

Markerless Pose Estimation for Resistance Training Technique Assessment

Joseph Turner, Jeff Clark^[0000–0003–0118–3999], and Nawid
Keshntmand^[0009–0008–5552–1395]

School of Engineering Mathematics and Technology, University of Bristol, UK
`nawid.keshntmand@bristol.ac.uk`

Abstract. Resistance training can be a high risk activity, and safe form is essential to avoiding injury. Laboratory-based movement analysis provides quantitative technique assessment, yet is not easily accessible. Markerless pose estimation infers body landmarks from images or video without physical markers and could offer a feasible alternative for technique assessment. We present a pose estimation framework to evaluate resistance-training technique from ordinary video footage. Using BlazePose [4], anatomical landmarks were extracted from squat, bench press, and deadlift videos and converted into joint-angle trajectories, with the squat serving as the primary case study. Trajectories were assessed against a defined reference repetition using root mean square error (RMSE). Results show that the framework recovers meaningful kinematic patterns for the squat and deadlift, enabling quantitative comparison between repetitions and identification of technique variability within a set. Performance depended strongly on camera orientation and visual occlusion, with non-sagittal views distorting 2D joint-angle estimates. The findings demonstrate that markerless pose estimation can support accessible biomechanical assessment outside laboratory environments.

Keywords: Pose estimation · Safety · Biomechanics

1 Introduction

Efficient and safe movement is central to sports performance, rehabilitation, and injury prevention [2]. In resistance training, technique is typically assessed through subjective coaching or, less commonly, laboratory-based biomechanical analysis [8]. Marker-based motion capture systems such as Vicon provide highly accurate kinematic measurements but require expensive specialised equipment and controlled environments, limiting their use in everyday training settings [3,15].

Recent advances in computer vision have made markerless pose estimation a promising alternative. Models such as BlazePose infer anatomical landmarks directly from ordinary video, enabling extraction of kinematic information from smartphone recordings at low cost [4,14]. This creates opportunities for accessible movement analysis in training and rehabilitation [12].

This paper aims to determine whether markerless pose estimation can provide a practical and interpretable basis for biomechanical assessment of resistance-training exercises outside laboratory environments. To investigate this, we developed and evaluated a framework for analysing resistance-training technique from standard video footage using pose estimation. Anatomical landmarks were extracted from recordings of the squat, bench press, and deadlift and converted into joint-angle trajectories. The squat served as the primary case study, with knee and trunk angles used for detailed biomechanical analysis, and quantitatively compared against gold standard reference repetitions. This work is intended as a feasibility study demonstrating the potential of markerless pose estimation for accessible resistance-training assessment rather than a fully validated biomechanical assessment system.

The main contributions of this paper are:

- Development and evaluation of a pose-estimation framework for analysing resistance-training technique from standard video footage
- Curation of a resistance-training dataset, including reference repetitions demonstrating good technique
- Investigation of the effects of camera viewpoint on joint-angle estimation

2 Related Work

Markerless approaches estimate anatomical landmarks directly from video, offering an accessible solution for coaching, rehabilitation, and performance monitoring [5]. Their accuracy, however, remains sensitive to camera viewpoint, occlusion, and image quality [16].

Human pose estimation has progressed from part-based models [18] to deep learning approaches such as HRNet [13], VideoPose3D [9], and transformer-based models such as ViTPose [17]. These methods have significantly improved landmark detection and, in some cases, enabled monocular 3D reconstruction. For practical applications, lightweight models such as BlazePose [4] are particularly attractive because they provide real-time full-body tracking from standard video.

For sports applications, OpenPose has been applied to elite long-jump competition footage, demonstrating that meaningful biomechanical measures can be extracted outside laboratory conditions [6]. More recently, AthletePose3D introduced a large-scale dataset of athletic movements to improve model performance on high-speed sports actions [19].

In resistance training, computer vision has been used to estimate external variables such as barbell velocity [1]. However, assessing movement quality requires joint-level biomechanical analysis. Squat biomechanics studies have shown that knee and trunk motion are key determinants of loading and technique [7]. Mercadal-Baudart et al. demonstrated that joint angles derived from single-camera pose estimation can provide interpretable metrics for strength exercises [8]. Despite these advances, limited work has evaluated how robustly markerless pose estimation can assess complex powerlifting movements under realistic gym

conditions. In particular, the effects of camera viewpoint, occlusion, and intra-set variability remain under explored, which this paper seeks to address.

3 Method

3.1 Dataset Curation and Preprocessing

Two publicly available Kaggle datasets [11,10] were combined ($n = 2618$ videos). The three target powerlifts (squat, bench press, and deadlift) were selected to focus on compound movements involving multiple joints and a higher risk of injury. As the original datasets contained a wide variety of gym exercises, the majority of videos were excluded during this initial selection step. The remaining recordings were then filtered to retain only videos suitable for reliable biomechanical analysis. Clips with poor lighting, severe occlusion, incomplete body visibility, or extreme camera viewpoints were removed, as these conditions prevent consistent landmark detection and can introduce substantial errors into 2D joint-angle estimation. Following this filtering process, the final dataset comprised 89 videos, which were subsequently normalised to a consistent frame rate and resolution before analysis.

To enable quantitative comparison between repetitions, a reference trajectory was established from a curated second set of instructional videos sourced from publicly available material, selected for technical soundness against widely accepted coaching standards. Each repetition was then evaluated against four biomechanical criteria. For the squat these were: (i) sufficient depth, indicated by a low minimum knee angle; (ii) a smooth, continuous U-shaped descent-ascent profile; (iii) stable trunk angle throughout; and (iv) minimal discontinuities attributable to pose estimation noise. The repetition best satisfying all four criteria was retained as the reference trajectory for subsequent comparisons. Reference repetitions could be tailored to an individual, providing quantified feedback tailored to the athlete’s specific anatomy and training style.

3.2 Pose-Based Biomechanical Analysis Framework

BlazePose is a lightweight convolutional neural network designed for real-time 2D human pose estimation [4]. This was implemented as the foundation of the analysis pipeline. For each frame t , the model returns a set of $M = 33$ anatomical landmarks

$$\mathbf{p}_i^{(t)} = (x_i^{(t)}, y_i^{(t)}), \quad i = 1, \dots, M, \quad (1)$$

where $(x_i^{(t)}, y_i^{(t)})$ are the 2D image-plane coordinates of the i -th keypoint. Joint angles were derived from landmark triplets using a standard vector formulation: for three points A, B, C where B is the joint of interest,

$$\theta = \cos^{-1} \left(\frac{(A - B) \cdot (C - B)}{\|A - B\| \|C - B\|} \right). \quad (2)$$

For the squat, the knee angle was computed from the hip-knee-ankle triplet, and trunk posture was quantified as the angular deviation of the shoulder-hip segment from the vertical axis. For the deadlift, trunk angle was derived using the same formulation, with a hip-hinge angle defined from the shoulder-hip-knee triplet to capture the degree of forward flexion at the hip. For the bench press, upper-limb kinematics were quantified using the shoulder-elbow-wrist triplet to compute elbow flexion angle, defined as the internal angle at the elbow joint.

3.3 Temporal Normalisation

Since repetitions differ in duration due to differences in pacing and frame count, direct frame-by-frame comparison is not meaningful without temporal alignment. Linear interpolation was applied to resample each joint-angle signal onto a common normalised time axis. For a repetition of T frames, the normalised coordinate was defined as:

$$\hat{\tau}_t = \frac{t}{T-1}, \quad t = 0, 1, \dots, T-1 \quad (3)$$

Each joint-angle signal was then linearly interpolated and resampled onto a common normalised time axis

$$\tau_i = \frac{i}{100}, \quad i = 0, 1, \dots, 100, \quad (4)$$

yielding 101 samples per repetition regardless of its original frame count. This process preserves the shape of each movement trajectory whilst removing variability due to execution speed. The reference was normalised using the same procedure, ensuring all comparisons were made on a common movement axis.

For multi-repetition analysis, a phase-aware variant was also applied. Since a multi-repetition recording contains several consecutive movement cycles, simple linear normalisation cannot distinguish between individual repetitions or align their phases meaningfully. Each repetition was therefore first segmented, then divided into descent, hold, and ascent phases, assigned fixed proportions of 40%, 20%, and 40% of the full movement cycle respectively, with each phase independently normalised to $[0, 1]$ prior to concatenation. The 40–20–40 proportions were selected as a simple heuristic to preserve the approximate temporal structure of a squat repetition, where descent and ascent typically occupy most of the movement and any pause at the bottom is comparatively brief. More adaptive phase alignment approaches, such as dynamic time warping, represent an important direction for future work.

3.4 Joint Angle Trajectory Comparison

Following reference repetition trajectory selection, deviation between each test repetition and the reference trajectory was quantified using RMSE, which was

computed independently for the knee-angle and trunk-angle trajectories. To express RMSE on a more interpretable scale, each value was converted to a similarity score ranging from 0 to 100:

$$S = \max(0, 100 - 2 \text{ RMSE}) . \quad (5)$$

The scaling factor of two was selected based on the typical RMSE values observed in the dataset, allowing the resulting scores to span an interpretable 0–100 range. The similarity score is not intended as a universally validated measure of exercise technique quality, but rather as a task-specific metric for quantifying consistency between a test repetition and a reference trajectory. This provides a simple and interpretable measure of movement agreement suitable for automated assessment using markerless pose estimation. An overall similarity score is reported as the mean of the knee and trunk scores.

3.5 Intra-Set Variability Analysis

Repetitions were segmented by detecting local minima of the knee-angle trajectory, each corresponding to the bottom position of a squat. Only complete descent-ascent cycles were retained for analysis. Each segmented repetition was independently time-normalised and per-repetition metrics were computed. Set-level summary statistics (mean RMSE, standard deviation, and best and worst repetitions by RMSE) were used to quantify intra-set technical consistency. The pipeline’s code will be released upon publication.

4 Results and Discussion

4.1 Pose Estimation Performance Across Exercises

The pipeline was evaluated on the dataset of squat, deadlift, and bench press recordings. Squats and deadlifts performed reliably, achieving mean usable-frame rates of 99.7% and 99.0% respectively. Bench press was substantially less reliable, with only 73.6% usable frames and 77.2% of videos successfully processed, reflecting the impact of barbell occlusion and supine body orientation on landmark extraction. These results indicate that markerless tracking performs best for upright exercises where the limbs of interest remain visible throughout the movement cycle. Consequently, the remaining analysis focuses primarily on squat and deadlift movements.

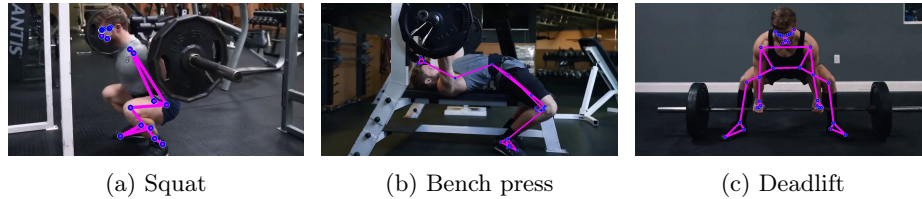


Fig. 1: Representative examples of markerless pose estimation across the three lifts. (a) Squat: the upright side-on view provides clear lower-limb visibility. (b) Bench press: upper-limb landmarks (i.e. shoulder, elbow, wrist) are difficult to track due to supine position and barbell occlusion during the repetition. (c) Deadlift: the upright setup supports consistent tracking of main joint landmarks.

4.2 Squat Trajectories and Reference Trajectory Comparison

Joint-angle trajectories extracted from squat videos demonstrated the expected biomechanical structure. When applied to multi-repetition analysis for a ten-repetition squat set, after temporal normalisation, most repetitions exhibited the characteristic U-shaped knee profile associated with squat descent and ascent, with minimum knee angles ranging approximately between 25° and 50° (Figure 2).

Most repetitions followed similar movement profiles and remained close to the reference repetition trajectory (dashed blue line, Figure 2). Quantitative results are reported in Table 1, with a mean RMSE of 17.56° and mean similarity score of 64.9/100. Rep 1 showed the largest deviation from the reference repetition trajectory (RMSE 37.50°), whereas Rep 8 produced the closest agreement (RMSE 13.18°). These findings demonstrate that the framework can identify technique variability both within individual repetitions and across an entire set.

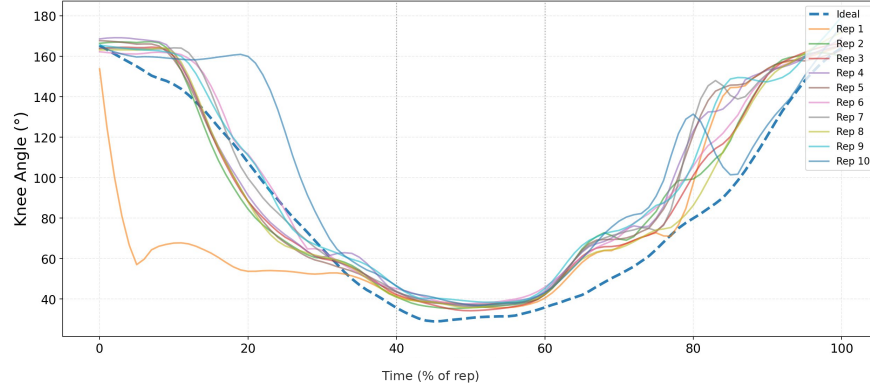


Fig. 2: Time-normalised knee trajectories for all ten repetitions within a squat set compared with the reference repetition trajectory (dashed line).

Rep	Knee Min (°)	Knee Max (°)	Knee RMSE (°)	Knee Similarity
1	35.5	173.9	37.50	25.0/100
2	35.4	166.9	15.62	68.8/100
3	34.1	168.3	14.15	71.7/100
4	36.3	169.9	19.24	61.5/100
5	37.3	168.1	19.80	60.4/100
6	37.6	165.2	15.61	68.8/100
7	36.4	165.1	19.45	61.1/100
8	34.8	172.3	13.18	73.6/100
9	38.0	176.7	18.74	62.5/100
10	36.9	179.3	22.33	55.3/100
Mean			17.56	64.9/100
SD			6.56	

Table 1: Per-repetition knee-angle metrics for the analysed squat set using phase-aware normalisation. RMSE is computed relative to the reference repetition trajectory.

4.3 Camera Sensitivity and Occlusion Effects

One of the strongest findings of this work was the sensitivity of joint-angle estimation to viewpoint, seen in the camera-rotation experiment (Figure 3).

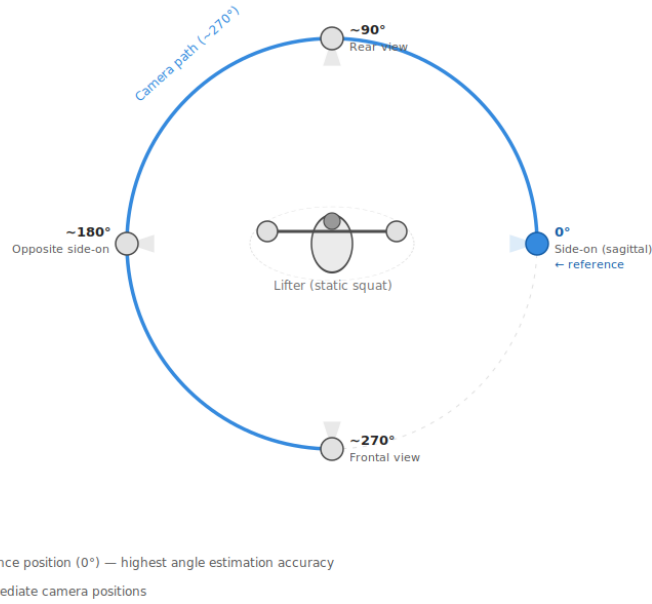


Fig. 3: Bird’s eye view of the camera rotation experiment. The lifter maintained a static squat posture at the centre whilst the camera was moved approximately 270° around the subject. The blue circle marks the sagittal side-on position (0°), which served as the reference angle due to its superior joint angle estimation accuracy. Key intermediate viewpoints rear (~90°), opposite side-on (~180°), and frontal (~270°) are also marked to illustrate the range of camera orientations evaluated.

It showed that estimated knee angle changed substantially even when the participant maintained a fixed squat posture. The effect became most pronounced inside a power rack where lower-limb landmarks became partially obscured (Figure 4). The largest deviations occur when the rack partially blocks the lower limbs. Consequently, apparent changes in knee angle may reflect projection distortion rather than genuine biomechanical differences, a limitation consistent with the occlusion and viewpoint sensitivity reported by Wade et al [16]. This

explains several examples in the dataset where visually acceptable squats failed to produce the expected trajectory shape. The results therefore identify sagittal recording as the preferred viewpoint for practical deployment.

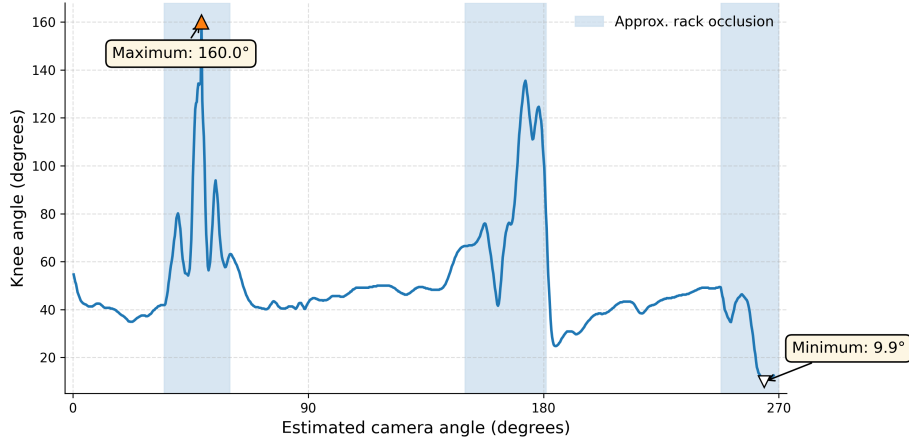


Fig. 4: Estimated knee angle during camera rotation around a static squat posture inside a power rack. Occlusion regions (shaded blue) produce larger deviations from the side-on reference estimate. See Figure 3 for details on the experimental setup and corresponding camera angles.

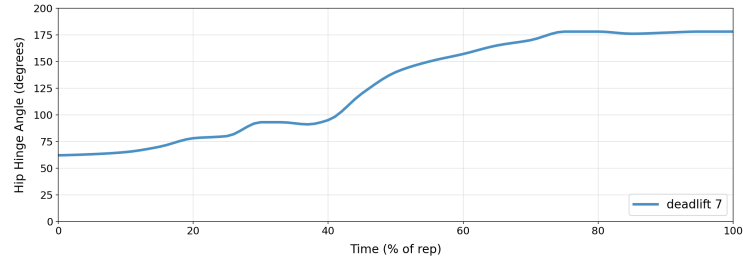
4.4 Joint-Angle Trajectories for a Representative Deadlift Trial

To demonstrate that the proposed framework generalises beyond the squat, a representative deadlift recording was analysed from a sagittal viewpoint. The same temporal normalisation procedure described for the squat was applied prior to trajectory comparison.

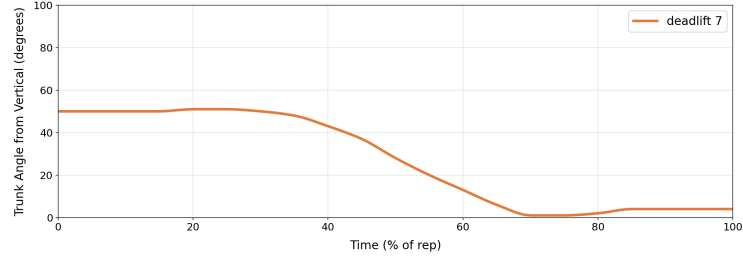
The hip-hinge angle increased smoothly throughout the movement, reflecting progressive hip extension from the bottom position to full lockout. The trunk angle began in a forward-leaning position and became more upright during ascent, indicating controlled trunk extension throughout the lift. Both signals exhibited a monotonic structure compared with the squat, reflecting the hinge-dominant nature of the movement. These results demonstrate that the framework can extract meaningful joint-angle trajectories beyond the squat, supporting its generalisability to compound movements with distinct biomechanical profiles.

4.5 Limitations and Future Work

The primary limitation of this framework is the sensitivity of 2D joint-angle estimation to camera orientation. As demonstrated in Section 4.3, even when the



(a) Time-normalised hip hinge angle.



(b) Time-normalised trunk angle from vertical.

Fig. 5: Joint-angle trajectories for a representative deadlift repetition.

underlying posture remains unchanged, viewpoint changes introduce substantial distortion in estimated joint angles. Preliminary work into 3D landmark reconstruction suggests partial improvement under non-sagittal conditions but did not fully recover the expected range of motion, as the landmarks are inferred from a single view and remain subject to perspective ambiguity. Reliable deployment in uncontrolled environments will therefore require either stricter guidance on camera placement or more robust depth-estimation techniques.

The RMSE metric, whilst interpretable, weights all phases of the movement equally. Deviations at biomechanically critical points, such as the bottom position of the squat, may carry greater practical significance than deviations during transitional phases, and a phase-weighted metric could better reflect how technique is evaluated in practice.

The dataset used is relatively small and specific. Real-world deployment would involve a wider range of lighting conditions, camera placements, and lifter experience levels, which the current pipeline does not fully account for. Expanding the dataset to include a wider range of lifters, experience levels, and recording environments would improve the external validity of the framework and support more robust evaluation of the similarity metric across populations.

5 Conclusion

This paper presents and evaluates a markerless pose estimation framework for assessing resistance-training technique from standard video footage, and successfully reconstructed joint-angle trajectories for the squat and deadlift, with mean usable frame rates of 99.7% and 99.0% respectively. A frame was considered usable when BlazePose successfully detected all landmarks required for the joint-angle calculation with valid coordinates. The usable-frame rate therefore represents the percentage of frames within a video that could be included in the biomechanical analysis. Squat trajectories demonstrated consistent biomechanical structure and strong agreement with a reference repetition under sagittal viewing conditions. Intra-set analysis captured movement quality variations across a set, with early and late repetitions showing the greatest deviation from the reference repetition, consistent with movement initiation effects and fatigue.

Camera viewpoint emerged as the strongest practical constraint. Rotation experiments confirmed that 2D joint-angle estimates are highly sensitive to perspective distortion, and that non-sagittal recordings can produce misleading trajectories even when the underlying movement is technically sound. The extension of the pipeline to the deadlift produced meaningful hip-hinge and trunk trajectories consistent with expected biomechanics, supporting the generalisability of the approach to compound movements with distinct mechanical profiles.

Overall, the results indicate that markerless pose estimation can provide accessible, interpretable biomechanical assessment outside laboratory environments. Whilst limitations around viewpoint sensitivity, occlusion, and metric design remain, the framework represents a practical step towards data-driven movement feedback for everyday training settings.

6 Code Availability

The code used to implement the proposed framework will be released upon publication at github.com/ac2771/Markerless-Pose-Resistance-Training.

Acknowledgments. NK was supported by the EPSRC LEAP Digital Health Hub grant EP/X031349/1.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Ågren, O., Palm, J.: Enhancing athletic training through AI: A comparative analysis of YOLO versions for image segmentation in velocity-based training. Student thesis, DiVA portal (2024), <https://www.diva-portal.org/smash/get/diva2:1872383/FULLTEXT01.pdf>

2. Bahr, R., Krosshaug, T.: Understanding injury mechanisms: a key component of preventing injuries in sport. *British journal of sports medicine* **39**(6), 324–329 (2005)
3. Barris, S., Button, C.: A review of vision-based motion analysis in sport. *Sports Medicine* **38**(12), 1025–1043 (2008). <https://doi.org/10.2165/00007256-200838120-00006>, <https://link.springer.com/article/10.2165/00007256-200838120-00006>
4. Bazarevsky, V., Kartynnik, Y., Vakunov, A., Raveendran, K., Grundmann, M.: BlazePose: On-device real-time body pose tracking. <https://arxiv.org/abs/2006.10204> (2020), accessed: 2025-11-06
5. Colyer, S.L., Evans, M., Cosker, D.P., Salo, A.I.T.: A review of the evolution of vision-based motion analysis and the integration of advanced computer vision methods towards developing a markerless system. *Sports Medicine – Open* **4**(1), 24 (2018). <https://doi.org/10.1186/s40798-018-0139-y>, <https://link.springer.com/article/10.1186/s40798-018-0139-y>
6. Cronin, N.J., Walker, J., Tucker, C.B., Nicholson, G., Cooke, M., Merlino, S., Bissas, A.: Feasibility of OpenPose markerless motion analysis in a real athletics competition. *Frontiers in Sports and Active Living* **5**, 1298003 (2024). <https://doi.org/10.3389/fspor.2023.1298003>, <https://pmc.ncbi.nlm.nih.gov/articles/PMC10796501/>
7. Escamilla, R.F.: Knee biomechanics of the dynamic squat exercise. *Medicine & Science in Sports & Exercise* **33**(1), 127–141 (2001). <https://doi.org/10.1097/00005768-200101000-00020>, <https://pubmed.ncbi.nlm.nih.gov/11194098/>
8. Mercadal-Baudart, C., Liu, C.J., Farrell, G., Boyne, M., González Escribano, J., Smolic, A., Simms, C.: Exercise quantification from single camera view markerless 3D pose estimation. *Heliyon* **10**(6), e27596 (2024). <https://doi.org/10.1016/j.heliyon.2024.e27596>, <https://www.sciencedirect.com/science/article/pii/S2405844024036272>
9. Pavlo, D., Feichtenhofer, C., Grangier, D., Auli, M.: 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*. pp. 7753–7762 (2019)
10. Philosopher0808: Gym workout/exercises video dataset. <https://www.kaggle.com/datasets/philosopher0808/gym-workoutexercises-video/data> (2023), accessed: 2025-11-06
11. Sinabutar, D.: Deadlift pose estimation dataset. <https://universe.roboflow.com/daniel-sinabutar/deadlift-pose-estimation> (2023), accessed: 2025-11-06
12. Souaifi, M., Dhahbi, W., Jebabli, N., Ceylan, H.İ., Boujabli, M., Muntean, R.I., Dergaa, I.: Artificial intelligence in sports biomechanics: A scoping review on wearable technology, motion analysis, and injury prevention. *Bioengineering* **12**(8), 887 (2025)
13. Sun, K., Xiao, B., Liu, D., Wang, J.: 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 (2019). <https://doi.org/10.1109/CVPR.2019.00585>
14. Ultralytics: What is pose estimation and where can it be used? <https://www.ultralytics.com/blog/what-is-pose-estimation-and-where-can-it-be-used> (2024), accessed: 2025-11-06
15. Vicon Motion Systems Ltd.: Vicon: Motion capture systems for biomechanics and sports (2026), <https://www.vicon.com/>, accessed: 1 April 2026

16. Wade, L., Needham, L., McGuigan, P., Bilzon, J.: Applications and limitations of current markerless motion capture methods for clinical gait biomechanics. *PeerJ* **10**, e12995 (2022). <https://doi.org/10.7717/peerj.12995>, <https://pmc.ncbi.nlm.nih.gov/articles/PMC8884063/>, published 25 February 2022
17. Xu, Y., Zhang, J., Zhang, Q., Tao, D.: ViTPose: Simple vision transformer baselines for human pose estimation. arXiv preprint arXiv:2204.12484 (2022). <https://doi.org/10.48550/arXiv.2204.12484>, <https://arxiv.org/abs/2204.12484>
18. Yang, Y., Ramanan, D.: Articulated pose estimation with flexible mixtures-of-parts. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. pp. 1385–1392 (2011). <https://doi.org/10.1109/CVPR.2011.5995741>, https://yangyi02.github.io/research/pose/pose_cvpr2011.pdf
19. Yeung, C., Suzuki, T., Tanaka, R., Yin, Z., Fujii, K.: AthletePose3D: A benchmark dataset for 3D human pose estimation and kinematic validation in athletic movements. arXiv preprint arXiv:2503.07499 (2025), <https://arxiv.org/abs/2503.07499>