Archery Human Pose Estimation System Based on DFPDeblur GAN Algorithm

Xiao Yan
School of Physical Education
Shanxi University, China
yx13154908967@126.com

Zhuohan Wu
School of Physical Education
Shanxi University, China
abcd1100866@163.com

Ting Wang
School of Physical Education
Shanxi University, China
kingtnnd@163.com

Abstract: *An Archery human pose estimation system based on Deep Feature Prior Deblurring Generative Adversarial Network (DFPDeblur GAN) is proposed. The system incorporates dynamic convolution, a technique that can adaptively adjust the parameters of the convolution kernel to extract multidimensional features, and a bidirectional Feature Pyramid Network (FPN), which effectively improves the image deblurring effect and multiscale feature fusion capability. Subsequently, the Improved High-Resolution Network (IHRNet), combined with the Coordinate Attention (CA)-mechanism, a mechanism that improves the accuracy of key point detection by focusing on the spatial location, is used to realize the precise localization of the key joints of the Archery action. The experimental results show that the proposed model achieves 94.3% key point localization accuracy, 9.1% Peak Signal-to-Noise Ratio (PSNR) improvement compared with the traditional method, up to 0.92 Structural Similarity Index (SSIM), and less than 2 seconds running time, which exhibits good real-time performance and robustness. The results show that the model performs well in a variety of lighting conditions and action phases, providing effective technical support for action analysis and training in Archery.*

Keywords: *Archery sport, attitude estimation, DFPDeblur GAN, HRNet, deblurring.*

Received January 15, 2025; accepted July 10, 2025
https://doi.org/10.34028/iajit/22/6/3

1. Introduction

1.1. Background and Motivation

As a competitive sport with a long history and high technical content, Archery requires athletes to complete precise, coordinated and complex movements in a very short time. During the key stages of drawing the bow, aiming and releasing the arrow, all the joints of the athlete's body need to be highly coordinated, and the subtle changes in the movements directly affect the accuracy and stability of the shooting. Traditional Archery training mainly relies on the coach's experience guidance and visual observation, which is difficult to objectively and real-time capture the details of the athlete's movements, resulting in a limited training effect [7, 29]. With the rapid development of computer vision and artificial intelligence technology, the intelligent Archery motion analysis system based on human posture estimation has gradually become a key technology to improve training science and efficiency. The system automatically captures athletes' body postures through video data, and combines machine learning models to achieve high-precision action recognition and evaluation, which greatly facilitates the quantitative analysis of sports performance and personalized training guidance.

However, in practical applications, the high-speed and complex dynamic scenes of Archery movements often lead to problems such as image blurring and low-lighting, which seriously affect the accuracy and stability of pose estimation [22, 23]. Image blurring not only reduces the recognition rate of the athlete's key points, but also makes it difficult to retain the detail information, which increases the difficulty of stance estimation. In addition, the Archery environment is variable and the lighting conditions are complex, which puts higher requirements on the robustness of the algorithm. Therefore, how to design a set of efficient and stable joint model of image deblurring and pose estimation has become a hot and difficult point in the current research, which is directly related to the practical effect and application promotion value of the Archery intelligent training system.

1.2. Related Work in Pose Estimation and Deblurring

Anvari *et al*. [3] proposed a new method inspired by classical regression models and trained on 3D motion capture data to achieve less parameter count in Archer Pose Estimation (APE) using virtual reality technology. The virtual reality simulation effect of pose estimation under this method is more realistic. Romero *et al*. [24] proposed a semi-automatic mechanism to enhance the training effectiveness of APE models, allowing synthetic humans to perform various actions to generate and edit visual scenes. The effectiveness of APE under this mechanism training was significantly improved, with shorter iteration times. Liu *et al*. [18] developed a method of inter-frame and intra-frame data

synchronization to lower the cost of pose estimation and enhance training efficiency. Compared to other methods, this method has improved the accuracy of action frame extraction by up to 8.3%. Abhishek and Tahir [2] proposed an adaptive boosting model to enhance the accuracy of context-based Archery action feature extraction. The average accuracy of this model for detecting ordinary sports datasets was 92.15%. Siaw *et al*. [27] proposed a marker-based optical Archery motion capture system. The system used a smartphone camera to capture the motion of the red marker, then tracked the coordinates of the marker, and calculated the angle. When the distance between the camera and the marker was 120 centimeters, the system had a low average root mean square error of 0.83°. Chen and Hu [6] concluded that the effectiveness of machine learning and deep learning techniques for activity recognition in sports such as swimming still needs to be improved, for this reason, the researchers proposed a new human swimming pose recognition model after combining reinforcement learning and inertial measurement units. Experimental results show that the new model has a balanced accuracy of 96.27% for human back, waist, and upper and lower limbs pose recognition. Tong and Wang [30] proposed a pose recognition model based on the combination of augmented Graph Convolutional Network (GCN) and Spatio-Temporal GCN (ST-GCN) for basketball player pose recognition. The model is able to handle graph-structured data with time series relationship, and extract the spatio-temporal features of human pose sequences by convolution operation. The results show that the ST-GCN model achieves 95.58% accuracy in basketball pose recognition.

Currently, deep learning methods based on Generative Adversarial Networks (GANs), especially the Deblur GAN algorithm, have provided a new approach to solving the problems of image blur and pose estimation accuracy [13]. The core of Deblur GAN is to use the GAN model to deblurr blurry images and generate clear high-quality images, thereby improving the accuracy of pose estimation. Lee *et al*. [16] proposed a deblurring network using

Deblur GAN to eliminate motion blur effects in UAV images. This method utilized a generative model to correct blur artifacts and generate clear images. The image quality evaluation under this model was higher than traditional methods. Jun [11] proposed a suitable method to eliminate blind blur caused by motion in images. This method was used for end-to-end learning of motion blur removal, describing the amount and type of blur caused by point light source imaging. This method provided the possibility of increasing real data and had a high blur elimination rate. Xiao *et al*. [33] considered the differences in feature abstraction levels extracted by different perceptual layers and used Deblur GAN based on weighted perceptual loss to deblur Unmanned Aerial Vehicle (UAV) images, thereby eliminating blur and restoring texture details of the images well. Kong *et al*. [14] found that due to the rapid movement between onboard cameras and fire targets, captured fire images often become dull and blurry. To this end, researchers have proposed a multiple input-output Deblur GAN architecture that fuses spatial and frequency domain message for image deblurring models. This model achieved an image similarity of 0.955 on a self-built dataset. Varela *et al*. [32] proposed a regression method using Deblur GAN to predict the parameters, length, and direction of linear motion blur kernels. In non-blind image deblurring methods, the cumulative histogram value error of the sum of squared differences of kernel parameters using this new method was higher than that of traditional methods.

1.3. Identified GAPs

Although there have been studies that have achieved certain results in the field of human pose estimation and image deblurring, there are still obvious shortcomings in the existing methods for highly dynamic and complex scenarios such as Archery sports. First, traditional deblurring algorithms rely on single-scale feature extraction, which is difficult to take into account the multi-scale dynamic changes of the moving objects, resulting in insufficient recovery of key details. Second, the existing pose estimation models are often difficult to realize efficient fusion of multi-resolution features when dealing with fast movements, and cannot adequately capture the coordinated motion characteristics of complex joints, which affects the accuracy and continuity of estimation. Furthermore, for the real-time tracking requirements in dynamic Archery scenarios, the traditional models have limited capabilities in target localization and time-series information fusion, making it difficult to meet the dual requirements of real-time and stability.

1.4. Suggested Contributions

To address these issues, the study proposes a joint model that fuses a Deep Feature Prior Deblur Generative Adversarial Network (DFPDeblur GAN) with an Improved High-Resolution Network (IHRNet). High-Resolution Network (IHRNet) as a joint model. Its innovativeness is mainly reflected in the combination of multi-dimensional dynamic convolution and bidirectional Feature Pyramid Network (FPN), which proposes a DFPDeblur GAN for fast dynamic actions in Archery, which significantly improves the clarity and detail restoration of the image; meanwhile, the IHRNet based on the Ghost module, the Sandglass module, and the Coordinate Attention mechanism (CA-mechanism), is designed, which effectively enhanced multi-scale feature fusion and coordinated recognition of complex joints; further introduced the target tracking module, realizing the real-time accurate capture of key points in dynamic scenes, so as to ensure the continuity and robustness of attitude estimation. The main

contributions of this study are that the proposed method effectively improves the image deblurring effect and the key point localization accuracy and robustness of stance estimation in complex dynamic scenes in Archery, realizes the efficient capture of complex joint coordination in Archery movements, and provides a more effective intelligent assistive method for Archery training and competitions.

2. Methods and Materials

Aiming at the problems of image blurring and difficult pose estimation caused by fast movements in Archery, the study firstly constructs an image deblurring model for Archery athletes based on the improved DFP Deblur GAN. Multi-scale features are extracted by full-dimensional dynamic convolution Omni-Dimensional Dynamic Convolution (ODConv) and bi-directional FPN, and combined with Iterative Attentionla Feature Fusion (IAFF) module for detail optimization. Secondly, the IHRNet for deblurring image input extracts multi-resolution features through Ghost and Sandglass modules as well as CA-mechanism, and combines with the target tracking module to capture dynamic targets and key joint features in real time. Finally, a novel Archery, Motion, High-Precision, Estimation (AMHPE) model is proposed that integrates deblurring and high-resolution pose estimation.

2.1. Image Deblurring Based on DFPDeblur GAN

2.1.1. Generator Design and Feature Extraction Path

Archery has unique characteristics compared to other sports. The main challenge lies in the high-precision requirements for athletes' postures, especially in the process of rapid and complex movements. How to capture the subtle posture changes of athletes is a key factor affecting Archery performance [28]. During the Archery process, athletes move at high speed and frequently during the moments of pulling the bow, aiming, and releasing the arrow, resulting in blurry images captured by the camera. The problem of image blur not only affects image quality and reduces the accessibility of detail information, but may also lead to pose estimation errors, thereby affecting the accuracy of training and competition [26]. Therefore, this paper designs a DFPDeblur GAN built on Deblur GAN. Unlike traditional GANs, DFPDeblur GAN introduces deep feature priors between the generator and discriminator. This prior information helps the generator to fully utilize the contextual information of the athlete's posture when restoring blurred images, ensuring that the image structure and details remain consistent during the deblurring process. Figure 1 shows the DFPDeblur GAN structure.

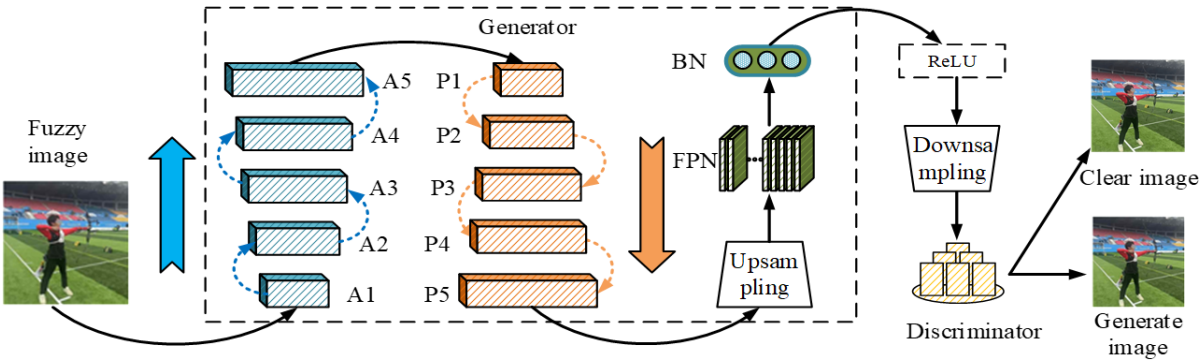


Figure 1. DFPDeblur GAN structure.

In Figure 1, the entire DFPDeblur GAN consists of a generator, discriminator, and attention mechanism module. The process first utilizes ODConv to improve the Mobilenetv2 network as the backbone network, and extracts features layer by layer through downsampling method to obtain five feature maps of different scales, namely A1 to A5. Next, these feature maps are processed through FPN's horizontal connection and top-down path to generate corresponding feature maps P1 to P5. On this basis, P1 to P5 are processed again through the bottom-up path of FPN to obtain feature maps N1 to N5. Subsequently, N1 to N5 are fused with the original image to generate the last image output. The generated image is input into the discriminator together with the target clear image, and the similarity probability value between the generated and clear images is calculated. The quality of the generated image is improved by optimizing the generator. Finally, through alternating training of the generator and discriminator, the performance of the network is gradually improved, completing the process of image deblurring. The generator structure of DFPDeblur GAN is shown in Figure 2.

In Figure 2, the generator includes multiple feature extraction modules, whose main function is to extract and fuse key features from the input blurred motion image to achieve deblurring processing. Firstly, the input blurred image is subjected to a five-layer 2D convolution operation from bottom to top through the backbone network, extracting feature maps of various scales. These feature maps are then fused through the top-down path of FPN to generate five fused features.

At the same time, the feature fusion process introduces horizontal connections and combines Batch Normalization (BN) layers and ReLU functions to enhance the expressive power and extraction efficiency of features. Then, IAFF dynamically focuses on the detailed features in the key region through the multilayer attention mechanism to reduce the interference of redundant information on the pose judgment. Compared with the traditional single-feature fusion method, IAFF is able to aggregate multi-scale features in multiple rounds, which significantly improves the ability to restore key details in blurred images [4, 17].

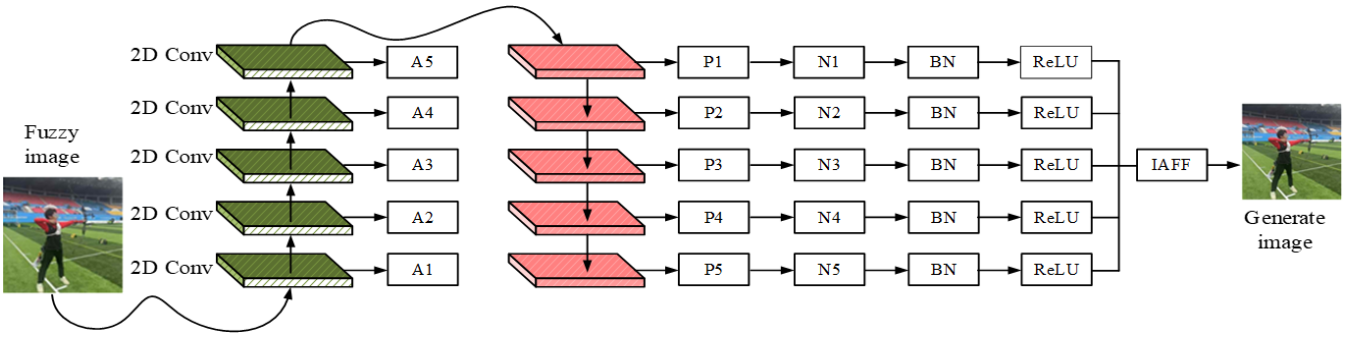


Figure 2. DFPDeblur GAN generator structure.

2.1.2. ODConv Module Principles and Benefits Analysis

Compared with traditional convolution (e.g., standard 3×3 convolution or group convolution), the use of ODConv for the backbone network can adaptively assign attention weights in spatial, channel, and input-output dimensions simultaneously, which enhances the expression fineness of the features and the dynamic modeling ability. Especially in the scenario of Archery sports where fast movements lead to image blurring, ODConv is able to extract multi-dimensional features more flexibly, enhancing robustness and structure preservation [35]. For example, Vancurik and Callahan [31] measured and introduced variables by combining wearable sensors with university tennis match video observations to detect abnormal movements in tennis, showing stable output results. The structure of ODConv is shown in Figure 3.

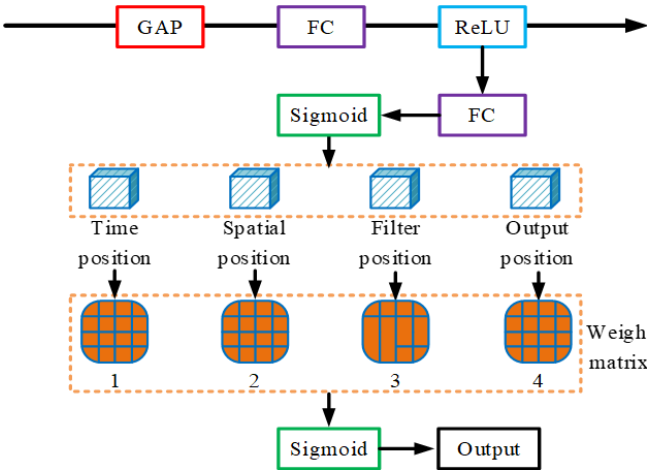


Figure 3. Structure diagram of ODConv.

In Figure 3, firstly, the input features are subjected to Global Average Pooling (GAP), which aggregates global information into feature vectors to quantitatively describe their feature lengths. Subsequently, the features are partitioned through a fully connected layer, generating four branches with different dimensions. Each branch corresponds to a different parameter configuration of the convolutional kernel, which is used to calculate the features of time position, spatial position, filtering position, and output position separately. Each branch calculates attention values through a specific weight matrix and extracts corresponding features. Subsequently, these features undergo specific mapping calculations to generate adjusted convolution kernel weights, further optimizing the feature extraction process. Finally, the sigmoid function normalizes the output and adjusts the weights to an appropriate range to ensure the stability and effectiveness of the output results [19]. The calculation for GAP is given by Equation (1).

$$z = \frac{1}{H \cdot W} \sum_{i=1}^{H} \sum_{j=1}^{W} x(i,j) \tag{1}$$

In Equation (1), x is the input feature map. $x(i, j)$ means the value of the feature map at position (i, j). H and W are the height and width of the feature map. z is the global eigenvector. The formula for generating branch feature weights is shown in Equation (2).

$$\alpha_k = \sigma(W_k \cdot y_k + b_k) \tag{2}$$

In Equation (2), σ and b_k are the weights and biases of branch k. y_k is the input feature of k. α_k denotes the attention weight of k. σ corresponds to the sigmoid function. At this point, the FPN connections from bottom to top and from top to bottom are shown in Equation (3).

$$\begin{cases} P_i = Conv_{1\times1}(A_i) + Upsample(P_{i+1}) \\ N_i = Conv_{3\times3}(P_i) + Downsample(N_{i+1}) \end{cases} \tag{3}$$

In Equation (3), P_i and N_i are feature maps fused from top to bottom and bottom-up. A_i is a feature map of

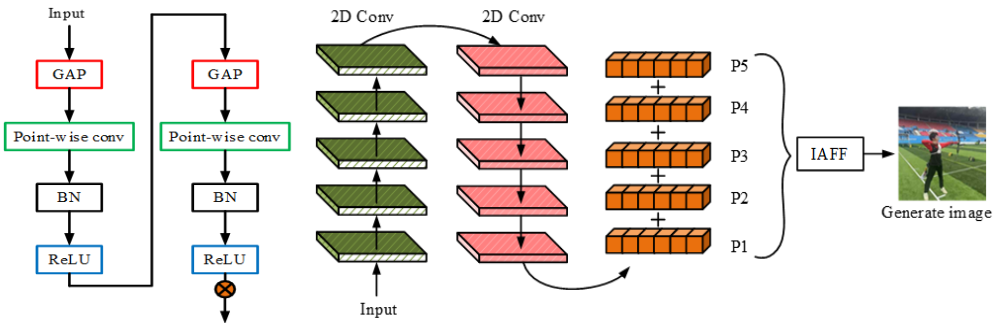
different scales extracted from the backbone network. *Upsample* and *Downsample* are upsampling and downsampling operations. The key features after focusing by the IAFF module are shown in Equation (4).

$$F_{IAFF}^{t} = F^{t} + \gamma \cdot Attention(Q^{t}, K^{t}, V^{t}) = F^{t} + \gamma \cdot Softmax\left(\frac{Q^{t}, K^{t^{T}}}{\sqrt{d}}\right) V^{t} \quad (4)$$

In Equation (4), F^t is the fusion feature of the t-th iteration. Q^t, K^t, and V^t are the query, key, and value matrix values in IFF at the t-th iteration. γ is the Feature Fusion Coefficient (FFC). $\sqrt{d}$ is the scaling factor. F_{IAFF}^{t} is the fused feature map after the t-th iteration.

2.1.3. IAFF Modular Structure and Iterative Feature Fusion Mchanisms

Specifically, the architecture diagram of IAFF module and bidirectional FPN is shown in Figure 4.



a) IAFF structure. b) Bidirectional FPN structure.

Figure 4. Architecture diagram of IAFF module and bidirectional FPN.

In Figure 4-a), the IAFF module is overlaid with two layers of Multi-Scale Channel Attention Module (MS-CAM) for initial feature fusion. The input features are subjected to two layers of point by point convolution and BN layer, as well as ReLU for global feature extraction, and then input to the right side for repeated operations. Finally, a weighted balanced feature map is obtained through matrix calculation. In Figure 4-b), in the top-down FPN structure, high-level semantic features are passed to low-level features to enhance their representational power, while ensuring feature continuity through horizontal connections. In the bottom-up FPN structure, the information of low-level features is transmitted to high-level features to supplement detailed information and optimize feature expression.

2.1.4. Hybrid Loss Function Design with Optimization Objective

In addition, to preserve the details of blurry images, this study constructs a novel hybrid loss function that integrates content loss, perceptual loss, and adversarial loss. The content loss formula is given by Equation (5).

$$L_{content} = \frac{1}{N}\sum_{i=1}^{N} \|G(x_i)y_i\|_1 \quad (5)$$

In Equation (5), $L_{content}$ is the content loss value. $G(x_i)$ is the image generated by the generator. $\|_\|_1$ is the norm of L_1. The perceptual loss is given by Equation (6).

$$L_{perceptual} = \frac{1}{N}\sum_{i=1}^{N}\sum_{l} \|\phi_l(G(x_i)) - \phi_l(y_i))\|_2^2 \quad (6)$$

In Equation (6), $L_{perceptual}$ is the perceptual loss value. $\phi_l(_)$ is the l-th layer feature map of deep networks. $\|_\|_2^2$ is the norm of L_2. The adversarial loss is shown in Equation (7).

$$L_{adversarial} = -\frac{1}{N}\sum_{i=1}^{N} \log D(G(x_i)) \quad (7)$$

In Equation (7), $L_{adversarial}$ is the adversarial loss value. $L_{adversarial}$ is a discriminator.

2.2. Construction of AMHPE Model Integrating Deblurring and IHRNet

2.2.1. Overall Structure and Flow of the IHRNet Model

After constructing the image deblurring model based on DFPDeblur GAN, it is found that the deblurring module can effectively improve the quality of the input image and provide clear and reliable basic image data for subsequent pose estimation. However, relying solely on deblurring images for pose estimation may still face the problem of insufficient capture of complex pose features. Especially, Archery movements have high technical complexity, involving coordination and synchronization of multiple joints [9, 25].

Therefore, it is necessary to introduce a network structure that can enhance the ability to capture complex poses while maintaining high resolution, further improving the accuracy and robustness of the model. HRNet, as an advanced attitude estimation framework proposed in recent years, can achieve high accuracy and robustness in attitude estimation tasks through parallel high-resolution and low-resolution feature representations, as well as multi-scale feature interaction mechanisms [20, 34]. Its core advantage lies in maintaining the continuity of high-resolution features throughout the entire feature extraction process, while

fusing features of different scales through information exchange between branches. It can capture both global pose information and attention to detailed joint features [8, 21].

However, to adapt to the continuity and coordination of joint movements in Archery, this study improves the HRNet structure and proposes an IHRNet, as shown in Figure 5.

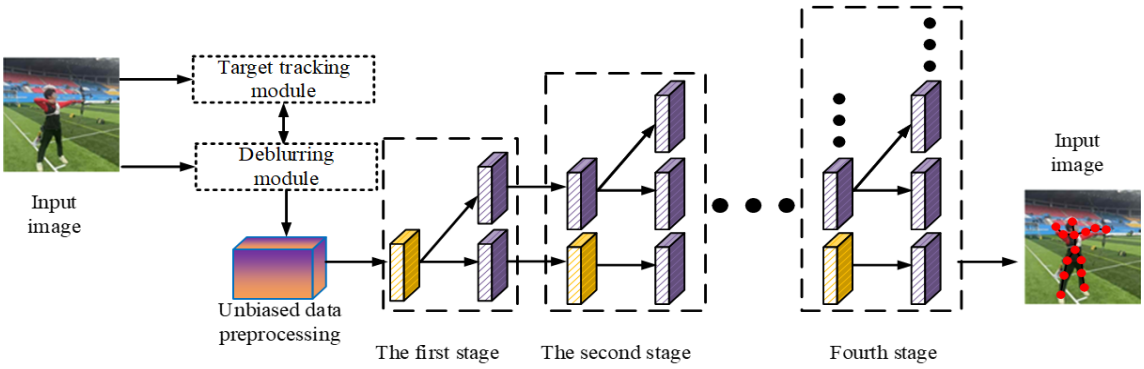


Figure 5. The network structure of IHRNet.

In Figure 5, the entire structure of IHRNet includes preprocessing and four different resolution network extraction stages. Firstly, in the preprocessing stage, motion blur removal module and target tracking module are used to collect camera image data of Archers, and unbiased data preprocessing is adopted to reduce the small errors caused by the initial review data. Then, the data are input into the HRNet of stage 1 for high-resolution feature extraction. Phase 1 consists of 4 layers, each of which integrates Ghost module, Sandglass module, and CA-mechanism. Then, the data are passed to the medium resolution network in stage 2 for feature extraction. At this point, the resolution network has more branches than the previous stage, capturing deeper level feature information through downsampling. After passing through four times in sequence, the data stream achieves multi-scale feature alignment and complementarity from high-resolution to low-resolution.

2.2.2. Architecture and Benefits of Ghost and Sandglass Modules

The Ghost module generates base features through a small number of standard convolutions, and then combines them with multiple linear transformations to quickly generate redundant features, significantly reducing the amount of model computation. Compared with the traditional convolution, it reduces the computational overhead by about 50% or more while maintaining the accuracy, which is especially suitable for real-time demanding tasks such as attitude estimation. The computational formula is shown in Equation (8) [10, 36].

$$Ghost(x) = F(x) + \sum_{i=1}^{b} \theta_i * F(x) \tag{8}$$

In Equation (8), $F(x)$ is the core feature generated by the main convolution. θ_i is the weight of the i-th lightweight operation. b is the amount of generated supplementary features. $Ghost(x)$ is the feature output by the Ghost module. The Sandglass module is structured to enhance the transfer efficiency of the model in maintaining high-resolution features by reversing the bottleneck structure of the standard residual block and combining the asymmetric feature flow with residual connections. Compared with conventional residual blocks, Sandglass is more suitable for detail modeling and position information retention in shallow networks. The computational formula is shown in Equation (9) [1].

$$S(x) = x + W_3 \cdot ReLU(W_2 \cdot ReLU(W_1 \cdot x)) \tag{9}$$

In Equation (9), W_1, W_2, and W_3 are the weights of the first to third layer point wise convolutions. $S(x)$ is the output of the Sandglass module.

2.2.3. Coordinate Attention Mechanisms and Feature Weighted Representation

CA generates globally perceived weighted features by separating horizontal and vertical information, as shown in Equation (10).

$$CA(x) = \sigma\big(W_h \cdot GAP_h(x)\big) \otimes Expand_h(x) + \sigma\big(W_w \cdot GAP_w(x)\big) \otimes Expand_w(x) \tag{10}$$

In Equation (10), W_h and W_w are the weights of the horizontal and vertical axis features. $GAP_h(x)$ and $GAP_w(x)$ are GAP along the horizontal and vertical directions. $Expand_h(x)$ and $Expand_w(x)$ are feature extension operations along the horizontal and vertical directions. $CA(x)$ is the output of CA. Throughout the process, the target tracking module plays an important role. By tracking the key parts of Archers in real-time, this module can effectively locate targets in dynamic scenes, ensuring that the feature extraction network maintains a high level of attention to key points in motion [5, 15].

2.2.4. Target Tracking Module Structure and Template Update Mechanism

The framework of the target tracking module is displayed in Figure 6.

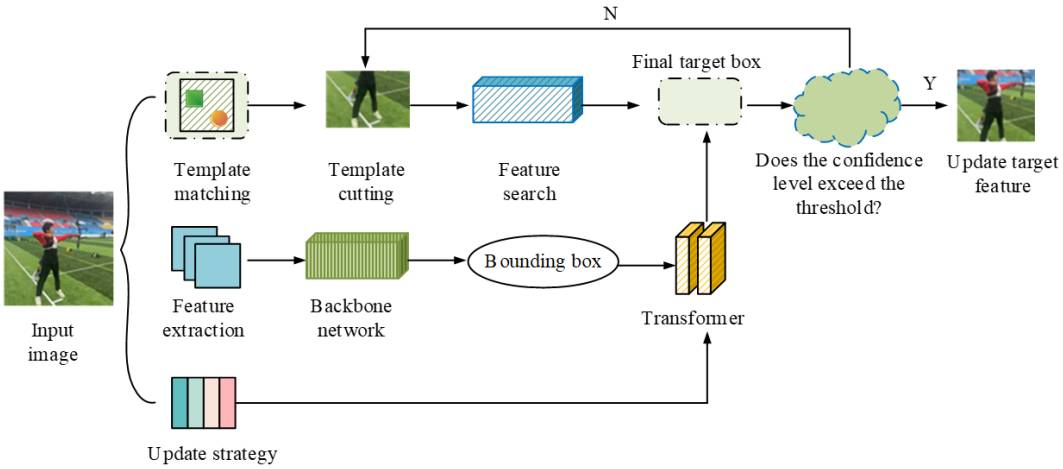


Figure 6. Object tracking module structure.

In Figure 6, the structure of the target tracking module mainly has three parts: Template matching, feature extraction, and update strategy. Firstly, the module prunes candidate regions from the video frame where the target is located by defining a search area, while extracting template regions for matching. The template matching stage adopts a feature level comparison method, extracts key features of the target through the backbone network, and combines the bounding box regression module to generate the initial position of the target. This method is similar to the approach taken by Kaloub and Abed Elgabar [12], who developed an emotion detection system for audio files by combining various machine learning classifiers such as sequence minimization, random forests, K-nearest neighbors, and simple logistic regression. Through a multi-ensemble approach, they selected the optimal feature-level objects from the target set. Next, the feature extraction part further processes the template features and search area features through a transformer encoder, integrating spatial and temporal information to enhance the discriminative capacity of the target. After extracting feature, the bounding box prediction module combines multi-scale features to generate the final target box. The update strategy ensures the tracking performance of the target in consecutive frames, and dynamically updates the template information based on the predicted matching degree between the target box and the template area. When the confidence level of the matching result exceeds a certain threshold, the template features are updated to adapt to the changes of the target during motion. At the same time, the branch module is responsible for determining whether to adjust or switch templates based on the tracking results to adapt to complex scenes such as occlusion, rapid motion, or target disappearance. The expression for template matching is shown in Equation (11).

$$S^*(i,j) = \sum_{u=1}^{H}\sum_{v=1}^{W} \varepsilon(T(u,v)) \cdot \varphi(I(i+u, j+v)) \tag{11}$$

In Equation (11), $S^*(i, j)$ is the similarity score of position (i, j) in the search area. $T(u, v)$ is the template feature at position (u, v). $I(i+u, j+v)$ is the feature value of the search area feature at the offset (i, j) position. ε and φ are both feature encoding functions. When the reliability of the target matching configuration exceeds the set threshold, the template features are updated to adapt to the dynamic changes of the target. The updated formula is shown in Equation (12).

$$T_{new} = \rho \cdot T_{old} + (1-\rho) \cdot F_{curr} \tag{12}$$

In Equation (12), T_{new} is the updated template feature. F_{curr} is the current template feature T_{old}. is the target feature extracted from the current frame. ρ is the Template Update Coefficient (TUC). In summary, this study combines an image deblurring model based on DFPDeblur GAN and an attitude recognition model based on IHRNet to construct a novel AMHPE model. The process of this model is shown in Figure 7.

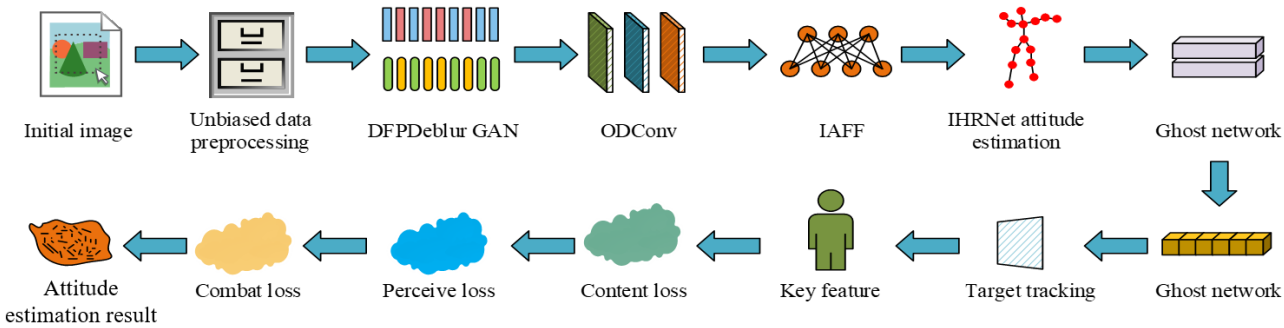


Figure 7. Flow of a new Archery human pose estimation model.

In Figure 7, firstly, the Archery video is preprocessed using the motion blur removal module and target tracking module. The target area is cropped and the quality of the input data is optimized. Then, DFPDeblur GAN is used to eliminate image blur, and multi-scale features are extracted and fused through ODConv and bidirectional FPN. IAFF module is used to focus on key details and generate high-quality images. Subsequently, the deblurred image is input into IHRNet, and multi-resolution features are extracted layer by layer using Ghost, Sandglass modules, and CA-mechanism. By combining the target tracking module to capture dynamic targets in real-time, precise positioning of key joints and capture of motion continuity have been achieved. Finally, a hybrid optimization based on content loss, perceptual loss, and adversarial loss optimizes the robustness and accuracy, providing scientific motion analysis and training support for Archers.

3. Results

This study first established an environment and conducted multidimensional testing using two classic datasets, with deblurring effect, pose estimation accuracy, and robustness as the core indicators. The experimental content covered hyperparameter selection, ablation testing, comparative testing, and multi-scenario simulation experiments. In addition, this study compared advanced human pose estimation algorithms and deblurring algorithms in the field to verify the true effectiveness of the research model.

3.1. Performance Testing of the AMHPE Model

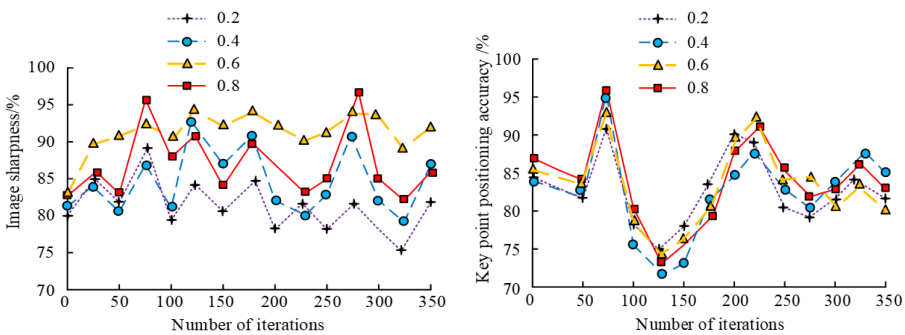
This study establishes a suitable experimental environment and uses the Common Objects in COntext keypoint detection dataset (COCO) and the Max Planck Institute for Informatics human pose dataset (MPII) as test data sources. Among them, the COCO dataset is a widely used human pose estimation dataset. It contains over 250,000 images and over 150,000 human body instances, annotated with 17 key points including head, shoulders, elbows, wrists, hips, knees, ankles, etc. The MPII dataset focuses on pose estimation of human activities, containing over 25,000 images and over 40,000 annotated human instances, covering 410 daily activity scenarios. Each human instance is annotated with 16 key points, making it particularly suitable for capturing dynamic poses and high-precision motion analysis. Table 1 provides detailed configuration parameters.

Table 1. Experimental parameter table.

Experimental equipment	Value
CPU	AMD Ryzen 7 5800H
GPU	NVIDIA RTX 3070
Memory	32GB DDR4
Graphics memory	8GB GDDR6
Development environment	Ubuntu 20.04, Python 3.8
Programming tools	PyTorch 1.12, CUDA 11.5
Initialise learning rate	0.0001
Learning rate batch size	32
Momentum parameters	0.9
Training period	200 epochs
Weight decay	5e-5
Optiming period	250 epochs

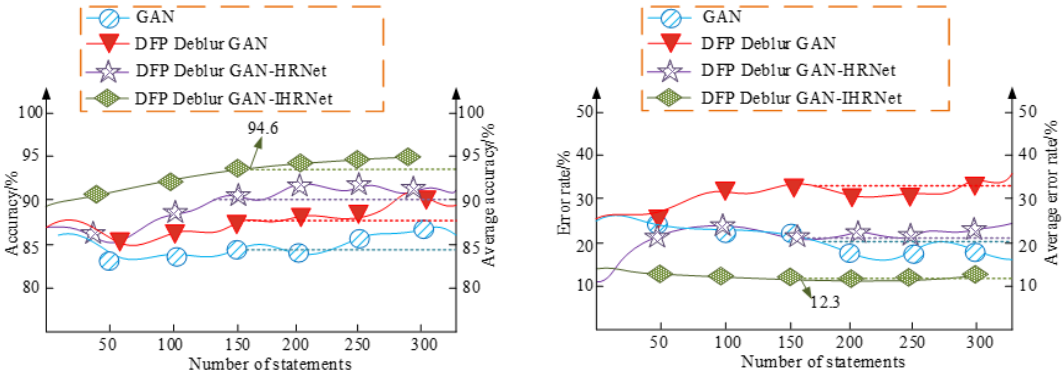
Based on Table 1, this study first conducted value selection tests on two types of hyperparameters that have a significant impact on deblurring and pose estimation, namely FFC and TUC, as shown in Figure 8.

In Figure 8-a), within the iteration range of 100 to 200, the image clarity at an FFC of 0.6 consistently outperforms other coefficient settings, with a stable image clarity of around 95%, demonstrating good stability and clarity improvement effect. When FFC is 0.8, although it performs well in some iteration intervals, the overall clarity is low and not suitable as the best choice. In Figure 8-b), TUC has a significant impact on the accuracy of keypoint localization. When the TUC is close to 0.6 and 0.8, both can maintain a keypoint positioning accuracy of over 94.3%, with the test result with a coefficient of 0.8 slightly better in most cases. In contrast, when TUC is 0.2, the positioning accuracy fluctuates greatly, making it difficult to ensure high stability. Therefore, this study determines that the model performance is optimal when the FFC is 0.6 and the TUC is 0.8. This study continues with ablation testing, as shown in Figure 9.



a) Feature fusion coefficient. b) Template updating coefficient.

Figure 8. Hyperparameter selection test result.



a) Accuracy testing. b) Error rate test.

Figure 9. Ablation test results.

Figure 9-a) shows the accuracy and average accuracy results of the model ablation test; Figure 9-b) shows the error and average error results of the model ablation test. From Figure 9-a), it can be seen that the base GAN model has the lowest accuracy throughout the test and fluctuates greatly, maintaining only between 80% and 85%, indicating that there is an obvious deficiency in its stability and recognition ability in dynamic scenes. After the introduction of the DFPDeblur GAN, the accuracy is significantly increased to about 90%, with an overall improvement of 6.5 percentage points, indicating that the image deblurring module has a positive effect on the enhancement of feature quality. After further overlaying HRNet, the average accuracy stabilizes above 93%, and the fluctuation is significantly reduced, indicating that high-resolution feature modeling effectively improves the attitude estimation accuracy. Finally, the model accuracy reaches 94.6% under the joint optimization of DFPDeblur GAN and IHRNet, which is the best performance, and the overall improvement is nearly 12 percentage points compared with the base GAN. As seen in Figure 9-b), the GAN model has the highest average error rate of about 35%, accompanied by obvious fluctuations, which makes it difficult to ensure the stability of keypoint localization; after the introduction of the DFPDeblur GAN, the error rate decreases to about 21.7%, which suggests that the improvement of the image clarity can help to reduce the recognition bias; after further combining with the HRNet, the error rate decreases to 14.9%, and the final introduction of the improved IHRNet, the average error is minimized to 12.3%, which is a 22.7% reduction in overall error compared to the original model. The above results verify the superimposed contribution of each key module to the performance improvement, and also show that the proposed model has better accuracy and robustness in complex dynamic scenarios. This study introduces other advanced deblurring pose estimation algorithms for comparison, such as Deblur GAN Version 2 (DeblurGAN-v2), Semantic Consistency GAN (SCGAN), and Pose Fixing Network (PoseFix). Table 2 presents test data using Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSIM), and runtime as indicators.

Table 2. Index test results of different models.

Data set	Model	PSNR (dB)	SSIM	Runtime (s)
COCO	DeblurGAN-v2	31.78	0.85	2.13
	SCGAN	33.62	0.87	2.45
	PoseFix	32.49	0.86	2.29
	Research model	34.81	0.91	1.87
MPII	DeblurGAN-v2	30.43	0.84	2.21
	SCGAN	32.56	0.86	2.39
	PoseFix	31.67	0.85	2.34
	Research model	34.32	0.92	1.94

In Table 2, the research model shows significant advantages in PSNR, SSIM, and runtime metrics. On the COCO and MPII datasets, the PSNR of the research model reaches 34.81dB and 34.32, which are about 9.1% and 7.4% higher than DeblurGAN-v2 and PoseMix. The SSIM index also reaches 0.91 and 0.92, far higher than other models. This indicates that it has better performance in image quality and structural restoration. In terms of runtime, the average runtime of the research model on both datasets is less than 2 seconds, which is about 0.5 seconds less than SCGAN and PoseMix, demonstrating its efficiency. Overall, the research model can maintain low computational costs while ensuring high-precision image quality, and has broad application prospects.

3.2. New AMHPE Model Simulation Testing

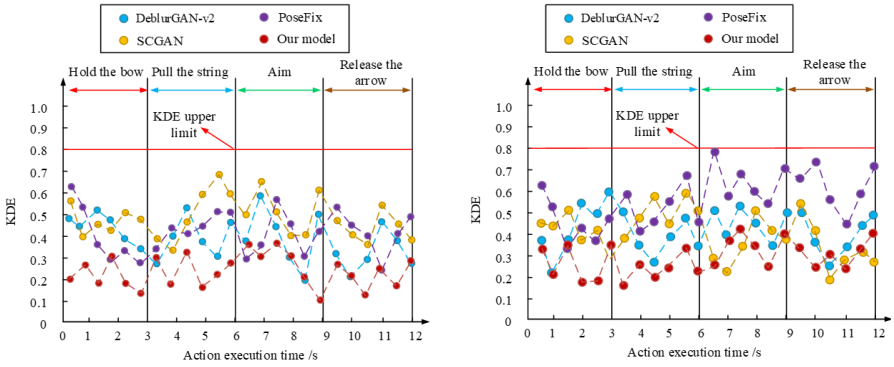
To validate the actual effectiveness of the research model in deblurring and Archery posture estimation, this study randomly selects two images from the MPII dataset for different model comparison tests. Firstly, it is necessary to ensure that these models have undergone image data preprocessing and maintain a certain level of data validity. The test results are shown in Figure 10.

Figures 10-a) to (d) show the deblurring effects of DeblurGAN-v2, SCGAN, PoseFix, and research model. Comparison shows that DeblurGAN-v2 has a good effect on restoring the basic structure of images, but it is slightly lacking in detail processing, with some joint edges still blurred. SCGAN has some improvement in semantic consistency, with overall image clarity higher than DeblurGAN-v2, but some details still appear slightly blurry, especially in complex background areas. PoseMix maintains a certain level of image clarity while correcting key point positions, but has limited ability to

restore blur in dynamic scenes, resulting in noticeable blurry traces in the background. The research model performs the best in deblurring and detail restoration, not only effectively restoring the overall clarity of the image, but also preserving the accuracy of joint positions in complex dynamic scenes. The character contours and background details are clear and natural, significantly better than other comparison models. This study tests four types of standard Archery movements (bow holding, string pulling, aiming, and shooting) and Keypoint Detection Error (KDE) as indicators, as shown in Figure 11.

a) DeblurGAN-v2. b) SCGAN. c) PoseFix. d) Our model.

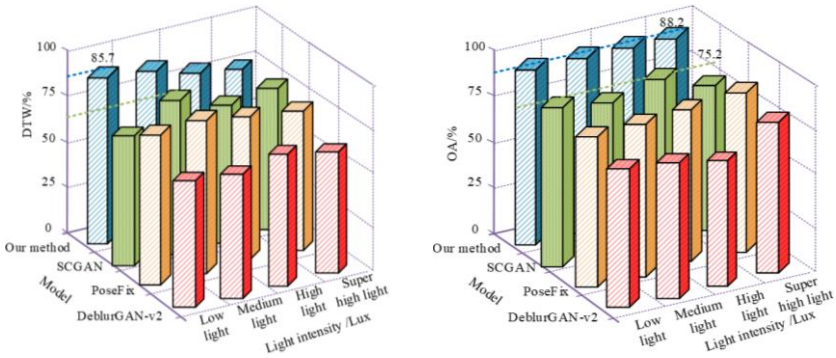
Figure 10. Comparison results of image deblurring effect of different methods.



a) COCO. b) MPII.

Figure 11. KDE test results of Archery pose of different models.

Figures 11-a) and (b) show the KDE detection results of four models on four types of Archery movements in the COCO and MPII datasets. In Figure 11-a), DeblurGAN-v2 has high KDE values in all stages of the action, with some stages, such as string pulling and aiming, approaching the upper limit of KDE at 0.9, indicating low accuracy in locating key points of the action. The SCGAN and PoseMix models perform slightly better in the bow and arrow stages, but there are still significant errors in the string pulling and aiming stages, indicating their limited robustness in dynamic motion capture. The research model maintains a KDE value of around 0.2 in all action stages, demonstrating excellent keypoint localization accuracy and stability. In Figure 11-b), the performance of PoseFix in the MPII dataset is abnormal, while DeblurGAN-v2 and SCGAN are relatively stable, but compared to the research model, they have poorer accuracy and stability in bow holding, string pulling, and aiming. The KDE mean of the research method is 0.23, which has a significant advantage over PoseMix's 0.51. This study uses Dynamic Time Warping (DTW) and Average Overlap (AO) as indicators for action trajectory similarity, as shown in Figure 12.



a) DTW test. b) OA test.

Figure 12. DTW and OA test results of models under different light intensities.

Figures 12-a) and (b) show the DTW and OA test results of four types of models under different light intensities. In Figure 12-a), DeblurGAN-v2 performs poorly under low light intensity conditions, with a DTW value of only about 45%, indicating that its ability to capture motion trajectories is significantly limited by lighting conditions. SCGAN and PoseMix exhibit significant fluctuations in DTW values between 65% and 80% at high light intensities (i.e., 501-2000Lux) and ultra-high light intensities (i.e., greater than 2000Lux), indicating insufficient stability. The research model maintains a high level of DTW values under all lighting intensities, especially at low light intensities where the DTW value reaches 85.7%, demonstrating high robustness and adaptability to motion trajectory capture. In Figure 12-b), DeblurGAN-v2 has the lowest AO value at low light intensity, only about 35.6%. With the increase of light intensity, the AO value has improved, but the overall performance is still unstable. SCGAN and PoseMix perform relatively well under medium light intensity conditions, with AO values reaching 72.3% and 75.2%, but they also exhibit fluctuations under high and ultra-high light intensities. The research model maintains excellent performance under all light intensity conditions, with an AO value consistently above 88.2%, and is able to maintain high accuracy and stability even under ultra-high light intensity. The results based on tracking error, Image Processing Speed (IPS), and mean Average Precision (mAP) are listed in Table 3.

Table 3. Multiple index test results of different models.

Data set	Model	Tracking error/%	IPS/FPS	mAP/%
COCO	DeblurGAN-v2	12.48	23.87	81.76
	SCGAN	10.34	24.92	84.32
	PoseFix	9.67	26.14	85.89
	Our model	7.43	29.61	89.23
MPII	DeblurGAN-v2	13.89	22.45	79.34
	SCGAN	11.52	23.67	82.17
	DEKR	10.11	25.21	84.68
	Our model	8.24	28.76	88.47

In Table 3, in terms of tracking error, the error of the research model on COCO and MPII is 7.43% and 8.24%, which is lower than other models, indicating that it has an advantage in the accuracy of target tracking. On IPS, the processing speed of the research model on two types of datasets is 29.61 Frames Per Second (FPS) and 28.76FPS, which is higher than the comparison model. Especially compared to DeblurGAN-v2's 22.45FPS, IPS has increased by about 30.1%, demonstrating high real-time performance and suitable for application requirements in dynamic scenarios. In terms of mAP, the research model achieves 89.23% on COCO and 88.47% on MPII, which is about 3% and 4% higher than PoseMix and Disentangled Keypoint Regression (DEKR), demonstrating its excellent performance in keypoint localization and attitude estimation. In summary, the research model has shown significant advantages in tracking error, IPS, and mAP, especially in complex dynamic scenarios with higher robustness and application potential.

4. Conclusions

Archery requires extremely high precision in motion capture and pose estimation, but existing technologies still have shortcomings in performance in dynamic scenes. Therefore, this study solved the problems of image blur and keypoint localization in complex dynamic scenes by incorporating fuzzy algorithms and HRNet, and ultimately proposed an AMHPE model based on DFPDeblur GAN. In the experiment, when FFC=0.6 and TUC=0.8, the image clarity of the model remained stable at around 95%, and the accuracy of keypoint localization reached 94.3%. After sequentially incorporating DFPDeblur GAN and IHRNet, the final model achieved an average keypoint detection accuracy of 94.6% and an average error of only 12.3%. Compared with DeblurGAN-v2 and PoseMix, the PSNR of the research model increased by 9.1% and 7.4%, the SSIM reached a maximum of 0.92, and the average runtime was less than 2 seconds. The proposed model performed the best in deblurring and detail restoration for four types of Archery actions: bow holding, string pulling, aiming, and shooting. At the same time, its KDE was the lowest at 0.2, which had a significant advantage over PoseMix's 0.51. Its DTW value reached 85.7% under low light intensity, and its AO value remained above 88.2% under all light intensities. The fastest IPS was 29.61FPS, and the lowest tracking error was 7.43%. In summary, the research model can well lift the accuracy and adaptability of pose estimation in complex dynamic scenes, providing reliable support for action analysis and training optimization in Archery. However, actual Archery testing is affected by differences in athletes' training quality, and this study has not yet tested Archery in a multi person environment. Subsequent research can further incorporate lightweight design to enhance real-time performance and deeply explore the key point capture problem in multi-target scenarios. Meanwhile, the proposed framework shows good versatility in image deblurring and high-precision pose estimation, which can be further extended to other motion-intensive scenarios, such as martial arts action recognition, biomechanical behavior analysis, and public safety video surveillance, to validate the model's adaptability and utility value in a wider range of complex dynamic environments.

References

[1] Abba Haruna A., Muhammad L., and Abubakar M., "Novel Thermal-Aware Green Scheduling in Grid Environment," *Artificial Intelligence and Applications*, vol. 1, no. 4, pp. 244-251, 2022. file:///C:/Users/HP/Downloads/AIA22023321%20(2).pdf

[2] Abhishek A. and Tahir S., "Human Verification over Activity Analysis via Deep Data Mining," *Computers, Materials and Continua*, vol. 75, no. 1, pp. 1391-1409, 2023. https://doi.org/10.32604/cmc.2023.035894

[3] Anvari T., Park K., and Kim G., "Upper Body Pose Estimation Using Deep Learning for a Virtual Reality Avatar," *Applied Sciences*, vol. 13, no. 4, pp. 1-22, 2023. https://doi.org/10.3390/app13042460

[4] Azadjou H., Błazkiewicz M., Erwin A., and Valero-Cuevas F., "Dynamical Analyses Show that Professional Archers Exhibit Tighter, Finer and more Fluid Dynamical Control than Neophytes," *Entropy*, vol. 25, no. 10, pp. 1-14, 2023. https://doi.org/10.3390/e25101414

[5] Beyaz O., Eyraud V., Demirhan G., Akpinar S., and Andrzej P., "Effects of Short-Term Novice Archery Training on Reaching Movement Performance and Interlimb Asymmetries," *Journal of Motor Behavior*, vol. 56, no. 1, pp. 78-90, 2024, https://pubmed.ncbi.nlm.nih.gov/37586703/

[6] Chen L. and Hu D., "An Effective Swimming Stroke Recognition System Utilizing Deep Learning Based on Inertial Measurement Units," *Advanced Robotics*, vol. 37, no. 7, pp. 467-479, 2023. https://doi.org/10.1080/01691864.2022.2160274

[7] Chen Q., Liu H., and Wei D., "The Relationship between Self-Efficacy for Rehabilitation and Kinesiophobia in Elderly Patients with Coronary Heart Disease Intervention," *International Journal of Life Science Study*, vol. 5, no. 1, pp. 28-33, 2024. http://doi.org/10.7508/ijlss.01.2024.28.33

[8] Ding W. and Li W., "High Speed and Accuracy of Animation 3D Pose Recognition Based on an Improved Deep Convolution Neural Network," *Applied Science*, vol. 13, no. 13, pp. 1-17, 2023. https://www.mdpi.com/2076-3417/13/13/7566

[9] Dominguez G., Alvarez E., Cordoba A., and Reina D., "A Comparative Study of Machine Learning and Deep Learning Algorithms for Padel Tennis Shot Classification," *Soft Computing*, vol. 27, no. 17, pp. 12367-12385, 2023. https://link.springer.com/article/10.1007/s00500-023-07874-x

[10] He Z., Sun Y., and Zhang Z., "Human Activity Recognition Based on Deep Learning Regardless of Sensor Orientation," *Applied Sciences*, vol. 14, no. 9, pp. 1-21, 2024. https://www.mdpi.com/2076-3417/14/9/3637

[11] Jun H., "Presentation of a Method for Removal of Motion Blur Effect in Images by Using GAN," *Journal of Optics*, vol. 53, no. 3, pp. 2469-2480, 2024. DOI: 10.1007/s12596-023-01408-2

[12] Kaloub A. and Abed Elgabar E., "Speech-based Techniques for Emotion Detection in Natural Arabic Audio Files," The International Journal of Information Technology, vol. 22, no. 1, pp. 139-157, 2025. https://doi.org/10.34028/iajit/22/1/11

[13] Khan R., Luo Y., and Wu F., "Multi-Scale GAN with Residual Image Learning for Removing Heterogeneous Blur," *IET Image Processing*, vol. 16, no. 9, pp. 2412-2431, 2022. https://doi.org/10.1049/ipr2.12497

[14] Kong X., Liu Y., Han R., Li S., and Liu H., "Forest Fire Image Deblurring Based on Spatial-Frequency Domain Fusion," *Forests*, vol. 15, no. 6, pp. 1-18, 2024. https://doi.org/10.3390/f15061030

[15] Lau J., Ghafar R., Zulkifli E., Hashim H., and Sakim H., "Comparison of Shooting Time Characteristics and Shooting Posture between High-and Low-Performance Archers," *Annals of Applied Sport Science*, vol. 11, no. 2, pp. 1-9, 2023. https://aassjournal.com/article-1-1115-en.html

[16] Lee J., Gwon G., Kim I., and Jung H., "A Motion Deblurring Network for Enhancing UAV Image Quality in Bridge Inspection," *Drones*, vol. 7, no. 11, pp. 1-20, 2023. https://doi.org/10.3390/drones7110657

[17] Lee R., Sivakumar S., and Lim K., "Review on Remote Heart Rate Measurements Using Photoplethysmography," *Multimedia Tools and Applications*, vol. 83, no. 15, pp. 44699-44728, 2024. https://doi.org/10.1007/s11042-023-16794-9

[18] Liu Y., Ai H., Xing J., Li X., and et al., "Advancing Video Synchronization with Fractional Frame Analysis: Introducing a Novel Dataset and Model," *in Proceedings of the 38th AAAI Conference on Artificial Intelligence and 36th Conference on Innovative Applications of Artificial Intelligence and 40th Symposium on Educational Advances in Artificial Intelligence*, Vancouver, pp. 3828-3836, 2024. https://doi.org/10.1609/aaai.v38i4.28174

[19] Luo X., Wu Y., and Wang F., "Target Detection Method of UAV Aerial Imagery Based on Improved YOLOv5," *Remote Sensing*, vol. 14, no. 19, pp. 1-25, 2022. https://www.mdpi.com/2072-4292/14/19/5063/review_report

[20] Lv Y., Li M., and Li B., "Athletic Sports Posture Measurement Algorithm Based on Multi-Sensor Combination," *Journal of Computational Methods in Sciences and Engineering*, vol. 22, no. 6, pp. 2065-2076, 2022. https://journals.sagepub.com/doi/10.3233/JCM-226393

[21] Maleki S., Raman A., Cheng Y., Crassidis J., and Schmid M., "Optimal Pose Estimation and Covariance Analysis with Simultaneous

Localization and Mapping Applications," *Journal of Guidance, Control, and Dynamics*, vol. 47, no. 2, pp. 187-202, 2024. https://arc.aiaa.org/doi/10.2514/1.G007301

[22] Okoro Y., Ayo-Farai O., Maduka C., Okongwu C., and Sodamade O., "Emerging Technologies in public Health Campaigns: Artificial Intelligence and Big Data," *Acta Informatica Malaysia*, vol. 8, no. 1, pp. 5-10, 2024. DOI: 10.26480/aim.01.2024.05.10

[23] Prasetyo Y., Pamungkas O., Prasetyo H., and Susanto S., "Analysis of Anthropometry, Physical Conditions, and Archery Skills as the Basis for Identification of Talent in the Sport of Arrow," *Sports Science and Health*, vol. 24, no. 2, pp. 183-188, 2022. https://doi.org/10.7251/SSH2202183P

[24] Romero A., Carvalho P., Corte-Real L., and Pereira A., "Synthesizing Human Activity for Data Generation," *Journal of Imaging*, vol. 9, no. 10, pp. 1-16, 2023. https://doi.org/10.3390/jimaging9100204

[25] Shen C. and Sun Z., "Research on Target Localization Recognition of Automatic Mobile Ball-Picking Robot," *Journal of Optics*, vol. 51, no. 4, pp. 866-873, 2022. https://link.springer.com/article/10.1007/s12596-021-00805-9

[26] Shin M., Lee D., Chung A., and Kang Y., "When Taekwondo Meets Artificial Intelligence: The Development of Taekwondo," *Applied Sciences*, vol. 14, no. 7, pp. 1-24, 2024. https://doi.org/10.3390/app14073093

[27] Siaw T., Han Y., and Wong K., "A Low-Cost Marker-based Optical Motion Capture System to Validate Inertial Measurement Units," *IEEE Sensors Letters*, vol. 7, no. 2, pp. 1-4, 2023. DOI: 10.1109/LSENS.2023.3239360

[28] Song J., Kim K., and Park J., "Multi-Muscle Synergies of Postural Control in Self-and External-Triggered Force Release During Simulated Archery Shooting," *Journal of Motor Behavior*, vol. 55, no. 3, pp. 289-301, 2023. https://doi.org/10.1080/00222895.2023.2187336

[29] Tokmakci H., Ozgur S., and Varol T., "Anxiety Sensitivity, Stress, and Postural Control: Their Implications on Archery Performance in 11-14-year-Olds," *Human Movement*, vol. 24, no. 4, pp. 80-89, 2023. https://doi.org/10.5114/hm.2023.133921

[30] Tong J. and Wang F., "Basketball Sports Posture Recognition Technology Based on Improved Graph Convolutional Neural Network," *Journal of Advanced Computational Intelligence and Intelligent Informatics*, vol. 28, no. 3, pp. 552-561, 2024. https://doi.org/10.20965/jaciii.2024.p0552

[31] Vancurik S. and Callahan D., "Detection and Identification of Choking Under Pressure in College Tennis Based Upon Physiological Parameters, Performance Patterns, and Game Statistics," *IEEE Transactions on Affective Computing*, vol. 14, no. 3, pp. 1942-1953, 2022. https://ieeexplore.ieee.org/document/9750870

[32] Varela L., Boucheron L., Sandoval S., Voelz D., and Siddik A., "Estimation of Motion Blur Kernel Parameters Using Regression Convolutional Neural Networks," *Journal of Electronic Imaging*, vol. 33, no. 2, pp. 1-16, 2024. https://doi.org/10.1117/1.JEI.33.2.023062

[33] Xiao Y., Zhang J., Chen W., Wang Y., and et al., "SR-DeblurUGAN: An End-to-End Super-Resolution and Deblurring Model with High Performance," *Drones*, vol. 6, no. 7, pp. 1-15, 2022. https://doi.org/10.3390/drones6070162

[34] Xu W. and Zhu Z., "Estimation for Human Motion Posture and Health Using Improved Deep Learning and Nano Biosensor," *International Journal of Computational Intelligence Systems*, vol. 16, no. 1, pp. 55-57, 2023. https://link.springer.com/article/10.1007/s44196-023-00239-0

[35] Yao J., Fan X., Li B., and Qin W., "Adverse Weather Target Detection Algorithm Based on Adaptive Color Levels and Improved YOLOv5," *Sensors*, vol. 22, no. 21, pp. 1-21, 2022. https://doi.org/10.3390/s22218577

[36] Zhao C., Li B., and Guo K., "Adaptive Enhancement Design of Non-Significant Regions of a Wushu Action 3D Image Based on the Symmetric Difference Algorithm," *Mathematical Biosciences and Engineering*, vol. 20, no. 8, pp. 14793-14810, 2023. DOI: 10.3934/mbe.2023662

Xiao Yan obtained a Bachelor's degree in Education from the School of Physical Education at Shanxi Normal University, China in 2017. She pursued a Master's degree at the School of Physical Education, Shanxi University, majoring in Physical Education and Training Science, China from 2017 to 2020. In 2020, she was admitted to the integrated Master's and Doctoral Studies at the School of Physical Education of Shanxi University. Her areas of interest are Archery Sports, Deep Learning and Computer Vision Application.

Zhuohan Wu obtained a Bachelor's degree in Education from the School of Physical Education at Xi'an Technological University, China in 2023. In 2024, he was admitted to the Master's Studies at the School of Physical Education of Shanxi University, His areas of interest are Archery Sports, Deep Learning and Computer Vision Application.

Ting Wang obtained his PhD in Sports Science in (2015) from Shanxi University, Tai Yuan, China. Presently, he is working as an associate professor at the School of Physical Education, Shanxi University. He used to be an Assistant Researcher at the Shanxi Provincial Institute of Sports Science. Since 2008, he has been invited to serve as the chief person in charge of scientific research projects for the Chinese National Archery team. He has published 8 academic articles and presided over 6 scientific research projects. His areas of interest include archery sports, deep learning and computer vision application.