Haolun Shi[1*]

Target Detection in Wushu Competition Video Based on Kalman Filter Algorithm of Multi-Target Tracking

Journal of Electrical Systems

Abstract: - In a Wushu competition video, target detection for multi-target tracking presents a formidable challenge due to the intricate and dynamic movements inherent in the sport. The task involves identifying and tracking multiple athletes across varying camera angles, lighting conditions, and potentially occluded views. This manuscript present the Dual Transformer Residual Network(DTRN) optimized with Elk Herd Optimizer (EHO) for target detection in wushu competition video of multi-target tracking (WCV-DTRN-EHO). This study utilizes data from the Martial Arts, Dancing and Sports (MADS) and employs pre-processing technique such as Inverse Unscented Kalman Filter (IUKF) based pre-processing process to enhance video quality followed by the Dual Transformer Residual Network (DTRN)is classified as strength, velocity, agility, flexibility, and stamina. The weight parameters of the DTRN are optimized using Elk Herd Optimizer (EHO).The suggested approach is built in Python, and multiple performances evaluation metrics are used to estimate the proposed technique WCV-DTRN-EHO's efficiency for strength, speed, flexibility, agility and endurance in terms of accuracy 26.95%, 28.95%,27.95%,28.95%, and 27.79%, sensitivity 26.38,29.95%,27.95%, 22.13%, and 28.59%, precision 24.87%, 26.83%, 26.95%,24.95%,and 31.23%, F1-score 32.21%, 37.53%, 24.95%,25.95%, and 39.47% and computational time 88.95%, 89.95%,84.35%, 96.47%, and 85.46%while comparing other existing methods such as wushu competition video founded on convolutional neural network (WCV-CNN), wushu competition video based on deep convolutional neural network (WCV-DCNN) and wushu competition video based on recurrent neural network (WCV-RNN) respectively.

Keywords: Wushu, Video Analysis, Multi-Target Tracking, Inverse Unscented Kalman Filter, Tracking Accuracy, Data Pre-Processing, Elk Herd Optimizer

I. INTRODUCTION

A technique where instructors provide relevant learning materials, feedback, and the flexibility to modify the practice routine at any time to improve performance and action level can be somewhat understood as Sports Wushu video teaching [1].Effective target recognition is crucial for impartial judging and precise performance evaluation in the world of Wushu tournaments, where quickness and accuracy are critical [2]. In this regard, the Kalman Filter Algorithm for Multi-Target Tracking offers a promising method [3]. The system is able to track the movements of several targets, including athletes or areas designated for scoring, in a dynamic and fast-paced Wushu routine by dynamically predicting their states [4].It is especially well-suited for real-time target recognition in Wushu competition recordings because of its capacity to adjust to noisy data and forecast future states, giving judges and coaches trustworthy insights into athletes' achievements [5].Wushu offers a range of action types, including transition, static pause, quick, slow, strong, soft, and the ability to convert a virtual situation into a real one [6].

There is a size to the action's range. The vibrant beat of the Wushu exercise is reflected in these many motions [7]. Strong physical attributes including speed, agility, flexibility, strength, and endurance are necessary for athletes to portray the rhythmic changes of Wushu routines. These characteristics are also essential for athletes' fundamental abilities, unique traits, and training proficiencies [8] Wushu action completion quality can be accurately assessed by extracting challenging action images in real time [9]. Wushu is a distinctive national traditional sport in China as well as a superb aspect of the country's traditional culture [10]. A Wushu routine consists of all the exercises that are based on Wushu, such as attack and defence advance and retreat, rigidity, suppleness, fragility, and actuality in slow motion [11]. Actions that involve transitions, static pauses, swift, slow, powerful, soft, and the transformation of a virtual world into a real one are all included in Wushu [12]. There is a size to the action's range. The vibrant beat of the Wushu exercise is reflected in these many motions.

Strong physical attributes including speed, agility, flexibility, strength, and endurance are necessary for athletes to portray the rhythmic changes of Wushu routines. Athletes' distinctive talents, training capacities, and fundamental skills all depend on these attributes [13, 14]. By extracting difficult action photos in real-time,

[1] *Institute of Physical Education, Yichun College, Yichun, Jiangxi, 336000, China

[1*]Email: haolunshiphd@gmail.com

Wushu action completion quality can be precisely evaluated. In order to extract the challenging photos in Wushu in real time, the action contour feature has to be extracted and compared to the database contour [15, 16].
Each state's probability distribution is used to assess the current state of knowledge, and in order to increase the effectiveness of decision-making, the decision space is investigated. In addition, the process becomes more challenging when attempting to discriminate distinct targets that might be dressed similarly. A mix of sophisticated computer vision methods, including object detection, motion analysis, and potentially even deep learning algorithms designed for handling fast-paced action scenarios, would be necessary to achieve reliable target detection [17].
The existing techniques of WCV-DTRN-EHO in sports based methods affected from various issues of diligence. To overcome these limitations, researchers are exploring innovative approaches such as advanced optimization algorithms to achieve better solutions for target detection in wushu competition video utilizing Dual Transformer Residual Network (DTRN) optimized with the Elk Herd Optimizer (EHO).There aren't many approach-based publications in the literature that address this issue; these shortcomings and issues are what spurred this study effort. This work's primary contribution is,
The study utilizes data from the MADS Dataset.

- Inverse Unscented Kalman Filter is utilized during pre-processing.
- After that, the pre-processed data which classify action based on their strength, velocity, adaptability, agility, and stamina are transmitted to the Dual Transformer Residual Network.
- The classification task is carried out using DTRN; the neural network's weight parameter is optimized using Elk Herd Optimizer.
- The proposed WCV-DTRN-EHO technique is implemented, and performance metrics such as computing time, accuracy, precision, and sensitivity are examined.

The remaining portions of this manuscript are organized as follows: sector 2 reviews the relevant literature; sector 3 explains the proposed approach; sector 4 presents the results and discussion; and sector 5 conclusions.

II. LITERATURE SURVEY

Numerous research studies on target detection in wushu competition videos based on different approaches and factors are available in the literature. Among them are the following reviews,
Liu et al. [18] have presented the target detection in wushu competition video of multi-target tracking. It outlines the design strategy for a target intelligent tracking video processing-based martial arts motion feedback system. It then optimizes the system's hardware configuration, adds a composite tracker, and builds a moving target tracking model employing deep recurrent neural networks (RNNs) based on the motion characteristic parameters computed from the article by combining real-time target intelligent tracking video processing technology.
Liu et al. [19] have presented the target detection in wushu competition video of multi-target tracking. The modelling and live updating techniques of 2D PCA have been thoroughly investigated, and the Convolutional neural networks are utilized in the 2D PCA approach to separate overlapped blocks into their foreground (CNNs). The study model for the martial arts deconstruction motion VR image's adaptive recognition technique was then established via the use of frame difference, optical flow, and backdrop difference study methodologies. Wushu routine has a distinct performance in addition to being distinct from other competition quality rules from the paper.
Rodrigo et al. [20] have suggested the target detection in wushu competition video of multi-target tracking. Deep convolutional neural network (DCNN) was used to provide an innovative method—the first, to the best of our knowledge, that is expressly created for this complicated sport—for the automated identification of key moments in user-made parody films. Most current methods for similar sports either use depth cameras for automatic recognition or use high-level semantics like preset camera angles or standard editing techniques. In the first step, important points of interest are located in the foreground and their motion within the video frames is estimated. These sites were categorized into regions of interest in the second stage according to their mobility and closeness from the paper.
Sun et al. [21] have presented the target detection in wushu competition video of multi-target tracking. The challenging movement technology of a Real-time VR picture extraction and modelling were used in martial arts practice. The pre-processing content of existing photos will be analysed in this article, taking into account the intricate peculiarities of martial arts movements. To improve the qualities of the gathered photos, median

filtering techniques were used, along with image enhancement and filtering. The original image's visual impact can be enhanced in this way, and the modified image help with the segmentation process later on from the paper. Ishac et al. [22] have presented the target detection in wushu competition video of multi-target tracking. Evaluating the impact force, velocity, impulse, momentum, and reverse-step punch of Taekwondo examined the kinematics characteristics of these methods and compared them to the three martial arts styles of Shaolin, Wushu, and Hapkido. Created a sensing system that included two to measure dynamic acceleration from the paper, the motion of the punching bag is recorded by one GoPro Hero 3 camera, while the movements of the practitioner are captured by the other. Kinovea motion analysis software and an IC Sensor Model 3140 accelerometer mounted on a punching bag were employed.

Lee and Jung [23] have presented the target detection in wushu competition video of multi-target tracking. The Human Action Dataset from the Taekwondo Unit technique is made up of series of multimodal images showing poomsae movements. 1936 action examples of eight unit processes, executed by 10 experts and recorded by two camera viewpoints, are included in TUHAD. Taekwondo action detection was handled using convolutional neural networks based on key frames; the accuracy of the system was verified across a variety of input configurations from the study.

Torigoe et al., [24] have presented the target detection in wushu competition video of multi-target tracking. The goal of Human Action Recognition (HAR) is to create a prototype automation client that can accurately classify human actions by analysing a video of a combat sports competition or workout. In this study, the potential use of HAR in combat sports will be examined. Deep Learning architectures were integrated into client-server systems for data storage and analysis using bespoke algorithms in order to examine Computer Vision (CV) architectures that analyse real-time video data streams. One of the main components of the project was the automation client, which taught and programmed CV robots to watch and mimic specific human activity patterns from the paper.

III. PROPOSED METHODOLOGY

In this section describe the proposed methodology WCV-DTRN-EHO. Data gathered from MADS. Using an Inverse Unscented Kalman Filter, the data is removed during the pre-processing phase. The pre-processing step is sent to the classification stage, where the strength, speed, flexibility, agility and endurance are categorized using Dual Transformer Residual Network (DTRN). The Elk Herd Optimizer (EHO)is employed to enhance the Dual Transformer Residual Network by optimizing the neural network's motion parameter. The Block schematic illustrating the proposed strategy is presented in Fig. 1.

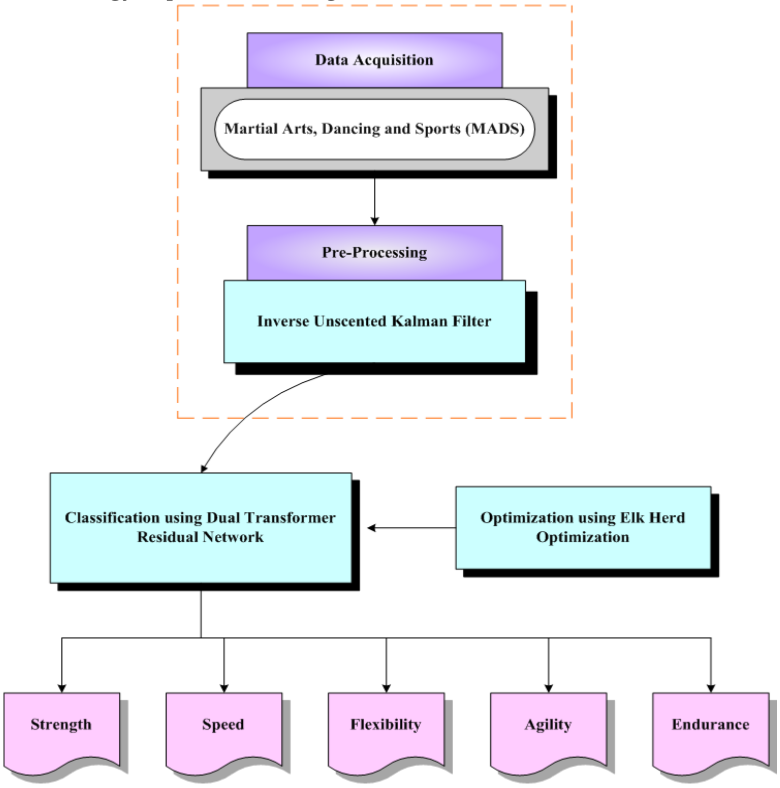

Fig 1: Block schematic illustrating the proposed strategy

A. Data Acquisition

The MADS dataset [25] is where the data was gathered. The information was captured in a studio setting with some background noise. Bumblebee-II cameras made by Point Grey were used to record the video footage. While the stereo images were taken from a single viewpoint, the multi-view data was gathered using three cameras positioned around the capture region. When the cameras were linked to the same hub, the multi-view data was automatically synchronized and recorded at a frame rate of 15 frames per second. At 10 or 20 frames per second, the depth data (stereo picture) was recorded. The image resolutions are 1024 × 768. The ground-truth pose data was recorded at 60 frames per second with a MOCAP device. Every video and piece of motion capture data is synchronized and calibrated to the same coordinate.

B. Data Pre-Processing Using Inverse Unscented Kalman Filter (IUKF)

In this, a MADS dataset is pre-processed using Inverse Unscented Kalman Filter in various engineering applications, including radar target tracking [26]. The UKF generates a set of $2nx+1$sigma points deterministically from the previous state estimate including the previous estimate itself as one of the sigma points. The state estimates are then derived as a weighted sum of the propagated sigma points after they have been propagated via the non-linear system model. In I-UKF, we assume that the attacker is employing a forward UKF to compute its estimate $\hat{x}_k$ with known state transition and observation. Using observation, the IUKF then deduces the estimate $\hat{\hat{x}}_k$ of $\hat{\hat{x}}_k$. In order to calculate the state estimate $\hat{x}_{k+1}$ and the corresponding error covariance matrix estimate, the attacker used forward UKF recursions.

$$\{s_{i,k}\}0 \le i \le 2n_x = S_{gen}\left(\hat{x}_k, \sum\nolimits_k \right), \quad (1)$$

$$s^*_{i,k+1|k} = f(s_{i,k}) \qquad \forall i = 0,1,\ldots 2n_x \quad (2)$$

Under the known forward UKF assumption, the inverse filter’s state transition as

$$\hat{x}_{k+1} = \sum_{i=0}^{2n_x} \omega_i \left(s_i^*, k+1|k - K_{k+1} q_i^*, k+1|k\right) + K_{k+1} h(x_k + 1) + K_{k+1} v_{k+1} \quad (3)$$

In this state transition, $\hat{x}_{k+1}$ is a recognised external input, and the process noise is represented by v_{k+1}. Since the functions $f(.)$ and $h(.)$ are known, the propagated sigma points $\{s_i^*, k+1|k\}$and $\{q_i^*, k+1|k\}$, and the gain matrix K_{k+1} relate to the initial set of sigma points $\{s_i^*, k\}$. These sigma points are then determined from the preceding state estimate in a deterministic manner. $\hat{x}_k$ and covariance matrix.

Denote the forward continuous-discrete UKF’s time update as $\hat{x}_{k+1|k} = X_1(\hat{x}_k)$ and $\sum_{k+1|k} X_2\left(\hat{x}_k, \sum_k \right)$. I-UKF’s state transition becomes

$$\hat{x}_{k+1} = X_1(\hat{x}_1) - \sum_{i=0}^{2n_x} \omega_i K_{k+1} q_i^*, k+1|k + K_{k+1} h(x_k + 1) + K_{k+1} v_{k+1} \quad (4)$$

Here, the propagated points $\{q_i^*, k+1|k\}$are once more derived deterministically from the anticipated state $\hat{x}_{k+1}|k$ and covariance estimate $\sum_{k+1}|k$, which in turn are functions of $\hat{x}_k$ and $\sum_k$ via solutions $X_1(.)$and $X_2(.)$. Hence, it simply becomes $\hat{x}_{k+1} = \tilde{f}\left(\hat{x}_k, \sum_k, x_{k+1}, v_{k+1}\right)$ with $\tilde{f}(.)$now denoting the modified state transition function. The pre-processed data’s are fed in to classification process.

C. Classification using Dual Transformer Residual Network (DTRN)

In this section, classification of wushu competition video data utilizing Dual Transformer Residual Network (DTRN) was discussed in this section of classification [27].Three pairs of embedding and an up-sampling block make up the transformer branch. This branch prioritizes the extraction and organization of global contextual information through multi-head attention within. The residual branch integrates both local and global data by feeding it feature maps.

The token embedding module is composed of cascaded convolution layers, normalization layers, and reshaping operations. The following may be expressed mathematically as the token embedding module's output:

$$e_\alpha = LN(\mathrm{Re}\,shape(Conv(e_\beta))), \tag{5}$$

Here e_α and e_β are the parameters, correspondingly. Additionally, LN represents the normalizing layer. Lastly, $\mathrm{Re}shape(.)$ and $Conv(.)$ represent the operators for convolution and reshape. The following provides the numerical S output of the convolutional transformer blocks (CTB) that is also named as A:

$$A_{out} = MLP(LN(B_{out})) + B_{out}, \tag{6}$$

Where,

$$B_{out} = B_{multi}(CP(A_{in})) + A_{in}, \tag{7}$$

With A_{in} and A_{out} serving as the A's input and output, respectively. Moreover, the multi-head attention action is called B_{multi}, and the output is called B_{out}. Convolution projection and multi-layer perception are denoted by CP and MLP, correspondingly.

The predicted output imbalance can be expressed as:

$$VI_{out} = Conv(c(Conv(VI_{in}))) \times S + VI_{in}, \tag{8}$$

Here VI_{in} and VI_{out} are indicated as the input and output are Imbalance, correspondingly. Furthermore, S is indicated as the residual scale while $\mathrm{Re}Lu$ is the $\mathrm{Re}Lu$ activation function. In this work, Elk Herd Optimizer for accurate video cover classification in deep learning, this method optimizes the DTRN optimum parameter e_α *and* e_β. Here, EHO is applied for tuning the weight parameter of DTRN. Finally the classification is done using dual transformer residual network.

D. Optimization of Elk Herd Optimizer (EHO)

The EHO [16] is discussed. It is used to accurately detect and track multiple targets simultaneously for performance evaluation and analysis. The term Elk Herd novel nature-inspired swarm-dependency optimization algorithm named EHO It draws inspiration from the elk herd's breeding program. Elks have two main seasons for mating, rutting, and calving.EHO efficiently converges to optimal solutions by mimicking EHO behaviour, leading to quicker convergence and reduced training time for DTRN. The flowchart of EHO is shown in fig 2.

Step 1: Initialization

Initially in EHO, a matrix representing the starting positional vectors of the search is identified. This matrix is originally established as random values inside a search space. Additionally, each position vector has a value for the initial fitness function. The steps of the suggested EHO initialization are mathematically expression is given in equation (10),

$$P = \begin{bmatrix} P_{11} & P_{12} & P_{13} & \cdots & P_{1n} \\ P_{21} & P_{22} & P_{23} & \cdots & P_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ P_{n1} & P_{n2} & P_{n3} & \cdots & P_{nn} \end{bmatrix} \tag{9}$$

Where P signifies set of current candidate solutions that are created randomly utilizing Fault diagnosis in robotics, $P_{n,b}$ is the number of ideal solution.

Step 2: Random generation

After start up, input factors are created at arbitrary. Therefore, optimization process of EHO and it transfer from exploration to exploitation steps utilizing various behaviours depend on this condition.

Step 3: Fitness Function

The outcome is determined by initialized judgments and random responses. The fitness is then computed using the equation (11),

$$Fitness\,Function = optimizing\ (e_\alpha, e_\beta) \tag{10}$$

Step 4: Exploration Phase

At this stage, robotics faults are diagnosed. The diving method comprises exploration phase, whirling method comprises exploitation phase of metaheuristic process. Thus, the each iteration is given in equation (12),

$$p_i(t+1) = \frac{f(y^j)(\varsigma_i)}{\sum_{j=1}^{k} f(y^k)(\tau_i)} \tag{11}$$

Where, $p_i(t+1)$ denotes selection probability next iteration of T that created by first search technique $f(y^j)$. $f(y^k)$ Is best-obtained solution until t^{th} iteration that reflects estimated place.

Step 5: Exploitation phase for optimizing $(w_\alpha \; and \; w_\beta)$

In the fourth strategy, the exploration phase in the Elk Herd Optimizer involves dynamically exploring the parameter space, potentially leveraging principles of collective behaviour inspired by starling murmuration's. This phase aims to discover new solutions efficiently while balancing exploration with exploitation. This behaviour is mathematically presented as in Equation (13),

$$p_i(t+1) = p_i^j(t) + \beta(p_i^{h_i}(t) - p_i^j(\varsigma_i)) + \gamma(p_i^r(t) - p_i^j(\tau_i)) \tag{12}$$

Where, $p_i^j(t)$ is an attribute of i, h_j denotes bull of harem j and β signifies coefficients.

Step 6: Termination Condition

The function values $\varsigma_i \; and \; \tau_i$ of generator from Node-Level Capsule Graph Neural Network is optimized with the use of EHO; continue from step 3 until the halting requirements are met. $P_N = P_N + 1$ is met. Then DTRN has accurately predicted and classified the target detection with MADS Dataset with higher tracking accuracy.

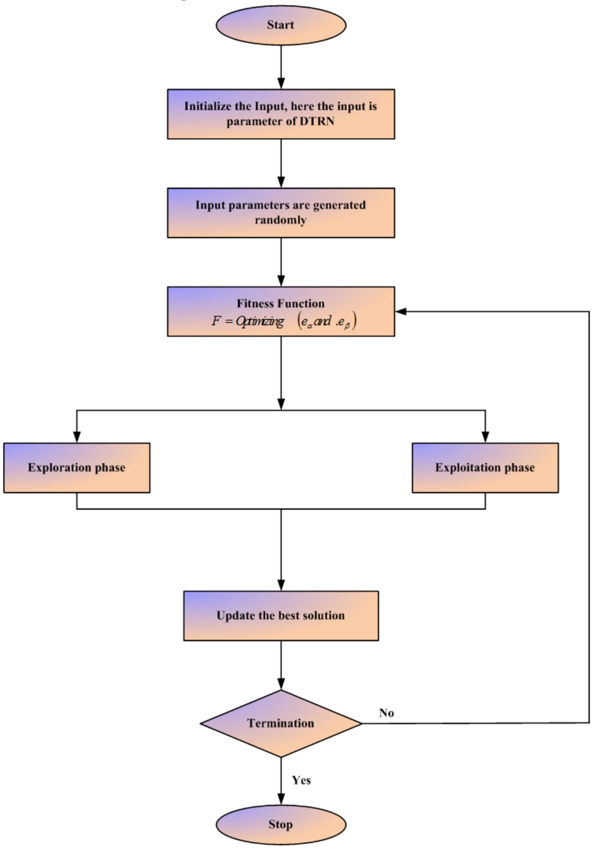

Fig 2: Flowchart of EHO

IV. RESULT AND DISCUSSION

This sector discusses the experimental outcomes of the suggested method. Next, Python is used to simulate the suggested method under the specified performance metrics. The proposed WCV-DTRN-EHO approach is implemented in Python using MADS dataset. The obtained outcome of the proposed WCV-DTRN-EHO approach is analysed with existing systems like WCV-CNN, WCV-DCNN, and WCV-RNN respectively.

A. Performance Measures

In order to choose the optimal classifier, this is an important task. Performance indicators like computation time, sensitivity, F1-score, accuracy, and precision are analyzed to assess performance. The confusion matrix will be used to scale the performance measures, it is decided. False Positive, False Negative, True Positive, and True Negative values are needed to scale the confusion matrix.

- True Positive (TP): Samples where the true class label is exact and the projected class label implies a positive value are counted.
- True Negative (TN): Number of examples where the true class label is exact and the predicted class label suggests a negative value.
- False Positive (FP): Number of samples when the true class label is imprecise and the predicted class label suggests a positive value.
- False Negative (FN): Number of samples where the true class label is imprecise and the predicted class label suggests a negative value.

1) Accuracy

It is the ratio of the number of accurate forecasts to the total number of predictions made for a dataset. Equation (13), which quantifies it, is used.

$$Accuracy = \frac{(TP + TN)}{(TP + FP + TN + FN)} \tag{13}$$

2) Precision

A statistic called precision counts the number of correctly predicted favorable outcomes. Equation (14) is used to scale this.

$$\Pr ecision = \frac{TP}{(TP + FP)} \tag{14}$$

3) Sensitivity

Sensitivity can also refer to true positive rate or recall. The equation (15), based on the sensitivity, calculates

$$Sensitivity(sen) = \frac{TP}{TP + FN} \tag{15}$$

4) F1 Score

The accuracy and precision weighted mean is called the F1-Score. It is expressed by equation (16),

$$F1Score = \frac{TP_\alpha}{\left(TP_\alpha + \frac{1}{2}\left[FP_\lambda + FN_\gamma\right]\right)} \tag{16}$$

B. Performance Analysis

The simulation results of the suggested WCV-DTRN-EHO approach are shown in Figure 3 to 7. Next, the suggested WCV-DTRN-EHO approach is contrasted with WCV-CNN, WCV-DCNN, and WCV-RNN, among other current approaches.

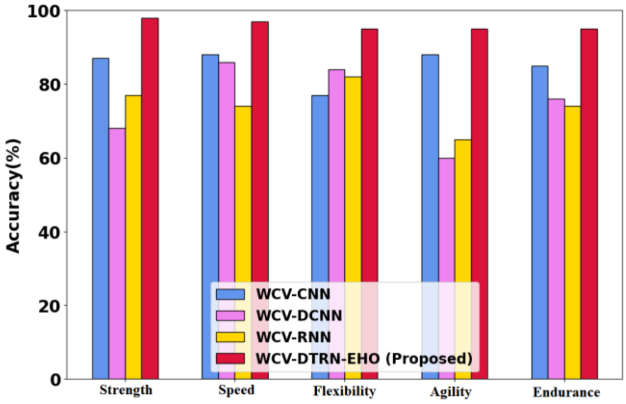

Fig 3: Accuracy value comparison between the proposed and existing methods

Fig. 3 shows Accuracy value comparison between the proposed and existing methods. The performance of the proposed technique results in accuracy that are 50.52%, 30.72%, 45.96% higher for the classification of strength, 40.42%, 57.52%, 49.52% higher for the classification of speed41.42%, 50.72%, 38.41% higher for the classification of flexibility, 40.42%, 55.52%, 43.52% higher for the classification of agility, 40.42%, 58.52%, 47.52% higher for the classification of endurance when compared to the WCV-CNN, WCV-DCNN, and WCV-RNN models that are currently in use, respectively.

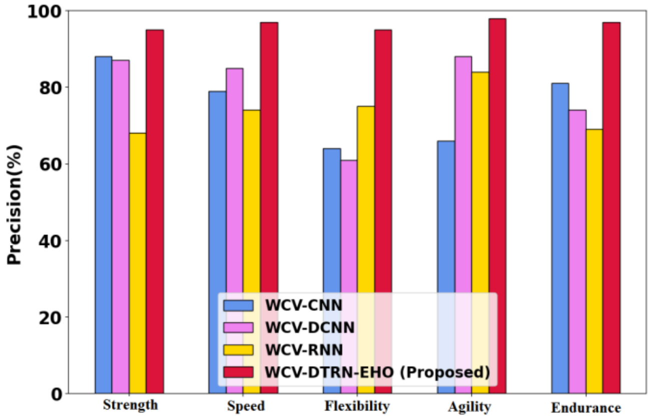

Fig 4: Examination of precision value using both the proposed and existing approach.

Examination of precision value using both the proposed and existing approach is depicts in Fig 4. The performance of the proposed technique results in precision that are 58.52%, 41.75%, 55.92%, higher for the classification of strength, 41.42%, 53.52%, 43.57% higher for the classification of speed, 41.42%, 50.72%, 38.42% higher for the classification of flexibility, 44.42%, 51.54%, 41.43% higher for the classification of agility and47.32%, 49.72%, 39.42% higher for the classification of endurance when compared to the WCV-CNN, WCV-DCNN, and WCV-RNN models that are currently in use, respectively.

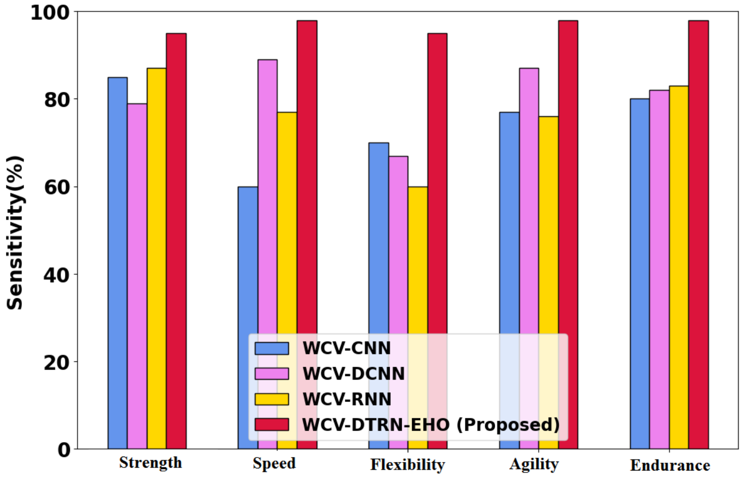

Fig 5: Sensitivity value performance using the proposed and existing methods.

Sensitivity value performance using the proposed and existing methods is depicts in Fig 5. The performance of the proposed technique results in sensitivity that are 50.52%, 41.72%, 53.92%, higher for the classification of strength, 41.49%, 53.54%, 43.57% higher for the classification of speed, 41.46%, 50.78%, 38.49% higher for the classification of flexibility, 44.43%, 52.51%, 43.58% higher for the classification of agility, 45.42%, 47.54%, 48.97% higher for the classification of endurance when compared to the WCV-CNN, WCV-DCNN, and WCV-RNN models that are currently in use, respectively.

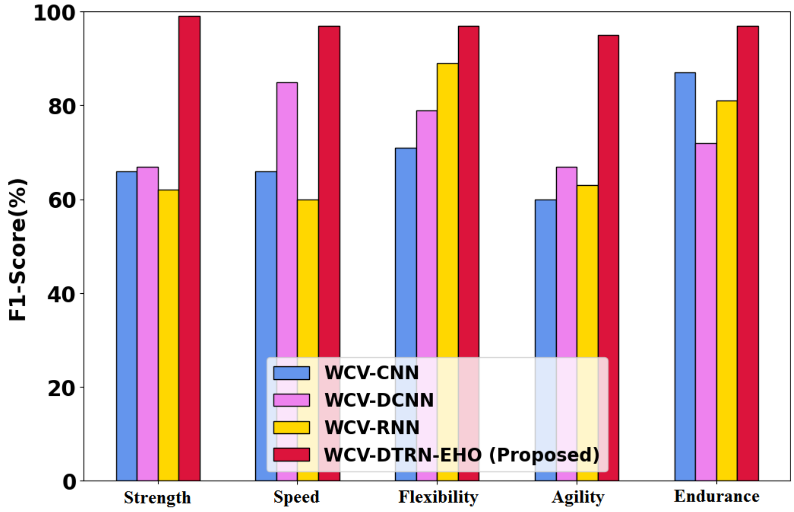

Fig 6: Comparing the F1-score value using the proposed and existing methods.

Comparing the F1-score value using the proposed and existing methods is depicts in Fig 6. The performance of the proposed technique results in F1-score that are 41.57%, 45.79%, 33.54%, higher for the classification of strength, 36.43%, 51.59%, 31.52% higher for the classification of speed, 42.43%, 47.71%, 56.49% higher for the classification of flexibility, 32.47%, 41.53%, 34.77% higher for the classification of agility and 52.23%, 42.37%, 46.24% higher for the classification of endurance when evaluated to the existing WCV-CNN, WCV-DCNN, and WCV-RNN models respectively.

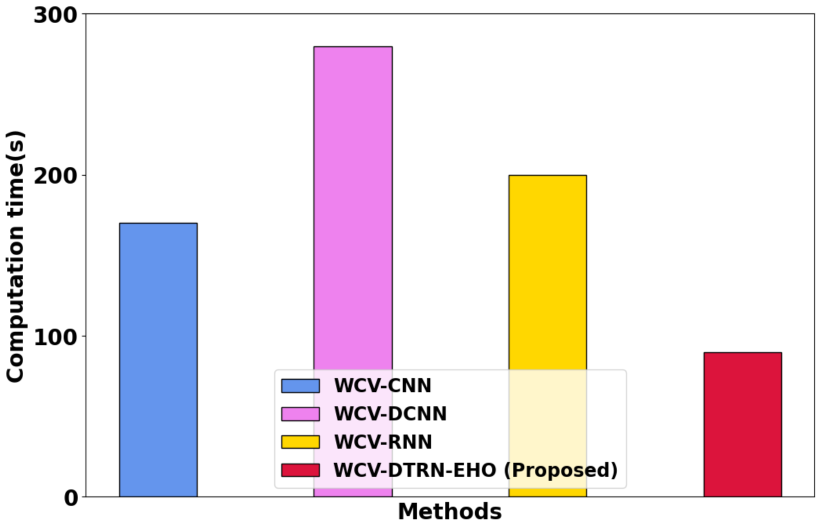

Fig 7: Computational time analysis using the proposed and existing methods.

Computational time analysis using the proposed and existing methods is depicts in Fig 7. In comparison to other current approaches such as WCV-CNN, WCV-DCNN, and WCV-RNN, the suggested WCV-DTRN-EHO method offers 170, 280, and 200 seconds of reduced computing time, respectively.

C. Discussion

The presented discussion underscores the effectiveness of the proposed Multi-Target Tracking using WCV-DTRN-EHO method in enhancing various aspects of object tracking performance compared to existing techniques such as WCV-CNN, WCV-DCNN, and WCV-RNN. Through comprehensive simulation results depicted in Figures 3 to 7, it is evident that WCV-DTRN-EHO achieves significant improvements in accuracy, tracking rate, F-score, precision and computation load. These enhancements demonstrate the method's capability to address key challenges in object tracking, including blurry targets, rapid object movements, and dynamic background changes, leading to more precise and efficient tracking results. Notably, the method outperforms existing techniques across multiple performance metrics, indicating its superiority in terms of accuracy, efficiency, and computational time. These findings highlight the potential of WCV-DTRN-EHO to advance surveillance systems by providing more reliable and comprehensive object tracking capabilities, thereby contributing to enhance security and monitoring applications in real-world scenarios.

V. CONCLUSION

In conclusion, the study on target detection in wushu competition video of multi-target tracking is analysed. The MADS is a crucial initial stage in the data collection process. Pre-processing involves processing the motion data using an inverse unscentred Kalman filter. The results of the pre-processing are sent to the classification, which uses the Dual Transformer Residual Network (DTRN) to effectively categorize the motion data's power, velocity, adaptability, agility, and endurance. The Elk Herd Optimizer (EHO) is introduced to enhance DTRN, which accurately classifies speed, strength, flexibility, agility and endurance. The suggested method is assessed using the Python working platform and contrasted with other existing methods.. The proposed method shows better results compared with existing WCV-CNN, WCV-DCNN, and WCV-RNN methods. The proposed WCV-DTRN-EHO method displays higher accuracy as 97%, lower error rate as 0.05%, greater F1-score as 98%,precision as 97%, sensitivity as 98%, and specificity as 96%thatis employed in the system's last stage and shows that it is capable of precisely identifying the motion of strength, speed, flexibility, agility and endurance. Based on the results, it is conclude that the proposed method shows better results compared with existing methods.

REFERENCE

[1] Chen, S., &Ren, Y. (2022). Small amplitude periodic solution of Hopf Bifurcation Theorem for fractional differential equations of balance point in group competitive martial arts. *Applied Mathematics and Nonlinear Sciences, 7*(1), 207-214.

[2] Ding, Q., & Ding, Z. (2021). Machine learning model for feature recognition of sports competition based on improved TLD algorithm. *Journal of Intelligent & Fuzzy Systems, 40*(2), 2697-2708.

[3] Abdelali, H.A., Derrouz, H., Zennayi, Y., Thami, R.O.H. and Bourzeix, F., 2021. Multiple hypothesis detection and tracking using deep learning for video traffic surveillance. *IEEE Access, 9*, pp.164282-164291

[4] Meimetis, D., Daramouskas, I., Perikos, I. and Hatzilygeroudis, I., 2023. Real-time multiple object tracking using deep learning methods. *Neural Computing and Applications, 35*(1), pp.89-118.

[5] Kumar, M., Ray, S. and Yadav, D.K., 2023. Moving human detection and tracking from thermal video through intelligent surveillance system for smart applications. *Multimedia Tools and Applications, 82*(25), pp.39551-39570.

[6] Ong, P., Chong, T. K., Ong, K. M., & Low, E. S. (2022). Tracking of moving athlete from video sequences using flower pollination algorithm. *The Visual Computer*, 1-24.

[7] Jianqiang, Z., Jialin, L., & Yan, L. (2022). TOPSIS missile target selection method supported by the posterior probability of target recognition. *Applied Mathematics and Nonlinear Sciences, 7*(2), 713-720.

[8] Tian, Z., Dong, F., Li, D., & Liu, C. (2024). 3D motion trajectory prediction based on optical image remote sensing data in sports training simulation. *Optical and Quantum Electronics*, *56*(3), 305.

[9] Hu, M., Zhu, X., Wang, H., Cao, S., Liu, C., & Song, Q. (2023). Stdformer: Spatial-temporal motion transformer for multiple object tracking. *IEEE Transactions on Circuits and Systems for Video Technology*.

[10] Pengyu, W., &Wanna, G. (2021). Image detection and basketball training performance simulation based on improved machine learning. *Journal of Intelligent & Fuzzy Systems*, *40*(2), 2493-2504.

[11] Chen, Y., Liu, M., Dong, L., Hui, M., Zhao, Y., & Kong, L. (2022, December). Tracking and shooting system based on trajectory prediction. *In Optical Metrology and Inspection for Industrial Applications IX* (Vol. 12319, pp. 356-361).SPIE.

[12] Sun, H., Lv, W., Khadidos, A. O., &Kharabsheh, R. (2021). Research on the influence of fuzzy mathematics simulation model in the development of Wushu market. *Applied Mathematics and Nonlinear Sciences, 7*(2), 341-352.

[13] Hasegawa, R., Uchiyama, A., Okura, F., Muramatsu, D., Ogasawara, I., Takahata, H., ...&Higashino, T. (2022). Close-Contact Detection Using a Single Camera for Sports Considering Occlusion. *IEEE Access, 10*, 15457-15468.

[14] Jia, L., & Han, X. (2024). Visual system based on optical sensor in Wushu training image trajectory simulation. *Optical and Quantum Electronics, 56*(4), 501.

[15] Ji, J., & Liang, F. (2022). Influence of embedded microprocessor wireless communication and computer vision in Wushu competition referees' decision support. *Wireless Communications and Mobile Computing*, 2022, 1-13.

[16] Liu, Y., Wang, L., Tang, Y., &Ren, B. (2022). Judgment of Athlete Action Safety in Sports Competition Based on LSTM Recurrent Neural Network Algorithm.*Mathematical Problems in Engineering, 2022*.

[17] Thành, N. T., &Công, P. T. (2019). An evaluation of pose estimation in video of traditional martial arts presentation. *Journal on Information Technologies & Communications*, *2019*(2), 114-126.

[18] Liu, X., Liu, J., Fan, J., & Li, X. (2022). Research on Wushu sports feedback system based on real-time target intelligent tracking video processing. *Security and Communication Networks, 2022*.

[19] Liu, W., Han, X., &Kamruzzaman, M. M. (2020). Adaptive recognition method for VR image of Wushu decomposition based on feature extraction. *IEEE Access.*

[20] Rodrigo, M., Cuevas, C., Berjón, D., &García, N. (2024). Automatic highlight detection in videos of martial arts tricking. *Multimedia Tools and Applications, 83*(6), 17109-17133.

[21] Sun, N., Sun, R., Li, S., & Wu, X. (2020). Martial arts routine difficulty action technology VR image target real-time extraction simulation. *IEEE Access, 8*, 155811-155818.

[22] Ishac, K., & Eager, D. (2021). Evaluating martial arts punching kinematics using a vision and inertial sensing system.*Sensors, 21*(6), 1948.

[23] Lee, J., & Jung, H. (2020). Tuhad: Taekwondo unit technique human action dataset with key frame-based cnn action recognition. *Sensors, 20*(17), 4871.

[24] Torigoe, Y., Nakamura, Y., Fujimoto, M., Arakawa, Y., &Yasumoto, K. (2020). Strike Activity Detection and Recognition Using Inertial Measurement Unit towards Kendo Skill Improvement Support System. *Sensors & Materials, 32*.

[25] Zhu, C., Wang, S., Yu, C., Zhou, H., Fernandez, C., & Guerrero, J. M. (2024). An improved Cauchy robust correction-sage Husa extended Kalman filtering algorithm for high-precision SOC estimation of Lithium-ion batteries in new energy vehicles. *Journal of Energy Storage, 88*, 111552.

[26] Sui, J., Ma, X., Zhang, X., & Pun, M. O. (2023, July). Dtrn: Dual transformer residual network for remote sensing super-resolution. *In IGARSS 2023-2023 IEEE International Geoscience and Remote Sensing Symposium* (pp. 6041-6044).IEEE.

[27] Hubálovská, M., Hubálovský, Š.,&Trojovský, P. (2024). Elk Herd Optimizer: A New Human-Based Metaheuristic Algorithm for Solving Optimization Problems. *Biomimetics, 9*(3), 137.