Journal of Computing and Electronic Information Management
ISSN: 2413-1660 | Vol. 22, No. 2, 2026

Physics-Guided One-Dimensional Convolutional Attention Network for Badminton Pose-Trajectory Refinement

Yiwen Gao *

Southwest Minzu University, Chengdu 610000, China
*** Corresponding author**

Abstract: Frame-wise 2D pose estimation produces jitter and anatomically inconsistent trajectories that can distort badminton motion analysis. We propose a physics-guided 1D CNN-CBAM denoising autoencoder for 64-frame, 17-joint pose sequences. The network receives normalized coordinates and confidence values and is trained against Butterworth-smoothed pseudo-labels using reconstruction, temporal-variation, bone-length, and acceleration losses. On a fixed 20-sequence validation split, the model reduced acceleration jitter from 0.032796 to 0.008095 and bone-length variance from 0.006314 to 0.001151 relative to raw trajectories. Compared with a five-frame moving average, acceleration jitter was similar, whereas bone-length variance was 80.60% lower. Ablations confirmed that bone and temporal constraints addressed distinct trajectory defects. The higher pseudo-label MSE than the moving average indicates a trade-off between filter agreement and structural consistency. These results support the model as a physics-consistent preprocessing stage, pending participant-independent evaluation and external pose ground truth.

Keywords: Human pose estimation; Badminton; Trajectory refinement; Denoising autoencoder; Convolutional block attention; Physics-guided learning.

1. Introduction

Camera-based human pose estimation (HPE) has made markerless motion analysis increasingly accessible for sports and physical exercise. Reviews of this field show that general-purpose pose estimators are commonly adapted to sport-specific tasks, while occlusion, fast motion, viewpoint variation, and limited public data remain persistent barriers [1,2]. Badminton is a particularly demanding setting because upper-limb motion is rapid, the racket-side wrist may be temporarily occluded, and small frame-to-frame keypoint errors propagate into joint angles, velocities, acceleration profiles, and event timing.

Recent badminton studies have used pose estimation for teaching support and skeleton-based keyframe detection [3,4]. MultiSenseBadminton further demonstrates the value of combining motion and physiological signals for forehand-clear and backhand-drive assessment [5]. These studies motivate automated coaching, but they also expose a practical gap between frame-wise keypoint detection and reliable trajectory-level analysis. A sequence can appear visually plausible in individual frames while containing high-frequency jitter, changing limb lengths, or an unstable estimate of the striking phase.

Classical smoothing filters provide a transparent remedy, and markerless kinematic workflows often include interpolation and filtering before downstream analysis [6]. Nevertheless, a fixed filter treats every joint and motion phase similarly and does not learn correlations among confidence, body structure, and temporal context. Denoising autoencoders (DAEs) instead learn to recover a cleaner representation from corrupted inputs [8]. Attention mechanisms such as the convolutional block attention module (CBAM) can further reweight informative channels and locations with low additional complexity [7]. Skeleton reconstruction research also suggests that temporal or categorical constraints can improve learned motion representations [8].

This work develops a lightweight trajectory-refinement model for 17-keypoint badminton pose sequences. The model uses residual one-dimensional convolutions to preserve temporal resolution, a channel-temporal adaptation of CBAM to emphasize informative joints and frames, and a composite objective that balances pseudo-label reconstruction with temporal smoothness, acceleration regularity, and bone-length stability. The model is intended as the denoising stage of a larger forehand-clear evaluation pipeline, before angle extraction and sequence comparison.

The contributions are threefold. First, the study provides an end-to-end preprocessing and learning pipeline for variable-length two-dimensional badminton skeleton sequences. Second, it introduces an interpretable physics-guided objective whose components correspond to reconstruction, first-order temporal variation, second-order acceleration, and kinematic consistency. Third, it evaluates the full model against raw trajectories and a moving-average baseline and reports component ablations. Claims are deliberately restricted to the current 100-sequence dataset and the Butterworth-derived pseudo-label protocol.

2. Related Work

2.1. Pose estimation for sports and badminton

Deep HPE estimates body-joint locations from images or videos and supports applications ranging from surveillance to sport analysis and joint-angle extraction [2,9]. In sport, markerless systems reduce setup burden, but their reliability depends on camera geometry, visibility, motion speed, and the accuracy of the underlying detector [1]. Pose2Sim illustrates how filtering and geometric processing can convert camera-

derived keypoints into robust kinematic estimates in multi-camera settings [6]. For single-camera applications, post-processing remains especially important because depth ambiguity and transient misses cannot be resolved geometrically. Representative systems such as OpenPose and HRNet established effective multi-person association and high-resolution keypoint representations [10,11]. Single-camera and smartphone-based studies further demonstrated the potential of video-derived landmarks for quantitative movement analysis [12,13], while comparisons with marker-based motion capture showed both practical value and configuration-dependent errors [14,15]. DeepLabCut and its multi-view extension also illustrate the broader development of customizable markerless tracking pipelines [16,17].

Badminton-specific work has focused on teaching interfaces, pose and joint-angle extraction, and the detection of semantically meaningful keyframes [3,4,9]. Sarwar et al. segmented badminton actions into start, ready, strike, and end phases using skeleton information [4], emphasizing the importance of temporally stable joint trajectories. The current study addresses an earlier stage: it refines the pose trajectory itself so that later event detection and biomechanical features are less sensitive to frame-wise noise.

2.2. Denoising autoencoders and attention

A DAE learns to reconstruct an uncorrupted target from a perturbed input and can therefore capture dependencies that are not encoded by pointwise filters [8]. Although the original DAE framework was developed for representation learning, reconstruction-based objectives have also been applied to skeleton sequences [8]. The present model differs from action-recognition DAEs by directly outputting refined joint coordinates and by incorporating losses defined on motion derivatives and limb-length variation. Temporal convolution has likewise proved effective for learning motion dependencies from pose sequences [18].

CBAM sequentially estimates channel and spatial attention maps from pooled feature statistics [7]. For one-dimensional pose trajectories, the spatial branch is naturally reinterpreted as temporal attention. Channel attention identifies informative latent motion channels, whereas temporal attention highlights frames that contribute strongly to reconstruction. This adaptation preserves the lightweight character of CBAM and remains compatible with residual temporal convolutions.

3. Materials and Methods

3.1. Dataset and pose-sequence preprocessing

The computational dataset comprised 100 video-derived badminton pose sequences stored in the project repository. Raw sequence lengths ranged from 45 to 324 frames, with a median of 120 frames. A lightweight YOLO pose checkpoint generated 17 COCO-format keypoints per frame. Each keypoint contained horizontal position, vertical position, and confidence, yielding a 51-dimensional frame vector. The processing files did not encode participant demographics or participant identifiers; therefore, the present split supports sequence-level validation but not participant-independent generalization.

Coordinates were treated as missing when confidence was below 0.3 or either coordinate was zero. Each coordinate trajectory was independently completed with a piecewise cubic Hermite interpolator; values before the first and after the last valid observation were held at the nearest valid endpoint. For translation and scale normalization, the midpoint between the left and right hips defined the pelvis origin, and the distance between the shoulder midpoint and pelvis midpoint defined torso length. Frames with invalid torso estimates were replaced by the median valid torso length, and extreme torso values were bounded to 0.3-3.0 times the median. Coordinates were then clipped to [-10, 10]. Finally, each sequence was resampled on normalized time to 64 frames; confidence values were clipped to [0, 1].

3.2. Pseudo-label construction and corruption process

The 34 coordinate channels were separated from the 17 confidence channels. Because independent motion-capture ground truth was unavailable, the reconstruction target was a pseudo-label generated by a fourth-order zero-phase Butterworth low-pass filter. After temporal normalization to 64 frames over an approximately three-second action window, the assumed sampling frequency was 64/3 Hz and the cutoff frequency was 4.5 Hz. If Butterworth filtering was unavailable or failed, a symmetric five-tap kernel [1, 2, 3, 2, 1]/9 was used as a fallback.

During training, independent Gaussian noise with standard deviation 0.03 was added to coordinate channels only. Confidence channels remained unchanged so that the network could use detector confidence as an input reliability cue. The 80 physical training sequences were exposed ten times per epoch with newly sampled noise, producing stochastic corrupted views without pretending that the project contained additional independent subjects.

3.3. Network architecture

The input tensor had shape 51 x 64 and the output had shape 34 x 64. The encoder first applied a kernel-5 convolution from 51 to 64 channels, batch normalization, rectified linear activation, and a two-convolution residual block. A kernel-3 convolution then expanded the representation to 128 channels, followed by another residual block. Replicate padding preserved sequence length and reduced boundary artifacts.

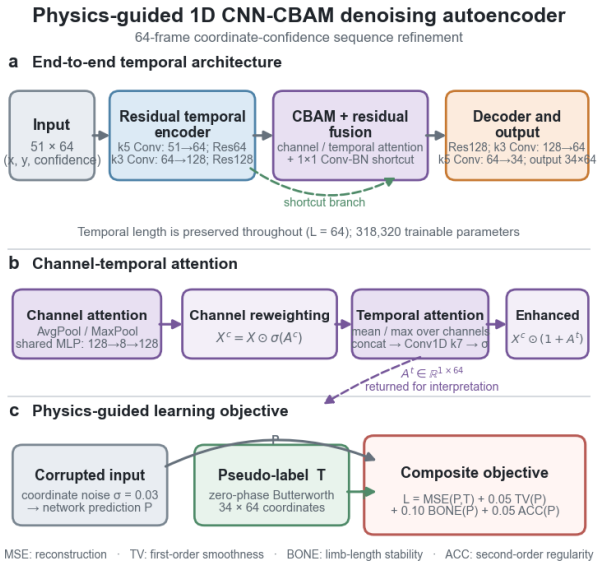


Fig. 1 Architecture of the physics-guided 1D CNN-CBAM denoising autoencoder. The model preserves the 64-frame temporal resolution and maps 51 coordinate-confidence channels to 34 refined coordinate channels.

At the bottleneck, channel attention combined global average- and max-pooled descriptors through a shared two-layer 1 x 1 convolutional multilayer perceptron with reduction ratio 16. Temporal attention pooled the enhanced feature map across channels by average and maximum operations, concatenated the two signals, and applied a kernel-7 convolution. The temporal response was used residually as x(1 + a), after which a 1 x 1 shortcut was added. The decoder contained a 128-channel residual block, a kernel-3 convolution to 64 channels, and a kernel-5 convolution to 34 coordinate channels. The resulting network contained 318,320 trainable parameters. As shown in Fig. 1, the network retains the 64-frame temporal resolution from input to output.

3.4. Physics-guided objective

The total objective combined four terms. Reconstruction loss measured mean-squared error between the prediction P and pseudo-label T. Total variation penalized the mean absolute first difference over time. Acceleration loss penalized the mean absolute second difference. Bone-length loss measured temporal variance for eight upper- and lower-limb segments: shoulder-elbow, elbow-wrist, hip-knee, and knee-ankle on both sides.

$$\mathrm{MSE}(P,T) = \mathrm{mean}[(P - T)^2] \quad (1)$$

$$\mathrm{TV}(P) = \mathrm{mean}_t\,|P(t) - P(t-1)| \quad (2)$$

$$\mathrm{ACC}(P) = \mathrm{mean}_t\,|P(t) - 2P(t-1) + P(t-2)| \quad (3)$$

$$\mathrm{BONE}(P) = \mathrm{sum}_b\,\mathrm{Var}_t[\mathrm{length}_b] \quad (4)$$

$$L = \mathrm{MSE} + 0.05\mathrm{TV} + 0.10\mathrm{BONE} + 0.05\mathrm{ACC} \quad (5)$$

The reconstruction term anchors the output to the filtered target, while the remaining terms encode complementary priors. Total variation suppresses frame-to-frame oscillation, acceleration regularization limits abrupt changes in velocity, and bone-length variance discourages anatomically implausible deformation. These priors do not establish absolute biomechanical correctness; they regularize trajectories within the normalized two-dimensional representation. This use of domain constraints as soft regularizers is consistent with the broader physics-informed machine-learning literature [19,20].

3.5. Training protocol

Files were grouped by filename prefix and split with seed 42 into 80 training and 20 validation sequences. The model was trained for 250 epochs with batch size 16 using AdamW, an initial learning rate of 3 x 10^-4, and weight decay of 1 x 10^-5. A ReduceLROnPlateau scheduler halved the learning rate after 20 validation epochs without improvement. Gradient norm was clipped at 1.0. Validation inputs were not corrupted, and the checkpoint with the lowest total validation loss was retained.

3.6. Baselines, metrics, and ablations

The full model was compared with the raw normalized coordinates and a centered five-frame moving average. The Butterworth output was retained as the pseudo-label reference. Evaluation used mean-squared error to the pseudo-label, mean absolute first difference (temporal variation), mean absolute second difference (acceleration jitter), mean temporal bone-length variance, and absolute anchor displacement in frames. The anchor was defined as the maximum racket-wrist height using a single sample-level handedness assignment.

Four ablations were trained under the same data split: removal of CBAM, removal of bone-length loss, removal of both temporal-variation and acceleration losses, and reconstruction-only training. Ablation models were trained for 120 epochs. Results are descriptive means and standard deviations over the 20 validation sequences; no inferential statistics were used because only one data split and one random seed were evaluated.

4. Results

4.1. Training convergence

Total validation loss decreased from 0.048041 in epoch 1 to its minimum of 0.007451 at epoch 226. At that checkpoint, reconstruction loss was 0.005173, temporal-variation loss was 0.018756, bone-length loss was 0.009358, and acceleration loss was 0.008095 before weighting. The final epoch produced a similar validation loss of 0.007529, indicating a stable plateau rather than divergence. The training loss was lower than validation loss after convergence, which is expected under the fixed sequence-level split and stochastic coordinate corruption. As shown in Fig. 2, both curves reach a stable plateau after the initial decline.

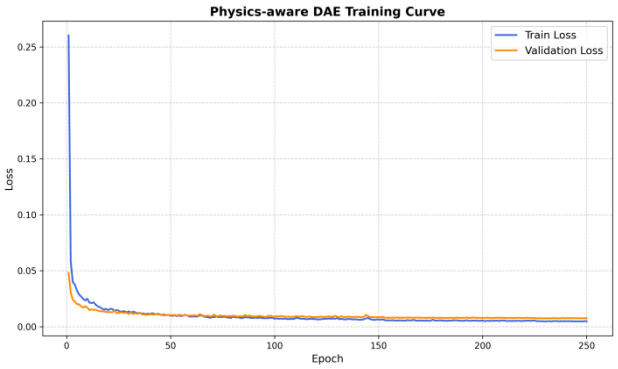


Fig. 2 Training and validation loss over 250 epochs. The best checkpoint was selected at epoch 226 using total validation loss.

4.2. Comparison with raw and moving-average trajectories

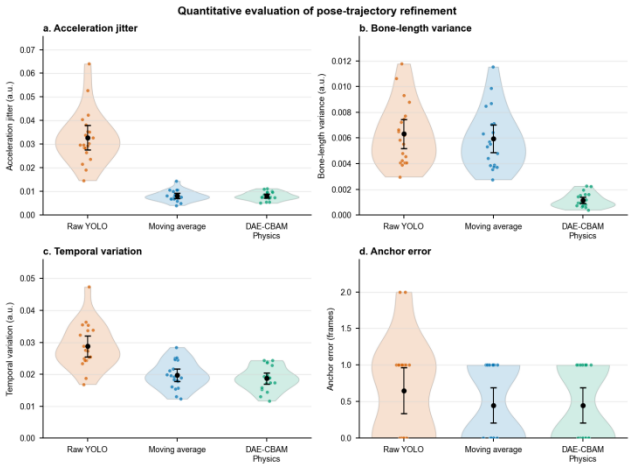


Fig. 3 Per-sequence distributions for acceleration jitter, bone-length variance, temporal variation, and anchor error. Black markers show the mean and uncertainty bars used in the project visualization.

Table 1 and Fig. 3 summarize validation performance. Relative to raw pose trajectories, the full model reduced acceleration jitter by 75.32%, temporal variation by 34.83%, and bone-length variance by 81.76%. Relative to the moving average, acceleration jitter was nearly unchanged (0.43%

lower) and temporal variation was 5.01% lower, whereas bone-length variance was 80.60% lower. Mean anchor error decreased from 0.65 frames for raw trajectories to 0.45 frames and matched the moving-average baseline.

The model did not minimize every metric. Its mean-squared error to the Butterworth pseudo-label was 0.005173, compared with 0.000481 for the moving average and 0.000630 for raw coordinates. This difference is important because the pseudo-label is a filtered transform of the same input rather than independent ground truth. The moving average can therefore closely approximate that target without learning structural correction, whereas the proposed model trades direct pseudo-label agreement for substantially improved limb-length consistency.

Table 1. Validation metrics (mean ± standard deviation; n = 20). Lower is better.

Method	MSE	TV	Acc. jitter	Bone var.	Anchor err. (frames)
Proposed	0.005173 ± 0.002806	0.018756 ± 0.003614	0.008095 ± 0.001697	0.001151 ± 0.000537	0.45 ± 0.50
Moving avg.	0.000481 ± 0.000265	0.019746 ± 0.004146	0.008129 ± 0.002243	0.005935 ± 0.002258	0.45 ± 0.50
Raw YOLO	0.000630 ± 0.000834	0.028779 ± 0.006778	0.032796 ± 0.010961	0.006314 ± 0.002351	0.65 ± 0.65

4.3. Ablation analysis

Table 2. Ablation metrics (mean; n = 20). Lower is better.

Variant	MSE	TV	Acc. jitter	Bone var.	Anchor error
Full model	0.005173	0.018756	0.008095	0.001151	0.45
w/o CBAM	0.005755	0.018801	0.008396	0.001282	0.40
w/o bone	0.004954	0.019302	0.009012	0.006141	0.75
w/o TV + acc.	0.005770	0.021395	0.012636	0.001247	0.40
Recon. only	0.005548	0.022352	0.013679	0.006579	1.40

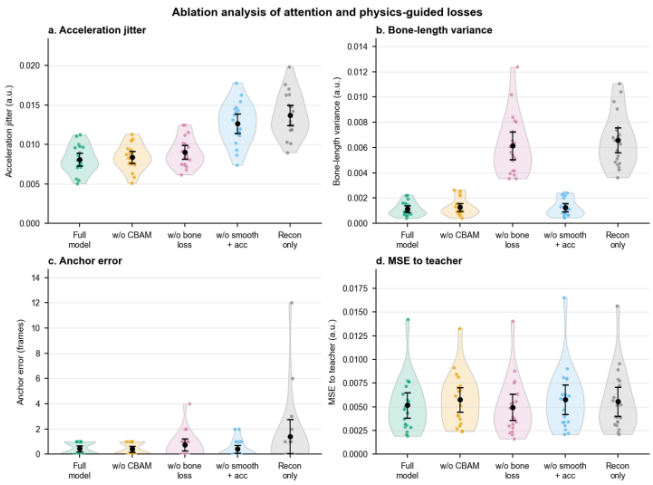


Fig. 4 Ablation analysis of attention and physics-guided losses. The largest anchor-error excursions occur in less constrained variants, particularly reconstruction-only training.

The ablation results separated the effects of the model components. Removing CBAM increased pseudo-label MSE from 0.005173 to 0.005755 and modestly increased acceleration jitter and bone variance. Removing bone-length loss increased bone variance by 433.32%, from 0.001151 to 0.006141. Removing temporal-variation and acceleration losses increased acceleration jitter by 56.10%. Reconstruction-only training produced the largest mean acceleration jitter, bone variance, and anchor error among the ablations. These results indicate that the physics-guided terms control different failure modes rather than acting as interchangeable smoothness penalties. The aggregate ablation metrics and per-sequence distributions are shown in Table 2 and Fig. 4, respectively

4.4. Anchor-error outliers

Anchor-error excursions were concentrated in the less constrained ablations. The largest displacement was 12 frames for reconstruction-only training, compared with maxima of 1 frame for the full model and the model without CBAM, 4 frames without bone-length loss, and 2 frames without temporal smoothness and acceleration losses. This pattern indicates that the large errors were not caused by CBAM. The current anchor is the hard argmax of racket-wrist height; when two local peaks are similar, small trajectory changes can switch the selected temporal lobe and produce a large discrete error. Anchor error was therefore treated as a diagnostic metric rather than a direct measure of pose accuracy. A phase-aware soft-argmax or a multi-cue anchor using wrist height, speed, and motion energy should be evaluated in future work.

5. Discussion

The principal finding is that the compact 1D CNN-CBAM model improved limb-geometry stability while preserving the temporal smoothness of frame-wise badminton pose trajectories. A five-frame moving-average baseline removed most acceleration jitter but produced little improvement in bone-length variance. By contrast, the proposed model achieved comparable acceleration smoothness and reduced bone-length variance by approximately fivefold relative to the moving-average baseline. This contrast suggests that learned cross-joint dependencies and structural constraints provide information that independent coordinate smoothing does not capture.

The ablation results further indicate that the loss terms address complementary trajectory defects. Removing bone-length regularization primarily increased geometric variability, whereas removing the temporal penalties

increased first- and second-order motion roughness. CBAM produced smaller but consistent improvements in reconstruction and structural metrics. Together, these findings position the full model as a balanced operating point across competing objectives, rather than as the optimum for any single metric.

However, the proposed model had a higher MSE to the Butterworth pseudo-label than the moving-average baseline. This result requires careful interpretation because the pseudo-label is itself a filtered transformation of the raw trajectory. Accordingly, MSE quantifies agreement with the selected filter rather than accuracy relative to true joint positions. Deviation from the pseudo-label may therefore coexist with improved smoothness and bone-length consistency, but it cannot be taken as evidence of greater anatomical accuracy. Without motion-capture references or expert-corrected landmarks, the present evidence supports only a narrower conclusion. Under the reported normalization and validation protocol, the model produces smoother and more bone-consistent two-dimensional trajectories.

The limited dataset also constrains the generalizability of these findings. Stochastic Gaussian corruption provides multiple noisy realizations of each sequence, but it does not introduce new participants, viewpoints, or movement styles. Because the filename-based split operates at the sequence level, clips from the same participant may occur in both the training and validation sets. This overlap could allow participant-specific movement patterns to influence validation performance. Future validation should therefore use participant-level partitioning, repeated participant-independent cross-validation, and multiple random seeds.

Other limitations arise from the use of normalized two-dimensional coordinates, the omission of racket and shuttlecock trajectories, restriction to one stroke family, and the fixed 64-frame representation. The present pipeline was evaluated offline and uses zero-phase filtering to construct its training targets. Validation against manually corrected or motion-capture reference trajectories is therefore a priority. Further experiments should test robustness to synthetic occlusion and viewpoint changes, uncertainty-aware confidence corruption, and comparisons with Savitzky-Golay, Kalman, and stronger sequence-learning baselines.

6. Conclusion

This study presented a physics-guided 1D CNN-CBAM denoising autoencoder for refining 17-keypoint badminton pose trajectories. The model combines stochastic coordinate corruption, a residual temporal encoder-decoder, channel-temporal attention, and reconstruction, temporal, acceleration, and bone-length objectives. On 20 held-out sequences, it reduced acceleration jitter, temporal variation, and bone-length variance relative to raw trajectories; its principal advantage over a moving average was bone-length stability. Ablations confirmed that bone and temporal penalties addressed distinct trajectory defects. The model should currently be regarded as a physics-consistent preprocessing component, not as a validated substitute for ground-truth kinematic measurement.

References

[1] Badiola-Bengoa, A., & Mendez-Zorrilla, A. (2021). A systematic review of the application of camera-based human pose estimation in the field of sport and physical exercise. Sensors, 21(18), 5996. https://doi.org/10.3390/s21185996

[2] Samkari, E., Arif, M., Alghamdi, M., & Al Ghamdi, M. A. (2023). Human pose estimation using deep learning: A systematic literature review. Machine Learning and Knowledge Extraction, 5(4), 1612-1659. https://doi.org/10.3390/make5040081

[3] Zhang, X., Zhao, P., & Cao, Y. (2022). Research on badminton teaching technology based on human pose estimation algorithm. Scientific Programming, 2022, 4664388. https://doi.org/10.1155/2022/4664388

[4] Sarwar, M. A., Lin, Y. C., Daraghmi, Y. A., Ik, T. U., & Li, Y. L. (2023). Skeleton based keyframe detection framework for sports action analysis: Badminton smash case. IEEE Access, 11, 90891-90900. https://doi.org/10.1109/ACCESS.2023.3307620

[5] Seong, M., Kim, G., Yeo, D., Kang, Y., Yang, H., DelPreto, J., Matusik, W., Rus, D., & Kim, S. (2024). MultiSenseBadminton: Wearable sensor-based biomechanical dataset for evaluation of badminton performance. Scientific Data, 11, 343. https://doi.org/10.1038/s41597-024-03144-z

[6] Pagnon, D., Domalain, M., & Reveret, L. (2021). Pose2Sim: An end-to-end workflow for 3D markerless sports kinematics: Part 1: Robustness. Sensors, 21(19), 6530. https://doi.org/10.3390/s21196530

[7] Woo, S., Park, J., Lee, J. Y., & Kweon, I. S. (2018). CBAM: Convolutional block attention module. In Proceedings of the European Conference on Computer Vision (ECCV) (pp. 3-19). Springer. https://doi.org/10.1007/978-3-030-01234-2_1

[8] Wu, Z., Weise, T., Zou, L., Sun, F., & Tan, M. (2020). Skeleton based action recognition using a stacked denoising autoencoder with constraints of privileged information. arXiv, arXiv:2003.05684. https://doi.org/10.48550/arXiv.2003.05684

[9] Wang, S., Zhang, X., Ma, F., Li, J., & Huang, Y. (2023). Single-stage pose estimation and joint angle extraction method for moving human body. Electronics, 12(22), 4644. https://doi.org/10.3390/electronics12224644

[10] Cao, Z., Hidalgo, G., Simon, T., Wei, S. E., & Sheikh, Y. (2021). OpenPose: Realtime multi-person 2D pose estimation using part affinity fields. IEEE Transactions on Pattern Analysis and Machine Intelligence, 43(1), 172-186. https://doi.org/10.1109/TPAMI.2019.2929257

[11] Sun, K., Xiao, B., Liu, D., & Wang, J. (2019). Deep high-resolution representation learning for human pose estimation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 5693-5703). IEEE. https://doi.org/10.1109/CVPR.2019.00584

[12] Kidzinski, L., Yang, B., Hicks, J. L., Rajagopal, A., Delp, S. L., & Schwartz, M. H. (2020). Deep neural networks enable quantitative movement analysis using single-camera videos. Nature Communications, 11, 4054. https://doi.org/10.1038/s41467-020-17807-z

[13] Uhlrich, S. D., Falisse, A., Kidzinski, L., Muccini, J., Ko, M., Chaudhari, A. S., Hicks, J. L., & Delp, S. L. (2023). OpenCap: Human movement dynamics from smartphone videos. PLoS Computational Biology, 19(10), e1011462. https://doi.org/10.1371/journal.pcbi.1011462

[14] Needham, L., Evans, M., Wade, L., Cosker, D. P., McGuigan, M. P., Bilzon, J. L., & Colyer, S. L. (2022). The development and evaluation of a fully automated markerless motion capture workflow. Journal of Biomechanics, 144, 111338. https://doi.org/10.1016/j.jbiomech.2022.111338

[15] Schmitz, A., Ye, M., Shapiro, R., Yang, R., & Noehren, B. (2014). Accuracy and repeatability of joint angles measured using a single camera markerless motion capture system. Journal of Biomechanics, 47(2), 587-591. https://doi.org/10.1016/j.jbiomech.2013.11.031

[16] Mathis, A., Mamidanna, P., Cury, K. M., Abe, T., Murthy, V. N., Mathis, M. W., & Bethge, M. (2018). DeepLabCut: Markerless pose estimation of user-defined body parts with deep learning. Nature Neuroscience, 21(9), 1281-1289. https://doi.org/10.1038/s41593-018-0209-y

[17] Nath, T., Mathis, A., Chen, A. C., Patel, A., Bethge, M., & Mathis, M. W. (2019). Using DeepLabCut for 3D markerless pose estimation across species and behaviors. Nature Protocols, 14, 2152-2176. https://doi.org/10.1038/s41596-019-0176-0

[18] Pavllo, D., Feichtenhofer, C., Grangier, D., & Auli, M. (2019). 3D human pose estimation in video with temporal convolutions and semi-supervised training. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 7753-7762). IEEE. https://doi.org/10.1109/CVPR.2019.00794

[19] Karniadakis, G. E., Kevrekidis, I. G., Lu, L., Perdikaris, P., Wang, S., & Yang, L. (2021). Physics-informed machine learning. Nature Reviews Physics, 3, 422-440. https://doi.org/10.1038/s42254-021-00314-5

[20] Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational Physics, 378, 686-707. https://doi.org/10.1016/j.jcp.2018.10.045