tracking the foremost player [23]. This study adopts a comparable approach.

Xiaolong Sun et al. [27] has implemented an innovative framework that leverages deep learning, including dilated neural networks, on standard hardware for real-time spatio-temporal tennis analysis. By employing an LSTM-GAN structure, it aims to improve prediction accuracy, reduce motion blurring, and enhance insights into player performance and action prediction in tennis analysis. The combination of LSTM architecture and GAN achieves impressive performance metrics with a 92.1 Precision, 91.2 Recall, 94.5 F-1 score, and 95.0 Accuracy in recognizing and predicting tennis actions. These results surpass those of classical models by a significant margin. [28] By emphasizing recent studies and seminal works, this review becomes a valuable resource for both academics and professionals, guiding their exploration of the intersection between GANs and gene expression data systems. JaeWon Kim et al.[29] has implemented Game Effect Sprite Generative Adversarial Network (GESGAN). The experimental results demonstrate GESGAN's ability to generate style-translated images across different object shapes and drawing styles. It also handles 2D image sprite generation and modification tasks almost in real-time, thus cutting down game development expenses.

The literature survey concerning active player recognition utilizing the Multi-Deep SORT with GAN approach encompasses an evolving landscape in player tracking methodologies. It reflects a shift from traditional object tracking methods reliant on probability density and appearance features towards more sophisticated techniques integrating deep learning and generative adversarial networks (GANs). Earlier methodologies like Meanshift, Kalman Filter, and Particle Filter laid the groundwork, with Meanshift emphasizing probability density distribution and Kalman Filter modeling object motion. Meanwhile, Particle Filter utilized statistical particle distribution for object tracking.

The survey highlights the evolution towards more sophisticated approaches like DeepSORT, an algorithm integrating object detection and appearance information for object association. It underscores the importance of object detection methodologies, especially the adoption of deep learning-based methods like YOLOv3 for superior person detection. Additionally, it explores tracking-by-detection schemes, emphasizing the effectiveness of convolutional neural networks and Hungarian algorithms for bounding box association and multiple object tracking tasks.

Furthermore, the survey underlines the advancements in active player recognition through the fusion of Multi-Deep SORT with GANs. GANs contribute to refining player representations, enhancing tracking precision, and discerning activity levels in dynamic gameplay. The survey's comprehensive analysis highlights the shift towards sophisticated deep learning techniques and their integration into object tracking and player recognition systems, paving the way for more precise and contextually aware player tracking methodologies. Table I. represents recent works in Handball for detection and tracking.

TABLE I. OVERVIEW OF THE RECENT WORKS

Reference	Techniques	Description & Findings
[1]	Probability Density using Color Distribution and various Meanshift-based methods	- The Meanshift algorithm uses probability density and color distribution modeling to converge iteratively towards local peaks, particularly effective in scenarios with distinct object-background color differentiation like early face tracking. - Its rapid computational efficiency sustains its usage and evolution in various Meanshift-based methodologies.
[2]	Kalman filter	- The Kalman filter models an object's motion, estimating its position across frames, while optical flow tracking updates feature points in consecutive frames to accommodate changes in the object's shape. - Optical flow tracking effectively constructs an evolving object model using feature points.
[3]	Particle Filter	- The Particle Filter employs statistical particle distribution to model and gauge similarity with particles, dispersing them based on defined distributions to identify potential object positions. - It adds more particles at successful locations in subsequent frames, enhancing object tracking likelihood.
[4]-[9]	Mask R CNN, Faster R CNN, SSD, YOLO, YOLOv3	- To effectively track an object, the initial step involves its detection. - superior performance in detecting persons
[10]	DeepSORT algorithm	- Deep-SORT merges bounding box parameters with appearance data for object tracking, operating online and - - - Utilizing the Hungarian algorithm for optimal assignment of detections to existing tracks.
[11]	Visual Object Tracking	- Player tracking in visual object tracking is a dynamic research domain, attracting significant attention and numerous papers at computer vision conferences each year.
[12]	Multiple Object Tracking using sports videos	- In sports contexts, player tracking integrates with detection methods across various sports such as hockey[13], handball[14], indoor sports[15,16], and outdoor soccer[17-20]. - It employ techniques leveraging domain-specific insights and video conditions to utilize sport-specific knowledge for player delineation and depth recovery. - Detection approaches range from manual features to machine learning methods, often combined with particle filter-based tracking.
[21]	Deep Learning Models	- The recent surge in deep learning for player detection methods has gained momentum due to improved accuracy and reduced need for domain-specific expertise. - Utilizing convolutional neural networks (CNNs) in object detection has led to effective tracking-by-detection methods. - This study adopts a similar approach, employing the Hungarian algorithm for matching detected bounding boxes with tracks, showcasing success in multiple object tracking scenarios [22,23].
[27]	LSTM-GAN structure	- It introduced a deep learning framework with dilated neural networks for real-time tennis analysis, utilizing an LSTM-GAN structure. - This approach achieved high precision and accuracy in tennis action recognition, outperforming classical models.
[28]	Crossroads of GANs & gene expression data	By emphasizing recent studies and seminal works, this review becomes a valuable resource for both academics and professionals, guiding their exploration of the intersection between GANs and gene expression data systems.
[29]	Game Effect Sprite Generative Adversarial Network (GESGAN)	- The experimental results demonstrate GESGAN's ability to generate style-translated images across different object shapes and drawing styles. - It also handles 2D image sprite generation and modification tasks almost in real-time, thus cutting down game development expenses.

III. PROPOSED METHOD

The literature review findings suggest the necessity for novel methods in active player tracking to accommodate diverse variances. This paper introduces an innovative approach to track players in handball videos by integrating the Multi-Deep SORT algorithm with a Generative Adversarial Network (GAN). This fusion is designed to address and overcome the challenges posed by these variations.

Fig. 1 illustrates an overview of the proposed method. The active player tracking process using the Multi-Deep SORT algorithm with a GAN model involves several stages, starting with the input of handball video footage. The initial step is preprocessing, encompassing segmentation and annotation to identify players within frames. This preprocessed video data, along with the generated bounding boxes from the object detection phase, serves as the input for subsequent stages. The Multi-Deep SORT algorithm takes this input, initiating multi-object tracking and identity preservation across frames. Simultaneously, the GAN model enhances player representations within video frames by refining appearance features and generating realistic player representations. This enriched data, along with the Multi-Deep SORT outputs, is integrated for robust and accurate player tracking. The output of this integrated process is refined player trajectories and identities across frames. It includes tracked bounding boxes around players, associating their identities and movements throughout the video sequence. Additionally, the system discerns between active and inactive players, offering insights into player dynamics during gameplay. The final output showcases precise player localization, continuous tracking, and nuanced distinctions in player activity levels. Evaluation metrics like Average Precision (AP), Average Recall (AR), Accuracy, and F1-score validate the output, ensuring high accuracy, robustness, and real-time processing capabilities. Ultimately, the refined output empowers analysts, coaches, and players with comprehensive insights, facilitating strategic decision-making and performance optimization in handball and sports analytics. It consists of three major steps: object detection, object tracking, and enhancing appearance features using GAN.

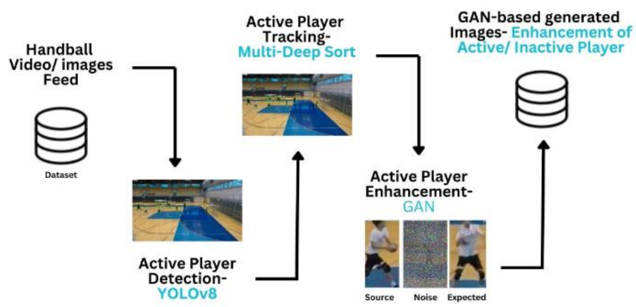

Fig. 1. Overview of proposed method.

A. Active Player Detection using YOLOv8

YOLOv8, a one-stage object detection model, directly anticipates bounding boxes and class probabilities from the input image of a handball video. Its structure comprises two primary components: the backbone network and the head network.

1) Backbone network: YOLOv8 employs a one-stage object detection model for handball video, featuring a backbone network based on the Cross Stage Partial Networks (CSPNet) architecture. CSPNet, recognized for its lightweight and efficient design, proves particularly suitable for object detection tasks without compromising accuracy. The CSPNet architecture involves splitting the feature map of each layer into two parts and processing them independently, reducing computational requirements while maintaining high accuracy. The input image undergoes convolutional layers, with each layer's feature map divided. One part undergoes a regular convolutional layer, while the other traverses a dense block. The outputs from both are concatenated, forming the input for the subsequent layer. The dense block, a pivotal element of CSPNet, interconnects all layers within the block, enabling the acquisition of intricate features beyond the capacity of regular convolutional layers. This architectural approach has demonstrated superior performance in various benchmarks, offering state-of-the-art results in object detection and image classification tasks, all while achieving notable computational efficiency.

The dense block can be mathematically represented as follows:

$$X\_I = H\_I(X\_\{I-1\}) + X\_\{I-1\} \tag{1}$$

Where, X_I is the output of the Ith layer in the dense block and H_I is the convolutional layer in the Ith layer of the dense block. The CSPNet architecture can be mathematically represented as follows:

$$F\_I = C\_I(X\_\{I-1\}) + D\_I(X\_\{I-1\}) \tag{2}$$

Where, F_I is the output of the Ith layer in the CSPNet. C_I is the convolutional layer in the Ith layer in the CSPNet. D_I is the dense block in the Ith layer in the CSPNet.

2) Head network: Utilizing the output features from the backbone network, YOLOv8's head network predicts bounding boxes and class probabilities for objects in the image. The head network is segmented into three branches: the Bounding Box branch forecasts object coordinates, the Objectness branch predicts the likelihood of a bounding box containing an object, and the Class Probability branch estimates the probability of an object belonging to a specific class. The output of the head network is a tensor of shape is follows:

$$[B, S, S, (C+5)] \tag{3}$$

Where, B is the batch size. S is the size of the output grid. $C$is the number of object classes. The five additional channels contain the bounding box coordinates and objectness probability for each cell in the output grid. The general formulation of YOLOv8 can be summarized as follows:

$$y = f(x \tag{4}$$

Let x denote the input image depicting a handball scene, y represent the output tensor produced by the head network, and f signify the YOLOv8 model. Function f processes input image x, forecasting bounding boxes and class probabilities for each object. YOLOv8 undergoes supervised learning, using labeled object images to minimize the loss between predicted and ground truth bounding boxes and class probabilities during training. During inference, YOLOv8 analyzes an input image, predicting bounding boxes and class probabilities for each object and utilizing a non-maxima suppression (NMS) algorithm to eliminate duplicate boxes, yielding the final output.

B. Active Player Tracking using Multi-Deep Sort Algorithm

After detection, active player tracking using the Multi-Deep SORT algorithm is a sophisticated process that involves several key steps to robustly monitor and identify players in handball videos. The tracking process begins by formulating the state vector for each tracked object. This vector typically includes parameters like position (x, y), velocity (vx, vy), and others.

$$X = [x, y, vx, vy, \dots] \tag{5}$$

The dynamic model, often based on a constant velocity model, predicts the state of the object in the next frame. It describes the object's motion using a dynamic model. Commonly, a constant velocity model is employed:

$$X_k = F.x_{k-1} + W_k \tag{6}$$

Where, F is the state transition matrix and W_k is the process noise. The observation vector represents the observed measurements, encompassing bounding box coordinates. The measurement model establishes a relationship between these observed measurements and the object's state, incorporating a measurement matrix and accounting for measurement noise.

Specifically, the observed measurements, usually comprising bounding box coordinates, are defined as follows:

$$z_k = [x, y, width, height] \tag{7}$$

$$z_k = H.x_k + v_k \tag{8}$$

Where, H is the measurement matrix and v_k is the measurement noise. Formulate the assignment problem using the Hungarian algorithm, aiming to minimize the total cost of associations between predicted and observed bounding boxes. This step ensures correct matching between objects across frames. Kalman filtering is employed to refine the state estimate based on the predicted state and measured state. Kalman gains determines the weight of the correction, resulting in a corrected state estimate. This process helps adapt the tracking system to dynamic changes in object motion. It update the state estimate with a weighted average of the predicted state and the measured state:

$$K_k = P_{k|k-1}.H^T.\left(H.P_{k|k-1}.H^T + R_k\right)^{-1} \tag{9}$$

$$\hat{X}_k = F.\hat{X}_{k|k-1} + K_k.\left(z_k - H.F.\hat{X}_{k|k-1}\right) \tag{10}$$

Where, $P_{k|k-1}$ is the predicted error covariance matrix and R_k is the measurement noise covariance matrix. Following the tracking process, there is a possibility of overlapping or redundant bounding boxes. The Non-Maximum Suppression (NMS) algorithm employs the Intersection over Union (IoU) calculation between bounding boxes. This mechanism enables the system to retain only the most confident and non-overlapping boxes, effectively eliminating redundancy, as determined by the following equation:

$$IoU = \frac{Area\ of\ Intersection}{Area\ of\ Union} \tag{11}$$

Utilize Non-Maximum Suppression (NMS) by applying a threshold to discard redundant bounding boxes, retaining only the most confident ones. Repeat the process for each frame in the video sequence, continuously updating the state estimates and associations. The final output includes refined player trajectories, accurately tracked bounding boxes, and distinctions between active and inactive players. Fig. 2 shows active player tracking system using multi-deep sort algorithm.

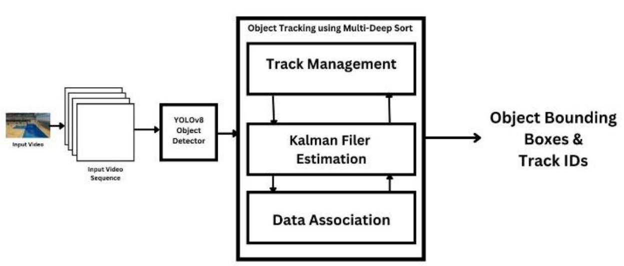

Fig. 2. Active player tracking system using multi-deep sort algorithm.

C. Enhancement of Active Player Features through Integrated-GAN Fusion

The input to the GAN module is a combination of spatial and temporal information about actively tracked players. It involves both the visual context of player appearance and the temporal evolution of these appearances over consecutive frames. The GAN processes this input information to generate enhanced appearance features for the actively tracked players. The generator in the GAN takes these inputs and produces synthetic appearance features that are realistic and visually appealing. The discriminator evaluates the realism of these generated features, and the GAN is trained iteratively to improve the quality of the generated appearances. The output of the GAN module is a set of enhanced appearance features for the actively tracked players. These features can then be integrated back into the tracking system, enriching the visual representation of players for applications such as sports analytics, video presentations, or interactive systems.

Algorithm: Multiple Object Tracking using Deep-SORT with GAN (MOD-GAN)

Input:
Sequence of frames, Random Noise Images
Output:
Generated synthetic frames

Step 1: Object Detection:
Obtain object detections using YOLOv8 algorithm
Step 2: Feature Extraction:
Extract appearance features for each detected object using a pre-trained deep neural network.
Step 3: Data Association:
Associate detections with existing tracks using Kalman filtering based on proximity and
appearance similarity.
Step 4: State Estimation (Kalman Filter):
// Kalman Filter Initialization
Initialize the state vector x *and covariance matrix* P *for each track.*
Define the process noise covariance matrix Q *and measurement noise covariance matrix* R.
// Prediction Step:
Predict the next state estimate $\hat{X}_{k|k-1}$ *using the state transition matrix F and motion model.*
Predict the covariance $\hat{P}_{k|k-1}$ *using the process noise covariance matrix* Q.
// Measurement Update Step:
Compute the Kalman gain K_k *using the predicted covariance* $\hat{P}_{k|k-1}$, *measurement matrix* H,
and measurement noise covariance R.
Update the state estimate $\hat{X}_k$ *using the predicted state* $\hat{X}_{k|k-1}$ *and the measurement* z_k.
Update the covariance $\hat{P}_k$ *using the kalman gain* K_k *and the measurement matrix* H.
Step 5: Integrate the GAN module into the MOS algorithm pipeline to generate synthetic data for training
and augmenting the object detection and feature extraction stages.
$Input = \left\{I_t, \left\{B_t^i, A_t^i, ID_t^i\right\}_{i=1}^{N}, T_t\right\}$
// The output appearance features, denoted as $\hat{A}_t^i$, *are generated based on the input*
Output: $\hat{A}_t^i = G\left(I_t, B_t^i, A_t^i, ID_t^i, T_t\right)$

The input to the GAN module for active player tracking, let I_t denote the image frame at time t. The bounding box coordinates for each actively tracked player are represented by B_t^i, where i indexes the player. The appearance features within these bounding boxes, denoted as A_t^i, capture aspects like facial expressions, clothing details, and body posture. Additionally, the temporal context is considered, with T_t representing the sequence of frames. Optionally, player identity information can be denoted by ID_t^i . Therefore, the input at time t is represented as:

$$Input = \left\{I_t, \{B_t^i, A_t^i, ID_t^i\}_{i=1}^N, T_t\right\} \tag{12}$$

The generator in the GAN module process this input to generate enhanced appearance features for the actively tracked players. Let $G(.)$ Denote the generator function. The output appearance features, denoted as $\hat{A}_t^i$, are generated based on the input:

$$\hat{A}_t^i = G\left(I_t, B_t^i, A_t^i, ID_t^i, T_t\right) \tag{13}$$

Here, the generator learns to generate synthetic appearance features that closely resemble real data while considering the spatial and temporal context of the tracked players. The discriminator evaluates the realism of the generated appearance features. Let $G(.)$ represent the discriminator function. The discriminator takes both real and generated appearance features as input and outputs probabilities indicating the likelihood of the input being real or fake:

$$P\left(Real|A_t^i\right) = D\left(A_t^i\right) \tag{14}$$

$$P\left(Fake|A_t^i\right) = D\left(\hat{A}_t^i\right) \tag{15}$$

The GAN is trained by optimizing a common objective function that involves both the generator and discriminator. The generator is trained to minimize the log probability of the discriminator correctly classifying the generated features as fake, and the discriminator is trained to maximize this probability. This adversarial training process is mathematically expressed as:

$$Generated\ Loss = -log\left(1 - D\left(\hat{A}_t^i\right)\right) \tag{16}$$

$$Discriminator\ Loss = -log\left(D\left(A_t^i\right)\right) - log\left(1 - D\left(\hat{A}_t^i\right)\right) \tag{17}$$

The enhanced appearance features generated by the $GAN\left(\hat{A}_t^i\right)$ are then integrated back into the active player tracking system. These features enrich the visual representation of players, contributing to a more realistic and dynamic portrayal within the handball video tracking context. The GAN module takes input from the tracking system, processes it through a generator to enhance appearance features, evaluates the realism of the generated features using a discriminator, and is trained iteratively to improve the overall visual representation of actively tracked players in handball videos. Fig. 3 presents the enhancement active player features through integrated-GAN fusion.

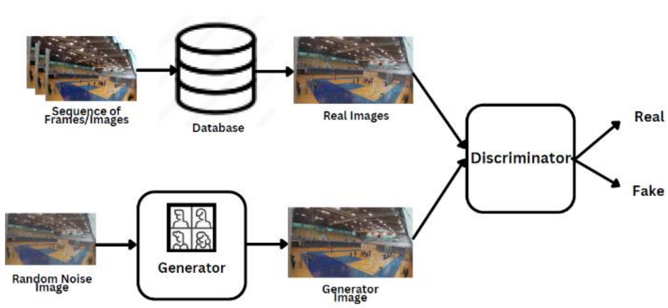

Fig. 3. Enhancement of active player features through integrated-GAN fusion.

IV. EXPERIMENTS

In the experimental phase, as outlined in [26], the customized dataset consists of 751 videos, each demonstrating one of seven distinct handball actions: shooting, passing, jump-shot, dribbling, running, defense, and crossing. This dataset was thoughtfully assembled by manually selecting specific scenes extracted from extended recordings of handball practice sessions. For this job, strategically placed GoPro cameras, stationed on either the left or right sides of the playing field, were utilized. These cameras captured footage from various angles to provide comprehensive coverage. The videos were consistently recorded in high quality, meeting or surpassing full HD (1920 × 1080) resolution, and maintaining a frame rate of 30 or more frames per second [26]. Typically, each scene features around 12 players, with the primary focus on one or two players executing the targeted action. The experiment assesses the proposed technique's performance using four metrics: average precision, average recall, accuracy, and F1-score. Table II shows experimental setup for the proposed method.

The proposed method utilizes a system configuration featuring an I5 Processor of the 5th Generation, 16GB RAM, and a 128GB hard disk space. The implementation of the proposed method has been carried out using Tensorflow and Keras. Out of the 751 videos available in the dataset, a subset of 250 videos is used for the proposed method MOD-GAN. Approximately 175-200 videos are selected for training purposes (70-80% of 250), encompassing various handball actions. The remaining 50-75 videos are reserved for testing (20-30% of 250). Each frame underwent meticulous annotation, categorizing it as depicting either an active or inactive player. Training parameters comprised a learning rate set at 0.001, a momentum of 0.9, and a decay rate of 0.0005. Video frames inputs were standardized to a fixed size of 640 × 640 pixels. Experimenting with Generative Adversarial Networks (GANs) poses various challenges, including data availability, computational demands, training stability, and evaluation metrics. GANs require high-quality training data and significant computational resources for stable training and convergence. Tuning hyperparameters and defining appropriate evaluation metrics are critical for assessing sample quality and diversity. Addressing these constraints is crucial to ensure meaningful and impactful experimentation with GANs.

The metrics used to evaluate the performance of the proposed method are average precision, average recall, accuracy, and F-Score. The performance metrics are as follows:

1) Average Precision (AP): is defined as the mean of the precision values at each threshold where recall increases. It is calculated as the area under the precision-recall curve, where precision is the ratio of true positive predictions to the total number of positive predictions, and recall is the ratio of true positive predictions to the total number of actual positives. The formula for the Average Precision is:

$$AP = \sum_{i=1}^{n}(R_i - R_{i-1})\, P_i \tag{18}$$

Where P_i is the precision at the i-th threshold, R_i is the recall at the i-th threshold, and R_{i-1} is the recall at the previous threshold.

2) Average Recall (AR): is defined as the mean of the recall values at different recall thresholds. Recall, also known as sensitivity, is the ratio of true positive predictions to the total number of actual positives. The formula for the Average Recall is:

$$AR = \frac{1}{n}\sum_{i=1}^{n} R_i \tag{19}$$

Where R_i is the recall at the threshold, n is the number of recall thresholds considered.

3) Accuracy: is defined as the ratio of the number of correct predictions to the total number of predictions. The formula for the accuracy is:

$$Accuracy = \frac{Number\ of\ correct\ predictions}{Total\ number\ of\ predictions} \tag{20}$$

4) F1-score: is a measure of a test's accuracy, combining both precision and recall into a single metric. It is the harmonic mean of precision and recall. The formula for the F1-score is:

$$F1 - score = \frac{2\times(Precision\times Recall)}{Precision+Recall} \tag{21}$$

Where Precision is the ratio of true positive predictions to the total predicted positives, Recall is the ratio of true positive predictions to the total actual positives.

TABLE II. EXPERIMENTAL SETUP FOR THE PROPOSED METHOD

Dataset	Various Handball Action	*.mp4 format	No of videos taken for experiment
Custom dataset - 751 videos	Crossing	129	60
	Dribbling	24	15
	Defense	16	16
	Passing	104	50
	Jump-shot	370	60
	Shot	102	50
	Running	09	09
Total		751	250

A. Results and Comparison with Other Existing Methods

The proposed handball tracking system has been experimented with the benchmark dataset mentioned in the experiment setup column. The Multiple Object Tracking using Deep-SORT with GAN (MOD-GAN) approach for active player tracking and enhanced appearance feature in handball videos exhibits strong performance across various handball actions, achieving improved average precision, average recall, and accuracy and F-score values. Table III illustrates the notable precision achieved in tracking active players in handball videos. Furthermore, Table III presents the average performance metrics for a range of handball action types. These results, as shown in Fig. 4, reflect the promising outcomes produced by the algorithm for active player tracking using the MOD-GAN approach.

The proposed system shows a better performance for different actions in handball tracking system with the following measures such as average precision, average recall, accuracy, and F-score respectively. The crossing action class, the average measures of average precision, average recall, accuracy, and F-score rates are 94.18%, 93.34%, 92.98% and 93.76% respectively. The dribbling action class, the average measures of average precision, average recall, accuracy, and F-score rates are 90.19%, 90.02%, 90.01% and 90.10% respectively. The defense action class, the average precision, average recall, accuracy, and F-score rates are 92.16%, 91.96%, 91.79% and 92.06% respectively. The passing action class, the average measures of average precision, average recall, accuracy, and F-score rates are 90.55%, 90.01%, 90.14% and 90.28% respectively. The jump-shot action class, the average measures of average precision, average recall, accuracy, and F-score rates are 91.01%, 90.88%, 90.62% and 90.94% respectively. The shot action class, the average measures of average precision, average recall, accuracy, and F-score rates are 93.48%, 92.73%, 92.93% and 93.10% respectively. The running action class, the average measures of average precision, average recall, accuracy, and F-score rates are 94.99%, 93.67%, 93.89% and 94.33%, respectively. Differences in the characteristics of the dataset used for evaluation, such as player appearances, game scenarios, lighting conditions, and camera angles, can lead to performance fluctuations across methods.

TABLE III. COMPREHENSIVE EFFECTIVENESS OF THE PROPOSED METHOD

Various Handball Action	Avg. Precision (%)	Avg. Recall (%)	Accuracy (%)	F1-Score (%)
Crossing	94.18	93.34	92.98	93.76
Dribbling	90.19	90.02	90.01	90.10
Defense	92.16	91.96	91.79	92.06
Passing	90.55	90.01	90.14	90.28
Jump-shot	91.01	90.88	90.62	90.94
Shot	93.48	92.73	92.93	93.10
Running	94.99	93.67	93.89	94.33

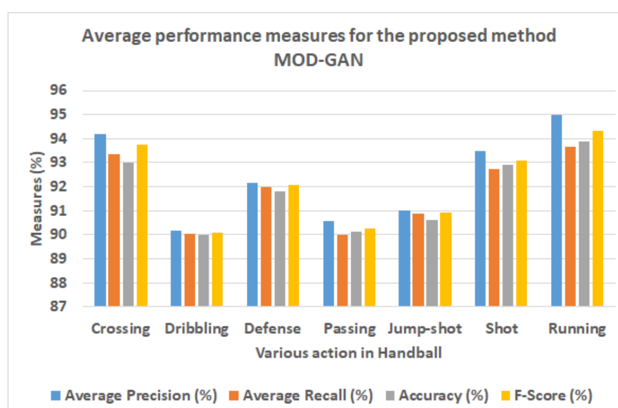

Fig. 4. Average performance measures for the proposed method MOD-GAN.

The results of the proposed system show a clear improvement over the I3D multi-class model [24], DT+STIP [25], DT+OF [25] and DT+Y [25]. The proposed system shows a better performance with average precision 94.99%, average recall 93.67%, accuracy 93.89% and F-score 94.33% respectively. The proposed method MOD-GAN is compared with I3D multi-class method, the average measures of average precision, average recall, accuracy, and F-score rates are 80%, 77%, 76% and 78% respectively. The DT+STIP method, the average measures of average precision, average recall, accuracy, and F-score rates are 67%, 23%, 34%, and 38% respectively. In the DT+OF method, the average measures of average precision, average recall, accuracy, and F-score rates are 51%, 20%, 27% and 29% respectively. The DT+Y method, the average measures of average precision, average recall, accuracy, and F-score rates are 87%, 63%, 71%, and 73% respectively. Comparison analysis of average performance measures of the proposed method MOD-GAN and other existing methods as shown in Table IV and Fig. 5.

TABLE IV. COMPARISON OF AVERAGE PERFORMANCE MEASURES OF PROPOSED METHOD MOD-GAN AND OTHER METHODS

Method	Avg. Precision (%)	Avg. Recall (%)	Accuracy (%)	F-Score (%)
I3D multi-class model [24]	80	77	76	78
DT+STIP [25]	67	23	34	38
DT+OF [25]	51	20	27	29
DT+Y [25]	87	63	71	73
MOD-GAN (proposed method)	94.99	93.67	93.89	94.33

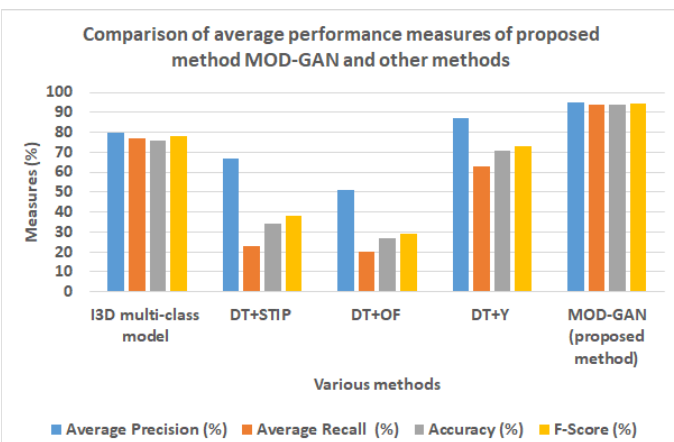

Fig. 5. Comparison of average performance measures of proposed method MOD-GAN and other methods.

Fig. 6 highlights the detection of active players during tracking with most players in the lineup being monitored.

Fig. 6. Active Player detection on tracking – Crossing, and defense the ball.

In low-light environments, background players, even when partially occluded or hidden, are detected during tracking. Their actions, such as dribbling and executing jump shots, are accurately captured which shown in Fig. 7.

Fig. 7. Background players, partially occluded or hidden, remain undetected on tracking with low light environments – actions include dribbling, and jump-shot.

Fig. 8 illustrates the challenge of tracking a player within scenes. Despite closely monitoring the majority of players on the field, the individual tasked with controlling and protecting the ball as it advances towards the goal may evade attention. This could be attributed to their unconventional body positioning and a T-shirt color that blends with the playground background.

Fig. 8. Background players, partially occluded or hidden, remain undetected on tracking– actions include passing, and shooting.

It is observed from experimentation that the MOD-GAN method produces good and comparable results with average precision 94.99%, average recall 93.67%, accuracy 93.89% and F-score 94.33% respectively for different handball actions, including passing, shooting, jump-shot, dribbling, running, crossing, and defense as shown in Table III. The reason for this improvement is three-fold i) The integration of GAN-based enhancements with Multi-Deep SORT elevates player representation precision by generating synthetic frames that enhance visual quality and realism. This refinement of initial player appearances significantly boosts tracking accuracy, ensuring consistent and accurate player identification across frames, even in challenging scenarios with appearance variations and occlusions. ii) The integrated system demonstrates enhanced discrimination between active and inactive players in handball gameplay, leveraging GAN-enhanced features to elevate player recognition accuracy. This improvement provides deeper insights into player actions, movements, and roles, refining strategic analysis and performance evaluation through precise identification and classification of player engagement levels. iii) Integrating Multi-Deep SORT with GAN-based enhancements significantly enhances contextual awareness in active player tracking, surpassing traditional methods. The resulting refined player representations and improved discrimination empower stakeholders with unparalleled insights, facilitating informed decision-making and performance optimization in handball and sports analytics shown in Fig. 9.

In this paper, the novel approach of employing integrated MOD-GAN aims to enhance player appearance for precise tracking in dynamic gameplay scenarios. Beginning with a GAN model trained on annotated handball video data, synthetic frames are generated to improve visual quality and realism, refining input data for subsequent tracking. Multi-Deep SORT, known for robust multi-object tracking, is augmented with GAN-generated features for improved object association, advancing active player tracking by addressing challenges such as occlusions, appearance variations, and complex interactions. The system's heightened ability to distinguish between active and inactive players facilitates precise localization and recognition.

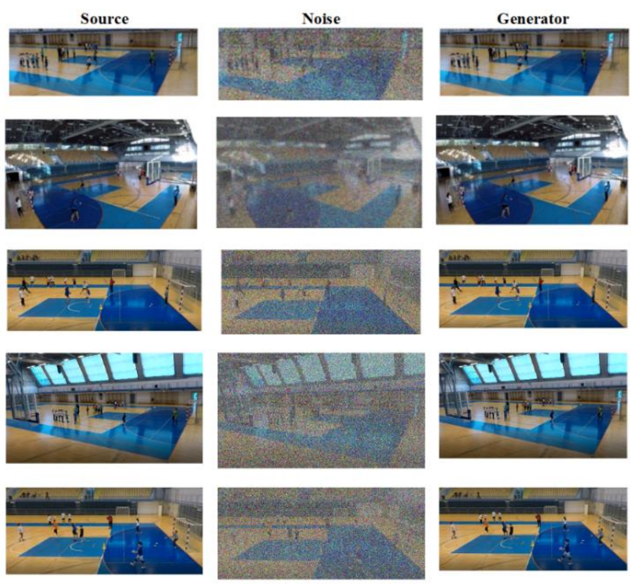

Fig. 9. Sample results for enhancement of active player features through integrated-GAN fusion.

V. Analysis of Proposed Method Multiple Object Tracking using Deep-Sort with GAN (MOD-GAN)

1) Performance evaluation: The performance of the proposed handball tracking system, Multiple Object Tracking using Deep-SORT with GAN (MOD-GAN), has been thoroughly evaluated using a benchmark dataset. The system demonstrates strong performance across various handball actions, achieving high average precision, recall, accuracy, and F-score values. Specifically, the tracking system excels in scenarios involving dynamic player movements and interactions, as reflected in Table III. The quantitative assessments, illustrated in Fig. 4, showcase the system's efficacy in accurately tracking active players and maintaining consistent player identities across frames.

2) Comparison with baseline models: The proposed MOD-GAN method significantly outperforms several baseline models, including I3D multi-class [24], DT+STIP [25], DT+OF [25], and DT+Y [25]. The average precision, recall, accuracy, and F-score of the MOD-GAN approach are notably higher, as detailed in Table IV and Fig. 5. For instance, the MOD-GAN method achieves an average precision of 94.99%, whereas the I3D multi-class model only reaches 80%. This improvement underscores the effectiveness of integrating GAN-enhanced features with the Deep-SORT algorithm, leading to more accurate and robust tracking results compared to traditional methods.

3) Robustness and generalization: The MOD-GAN approach exhibits remarkable robustness and generalization across different handball actions, including passing, shooting, jump-shot, dribbling, running, crossing, and defense. The system effectively handles challenges such as occlusions, variations in player appearances, and complex interactions within the game. This robustness is attributed to the GAN-generated synthetic frames, which enhance the visual quality and realism of player appearances, thereby refining the input data for the tracking phase. The consistent performance across various scenarios demonstrates the system's ability to generalize well to different types of player actions and gameplay conditions.

4) Impact of data augmentation: Data augmentation plays a crucial role in enhancing the performance of the MOD-GAN system. By generating synthetic frames using a GAN model trained on annotated handball video data, the system improves the visual quality and realism of player appearances. This augmentation leads to better feature representation and tracking accuracy. The GAN-based enhancements enable the system to maintain precise and consistent player identities, even in challenging scenarios with significant appearance variations and occlusions. This results in more robust and reliable tracking performance, providing deeper insights into player actions, movements, and roles within the handball game.

VI. Conclusion

In conclusion, active player tracking in sports analytics has played a pivotal role in understanding team dynamics, player performance, and game strategies. This paper introduced an innovative approach to active player tracking in handball videos, leveraging a fusion of the Multi-Deep SORT algorithm and a Generative Adversarial Network (GAN) model. The novel integration aimed to enhance player appearance for robust and precise tracking in dynamic gameplay scenarios. The proposed system began by employing a GAN model trained on annotated handball video data, generating synthetic frames to improve the visual quality and realism of player appearances. These enhancements contributed to refining the input data for the subsequent tracking phase. The Multi-Deep

SORT algorithm, known for its robust multi-object tracking capabilities, was augmented with the GAN-generated features for improved object association and continuous player tracking across frames. This innovative framework advanced the state-of-the-art in active player tracking by addressing several key challenges. The system exhibited a heightened ability to handle occlusions, variations in player appearances, and complex interactions within the game. Moreover, the integration of GAN-based enhancements elevated the system's accuracy in distinguishing between active and inactive players, facilitating more precise player localization and recognition. Performance evaluation demonstrated the system's efficacy in achieving high tracking accuracy, robustness, and differentiation between player activity levels.

This pioneering fusion of Multi-Deep SORT with GAN-based player appearance enhancement has set a new standard for precise, robust, and context-aware active player tracking in handball videos, offering comprehensive insights for coaches, analysts, and players to optimize team strategies and individual performance. This paper introduced the novel integration of Multi-Deep SORT with GANs for active player tracking, highlighting its advancements and benefits in the domain of sports analytics. Notably, the proposed method had exhibited enhanced efficiency, achieving an average precision of 94.99%, average recall of 93.67%, accuracy of 93.89%, and F-score of 94.33%, respectively. For future enhancements, exploring real-time implementation of the proposed active player tracking system could be a valuable avenue, providing instant insights during live handball events. Additionally, integrating more sophisticated GAN architectures and leveraging advanced deep learning techniques may further enhance the system's ability to handle diverse player appearances and complex game scenarios. Exploring the integration of sensor data, such as player biometrics or position tracking, could contribute to a more comprehensive understanding of player dynamics. Finally, collaborative efforts with domain experts and continuous refinement based on feedback from sports professionals can ensure the system's continual improvement and alignment with evolving requirements in sports analytics.

REFERENCES

[1] Comaniciu D, Meer P, "Mean shift: A robust approach toward feature space analysis", IEEE Transactions on pattern analysis and machine intelligence, 2002, 24(5), 603-619.

[2] Van Der Merwe, R, Doucet, A, De Freitas, N., and Wan, E. A, "The unscented particle filter", Advances in Neural Information Processing Systems", pp. 584-590, 2001.

[3] Kalman, R. E, "A new approach to linear filtering and prediction problems", 1960.

[4] He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask r-CNN. In Proceedings of the IEEE international conference on computer vision (pp. 2961-2969).

[5] Girshick, R. (2015). Fast r-CNN. In Proceedings of the IEEE international conference on computer vision (pp.1440-1448).

[6] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. (2016, October). SSD: Single shot multibox detector. In European conference on computer vision (pp. 21-37). Springer, Cham.

[7] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 779-788).

[8] Burić, M., Pobar, M., & Ivašić-Kos, M. (2018, January). Ball detection using YOLO and Mask R-CNN. In 2018 International Conference on Computational Science and Computational Intelligence (CSCI).

[9] Redmon, J., & Farhadi, A. (2018). Yolov3: An incremental improvement. arXiv preprint arXiv:1804.02767.

[10] Wojke, N., Bewley, A., and Paulus, D, "Simple online and real-time tracking with a deep association metric", IEEE International Conference on Image Processing (ICIP) pp. 3645-3649, Sep 2017.

[11] Kristan, M., Leonardis, A., Matas, J., Felsberg, M., Pflugfelder., R, Cehovin Zajc, L., Vojir, T., Hager., G Lukezic, A., Eldesokey., A" The visual object tracking challenge results", IEEE International Conference on Computer Vision Workshops, Venice, Italy, 22–29 October 2017; pp. 1949–1972.

[12] Dendorfer, P., Rezatofighi, H., Milan, A., Shi, J., Cremers, D., Reid, I., Roth, S., Schindler, K., Leal-Taixe, L., "CVPR19 Tracking and Detection Challenge: How crowded can it get?", arXiv 2019, arXiv:1906.04567 2019.

[13] Okuma, K., Taleghani, A., De Freitas, N., Little, J.J., Lowe, D.G., "A boosted Particle Filter: Multitarget Detection and Tracking", In European Conference on Computer VISION, Springer: Berlin/Heidelberg, Germany, pp. 28–39, 2004.

[14] Pers, J., Kovacic, S., "Computer vision system for tracking players in sports games", In Proceedings of the First International Workshop on Image and Signal Processing and Analysis, Conjunction with 22nd International Conference on Information Technology Interfaces, IWISPA 2000, Pula, Croatia, pp. 177–182, 14–15 June 2000.

[15] Needham, C.J., Roger, D.B., "Tracking multiple sports players through occlusion, congestion and scale", BMVC 2001, 1, 93–102, 2001.

[16] Erikson, M., Ferreira, A., Cunha, S.A., Barros, R.M.L., Rocha, A., Goldenstein, S., "A multiple camera methodology for automatic localization and tracking of futsal players", Pattern Recognit. Lett., 39, 21–30, 2014.

[17] Bebie, T., Bieri, H., "SoccerMan-reconstructing soccer games from video sequences", In Proceedings of the 1998 International Conference on Image Processing. ICIP98 (Cat. No.98CB36269), Chicago, IL, USA, Volume 1, pp. 898–902 Oct 1998.

[18] Xu, M., Orwell, J., Jones, G., "Tracking football players with multiple cameras", In Proceedings of the 2004 International Conference on Image Processing, ICIP'04, Singapore; Volume 5, pp. 2909–2912, 24–27 Oct 2004.

[19] Jia, L., Tong, X., Li W., Wang, T., Zhang, Y., Wang, H., "Automatic player detection, labeling and tracking in broadcast soccer video", Pattern Recognit. Lett., 30, 103–113, 2009.

[20] Zhu, G., Xu, C., Huang, Q., Gao, W., "Automatic multi-player detection and tracking in broadcast sports video using support vector machine and particle filter", In Proceedings of the 2006 IEEE International Conference on Multimedia and Expo, Toronto, Canada, pp. 1629–1632, 9–12 July 2006.

[21] Lehuger, A., Duffner, S., Garcia, C., "A Robust Method for Automatic Player Detection in Sport Videos", Orange Labs: Paris, France, Volume 4, 2007.

[22] Bewley, A., Ge, Z., Ott, L., Ramos, F., Upcroft, B., "Simple online and real-time tracking", In Proceedings of the 2016 IEEE International Conference on Image Processing (ICIP), Phoenix, AZ, USA, pp. 3464–3468, 25–28 September 2016.

[23] Pobar, M., Ivaši´c-Kos, M., "Detection of the leading player in handball scenes using Mask R-CNN and STIPS", In Proceedings of the Eleventh International Conference on Machine Vision (ICMV 2018), Munich, Germany, International Society for Optics and Photonics: Bellingham, Volume 11041, pp. 110411V, 1–3 November 2018.

[25] Kristina Host, Miran Pobar, Marina Ivasic-Kos, “Analysis of Movement and Activities of Handball Players Using Deep Neural Networks”, Journal of Imaging, 9, 80, 2023.

[26] Miran Pobar, Marina Ivasic-Kos, “Active Player Detection in Handball Scenes Based on Activity Measures”, Journal of Sensors, 20, 1475, 2020; doi:10.3390/s20051475.

[27] Marina Ivasic-Kos, Miran Pobar, Jordi Gonzàlez, “Active Player Detection in Handball Videos Using Optical Flow and STIPs Based Measures”, 13th International Conference on Signal Processing and Communication Systems (ICSPCS), Gold Coast, QLD, Australia, 2019, pp. 1-8, doi: 10.1109/ICSPCS47537.2019.9008460.

[28] Xiaolong Sun, Yong Wang, Jawad Khan, “Hybrid LSTM and GAN model for action recognition and prediction of lawn tennis sport activities”, Soft Computing 27(23):1-20, Sep 2023. DOI: 10.1007/s00500-023-09215-4.

[29] Minhyeok Lee, “Recent Advances in Generative Adversarial Networks for Gene Expression Data: A Comprehensive Review”, Mathematics, 2023, 11(14), 3055; https://doi.org/10.3390/math11143055.

[30] JaeWon Kim, KyoHoon Jin, SooJin Jang, ShinJin Kang, YoungBin Kim,"Game effect sprite generation with minimal data via conditional GAN", Expert Systems with Applications,Volume 211,118491,ISSN 0957-4174,2023. https://doi.org/10.1016/j.eswa.2022.118491.