Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

A Posture-Wise Comparative Analysis of Human Pose Estimation Models for Physical Rehabilitation Monitoring

Deepak Shukla[1], Maya Rathore[2]
[1]Department of Computer Science & Engineering, Faculty of Engineering, Oriental University, Indore, India. deepakactive@gmail.com
[2]Department of Computer Science & Engineering, Faculty of Engineering, Oriental University, Indore, India. mayarathore114@gmail.com

ARTICLE INFO

Received: 02 Nov 2024
Revised 17 Dec 2024
Accepted: 28 Dec 2024

ABSTRACT

Introduction: Marker-less human pose estimation (HPE) is increasingly proposed for monitoring physical rehabilitation, yet most models are trained on upright, everyday poses, whereas physiotherapy involves lying, sitting and slow, occluded movements.

Objectives: To determine, posture by posture, how widely used 2D and 3D HPE models behave on rehabilitation exercises, how stable their accuracy is across postures, and which requirements a rehabilitation-specific estimator must meet.

Methods: A secondary quantitative analysis of the published UCO Physical Rehabilitation benchmark (27 subjects, OptiTrack ground truth) covering seven models: AlphaPose, MediaPipe, KAPAO, StridedTransformer, HybrIK, PoseBERT and VideoPose3D. Beyond the reported 2D keypoint error, 3D joint-angle mean absolute error (MAE) and throughput, we derive a Posture Sensitivity Index, rank stability, rotation gain for supine frames, an accuracy–throughput Pareto front and the gap to a ~5° clinical reference.

Results: No model leads in every posture. For 2D keypoints AlphaPose is best on supine frames (3.86×10^{-2} couch-normalized units) and KAPAO on seated (3.24) and standing (3.34) frames; supine error is 14–103% higher than the best upright error for every model. For 3D joint angles HybrIK is best for supine (6.92°) and seated (7.29°) exercises and MediaPipe for standing (11.96°). The best 3D error exceeds the ~5° reference by 1.9°, 2.3° and 7.0° in the three postures, and the most accurate models are the slowest (13–16 FPS).

Conclusions: General-purpose HPE models are not yet sufficient for posture-agnostic rehabilitation monitoring. Estimators that reason jointly over all body joints, remain robust to non-upright orientation and run in real time are needed; joint-token transformer designs are a promising direction.

Keywords: human pose estimation; physical rehabilitation; joint-angle estimation; posture robustness; comparative analysis; markerless motion capture

INTRODUCTION

Physical rehabilitation restores motor function after injury or surgery, and clinicians judge progress mainly through posture and joint range of motion. Much of this therapy is performed at home without supervision, so objective, low-cost monitoring is valuable [1, 2]. Marker-less human pose estimation (HPE) from ordinary RGB cameras offers such monitoring and has been proposed for tele-rehabilitation, gait analysis and motor assessment [3, 4]. Depth cameras such as the Kinect were an earlier option but have known limits outside the laboratory [5].

Modern HPE models, from convolutional estimators such as OpenPose [6] and HRNet [7] to recent transformer-based designs, are, however, trained and benchmarked on general-purpose data such as COCO [8] and Human3.6M [9], in which people are mostly upright. Rehabilitation differs: many exercises are performed lying on a couch (supine) or seated, limbs occlude one another, and the clinically relevant output is a joint angle rather than a keypoint. Studies comparing OpenPose-type estimators on upper-limb exercises [10], 3D joint-centre localisation [11] and gait kinematics [12] all report that accuracy depends strongly on the model and the activity. The UCO Physical Rehabilitation study of Aguilar-Ortega et al. [1] provided the first multi-model benchmark on rehabilitation exercises with optical ground truth. It showed that posture and camera viewpoint matter, and it reported per-posture results for eight estimators. What that study did not provide is an integrated reading of the results in terms of how stable

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

each model is across postures, how accuracy trades against speed, and how far the best models remain from clinically useful accuracy. This paper addresses that gap through a structured secondary analysis of the published benchmark, and uses the analysis to derive concrete requirements for rehabilitation-specific HPE.

OBJECTIVES

The study pursues three objectives:

O1. To compare seven representative 2D and 3D HPE models posture by posture (supine, seated, standing) on 2D keypoint error, 3D joint-angle error and throughput.

O2. To quantify robustness through derived indices: a Posture Sensitivity Index, rank stability across postures, the effect of rotating supine frames, and the accuracy–throughput Pareto front.

O3. To relate the best achievable errors to a clinical reference and derive design requirements for a rehabilitation-specific estimator.

METHODS

Study design and data source

This is a secondary quantitative analysis; no new recordings or model runs were made. All primary values are taken, with attribution, from the benchmark of Aguilar-Ortega et al. [1] on the UCO Physical Rehabilitation dataset. That dataset contains 2160 videos of 27 healthy adults performing eight exercises (lower and upper limb, left and right sides) prescribed by surgeons and physiotherapists, recorded by five RGB cameras at 1280 × 720 pixels, with joint ground truth from an OptiTrack infrared system. Exercises are grouped by body position into supine (lying leg exercises), seated and standing groups. Fig. 1 summarizes the analysis pipeline.

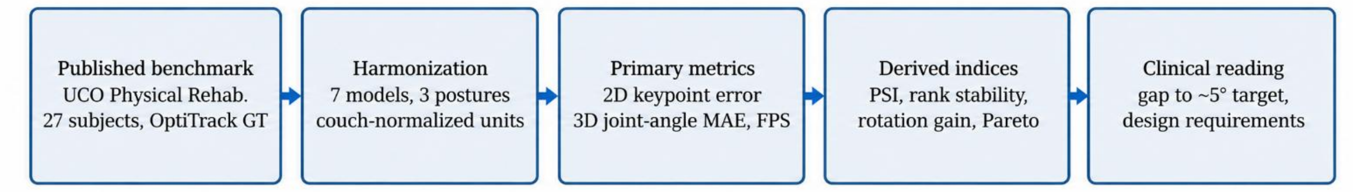

Figure 1. Analysis framework used in this study

Models under study

Seven models were retained to cover the main design families (Table 1). HMR [13], also evaluated in [1], was excluded because it had the highest error in every posture and has been superseded by HybrIK, which also recovers an SMPL body [14]. All models were used in [1] with their public default configurations and without fine-tuning.

Table 1. Models included in the analysis

Model	Paradigm	Output used	FPS in [1]
AlphaPose [15, 16]	Top-down CNN, regional multi-person	2D keypoints	15.93
MediaPipe / BlazePose [17, 18]	Lightweight detector–tracker	2D + 3D	66.58
KAPAO [19]	Single-stage keypoints-as-objects	2D keypoints	41.43
StridedTransformer [20]	Spatio-temporal transformer lifting	2D + 3D	21.62
VideoPose3D [21]	Dilated temporal convolution lifting	2D + 3D	19.09
HybrIK [22]	Analytical–neural inverse kinematics	3D	13.21
PoseBERT [23]	Transformer temporal refinement	3D	24.30

Primary metrics

Following [1], the 2D error of a frame is the mean Euclidean distance over the three joints of the exercised limb (hip–knee–ankle or shoulder–elbow–wrist), normalized by the projected length of the treatment couch so that errors are independent of subject distance and resolution:

$$e_{2D} = (1/3)\, \Sigma_j \, \lVert \hat{k}_j - k_j \rVert \, / \, L_{couch} \quad (1)$$

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

The joint angle at the middle joint b of the triplet (a, b, c) and its mean absolute error over N frames are

$$\theta = \arccos[(\mathbf{v}_1 \cdot \mathbf{v}_2) / (\|\mathbf{v}_1\| \, \|\mathbf{v}_2\|)], \;\; \mathbf{v}_1 = k_a - k_b, \;\; \mathbf{v}_2 = k_c - k_b \quad (2)$$

$$\mathrm{MAE}_\theta = (1/N) \, \Sigma_i \, |\hat{\theta}_i - \theta_i| \quad (3)$$

2D errors are reported ×100 for readability. Supine values refer to the rotated-video protocol of [1], which gave lower error for most models.

Derived indices

For each model with per-posture error E_p, the **Posture Sensitivity Index** is the ratio of its worst to its best posture error; a value of 1 means perfectly posture-invariant behaviour:

$$\mathrm{PSI} = \max_p E_p \, / \, \min_p E_p \quad (4)$$

The **rotation gain** is the relative error reduction when supine videos are rotated by 90° so that the subject appears upright, computed from the original and rotated means in [1]:

$$G_{rot} = (E_{orig} - E_{rot}) \, / \, E_{orig} \times 100\% \quad (5)$$

Rank stability is assessed by ranking models within each posture by mean error. The **accuracy–throughput Pareto front** contains the models for which no other model is both faster and more accurate. Finally, the **clinical gap** is the difference between the best 3D angle MAE in each posture and a ~5° reference, taken from the inertial-sensor knee-flexion error cited in [1].

RESULTS

2D keypoint localization

Table 2 and Fig. 2 give the 2D keypoint error. AlphaPose is the most accurate on supine frames (3.86) and has the lowest mean over all postures (3.58) and the lowest PSI (1.14), making it the most posture-robust 2D model. On seated and standing frames KAPAO (3.24, 3.34) and StridedTransformer (3.26, 3.34) are practically tied for first place; the Friedman ranking in [1] places StridedTransformer first for seated exercises. Every model is worst on supine frames: supine error exceeds the model's best upright error by 14.2% (AlphaPose), 45.6% (VideoPose3D), 51.5% (KAPAO), 59.9% (MediaPipe) and 102.5% (StridedTransformer).

Table 2. 2D keypoint error (×100, couch-normalized; mean ± SD) reported in [1], with derived mean and PSI

Model	Supine	Seated	Standing	Mean	PSI
AlphaPose	3.86 ± 3.01	3.38 ± 1.40	3.49 ± 1.54	3.58	1.14
MediaPipe	5.58 ± 6.75	3.99 ± 1.81	3.49 ± 1.35	4.35	1.60
KAPAO	4.91 ± 6.47	3.24 ± 1.19	3.34 ± 1.03	3.83	1.52
StridedTransformer	6.60 ± 6.68	3.26 ± 1.63	3.34 ± 1.02	4.40	2.02
VideoPose3D (2D stage)	4.92 ± 4.93	3.38 ± 1.35	3.45 ± 1.07	3.92	1.46

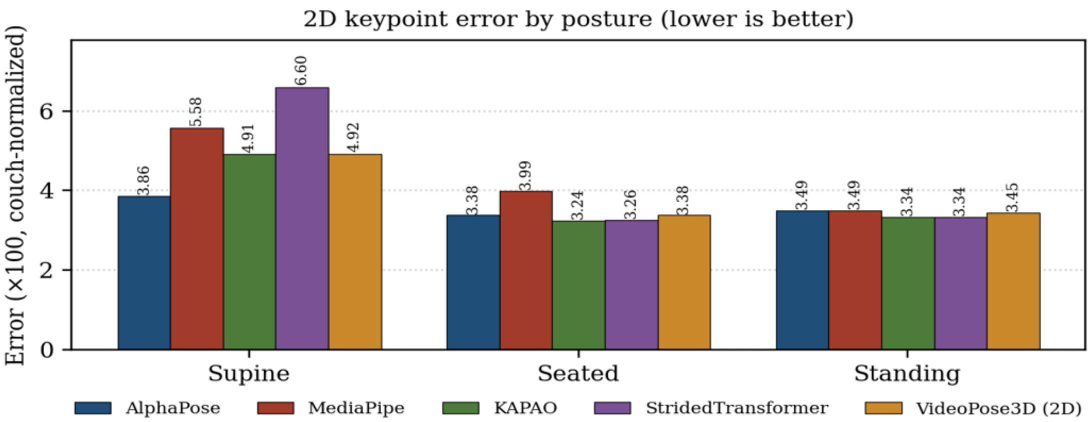

Figure 2. 2D keypoint error by posture (data from [1])

3D joint-angle estimation

Table 3 and Fig. 3 give the 3D joint-angle MAE. HybrIK is clearly best for supine (6.92°) and seated (7.29°) exercises, where its kinematic body prior constrains implausible limb configurations, but its error almost doubles for standing upper-limb exercises (13.70°; PSI 1.98). MediaPipe is best for standing (11.96°) and is the most posture-stable 3D model (PSI 1.11), although its absolute error is higher than HybrIK's on supine and seated exercises. The lifting-based models VideoPose3D and StridedTransformer have the highest angle errors (14.1°–25.1°), because 2D errors propagate into the lifted 3D pose.

Table 3. 3D joint-angle MAE (°; mean ± SD) reported in [1], with derived mean and PSI

Model	Supine	Seated	Standing	Mean	PSI
HybrIK	6.92 ± 10.05	7.29 ± 11.18	13.70 ± 12.35	9.30	1.98
PoseBERT	9.70 ± 11.41	12.52 ± 9.71	19.75 ± 16.00	13.99	2.04
MediaPipe (3D)	10.75 ± 16.06	11.33 ± 9.79	11.96 ± 9.99	11.35	1.11
VideoPose3D	18.84 ± 18.95	14.07 ± 14.80	21.62 ± 18.98	18.18	1.54
StridedTransformer	15.75 ± 17.75	18.78 ± 18.56	25.05 ± 19.66	19.86	1.59

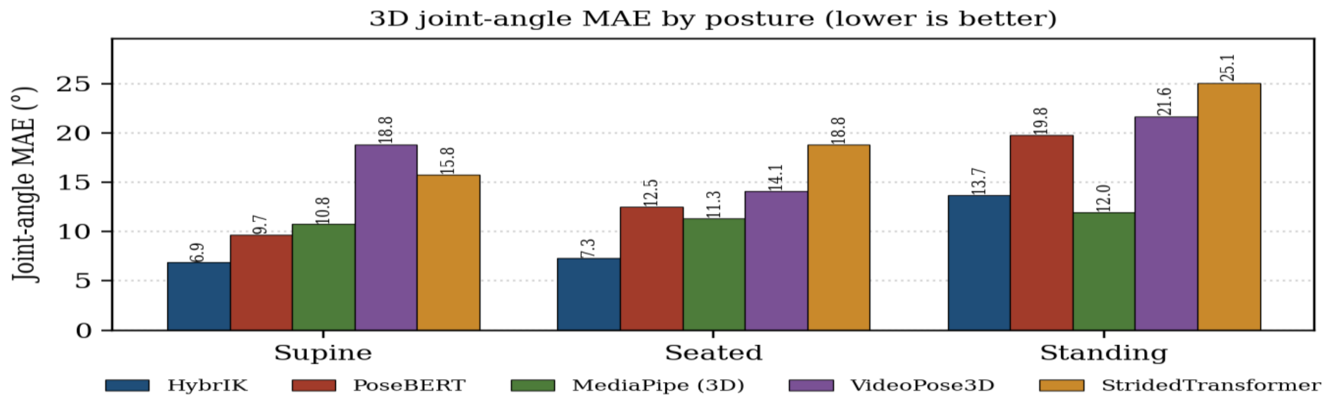

Figure 3. 3D joint-angle MAE by posture (data from [1])

Effect of rotating supine videos

Table 4 shows that rotating supine videos, so that the subject appears upright, reduces 2D error for four of five models (by up to 28.4% for VideoPose3D) and reduces HybrIK's 3D angle error by 61.7%. The effect is not universal: StridedTransformer's 2D error doubles and MediaPipe's 3D error rises slightly. The large gains confirm that the failures on supine frames are largely caused by orientation bias in the training data rather than by the exercises themselves.

Table 4. Supine exercises, original vs rotated videos (values from [1]; gain derived with Eq. 5)

Model	2D error orig. → rot.	G_rot (2D)	3D MAE orig. → rot. (°)	G_rot (3D)
AlphaPose	4.4 → 3.9	+11.4%	–	–
MediaPipe	5.1 → 5.0	+2.0%	10.30 → 10.76	−4.5%
KAPAO	6.7 → 5.0	+25.4%	–	–
StridedTransformer	3.4 → 6.8	−100.0%	19.43 → 15.42	+20.6%
VideoPose3D	7.4 → 5.3	+28.4%	17.81 → 18.17	−2.0%
HybrIK	–	–	17.93 → 6.86	+61.7%
PoseBERT	–	–	9.87 → 9.62	+2.5%

Accuracy–throughput trade-off

Fig. 4 plots the mean 2D error against the throughput reported in [1]. Only AlphaPose, KAPAO and MediaPipe lie on the Pareto front; VideoPose3D and StridedTransformer are dominated by KAPAO, which is both faster and more accurate. The most accurate 2D model (AlphaPose, 15.9 FPS) and the most accurate 3D model for lying and sitting

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

exercises (HybrIK, 13.2 FPS) are also the slowest, while the fastest model (MediaPipe, 66.6 FPS) is the least accurate in 2D.

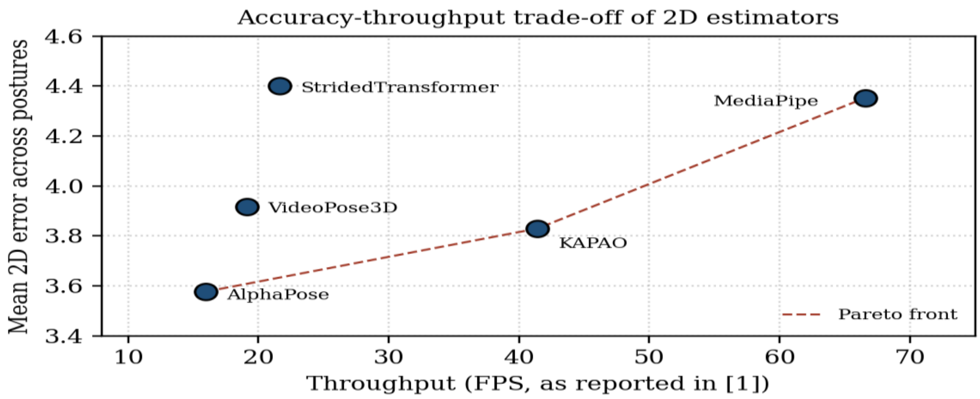

Figure 4. Accuracy–throughput trade-off of the 2D estimators (derived from [1])

DISCUSSION

The analysis leads to four findings. **First, no single model is best across postures.** The 2D leader changes from AlphaPose (supine) to KAPAO/StridedTransformer (seated, standing), and the 3D leader from HybrIK (supine, seated) to MediaPipe (standing). A clinic choosing a model from upright benchmarks such as COCO [8] could therefore select one that performs poorly on the lying exercises that make up much of lower-limb rehabilitation.

Second, posture sensitivity is large and systematic. Supine error exceeds upright error for every 2D model, and rotation experiments show that much of this penalty disappears when the image is turned upright. This points to orientation bias in general-purpose training data [8, 9] and is consistent with earlier reports that estimator accuracy depends on activity and viewpoint [11, 12]. Models that impose structure on the whole body, either through a kinematic prior (HybrIK) or through top-down person normalization (AlphaPose), are the most robust.

Third, the best models remain above clinical accuracy. Relative to a ~5° reference, the best 3D angle errors leave gaps of 1.9° (supine), 2.3° (seated) and 7.0° (standing), and the frame-level standard deviations reported in [1] are larger than the means, indicating frequent large failures under occlusion.

Fourth, accuracy and deployability conflict. The most accurate models run at 13–16 FPS, below comfortable real-time feedback, while the fastest model is the least accurate. Home rehabilitation needs both.

These findings translate into design requirements for a rehabilitation-specific estimator: (i) global reasoning over all joints so that occluded joints are inferred from visible ones; (ii) robustness to non-upright orientation, through training data or architecture rather than manual rotation; (iii) direct support for joint-angle output with a light temporal lifting stage; (iv) configurable joint sets for exercise-specific assessment; and (v) real-time throughput on mid-range hardware. Transformer designs that represent each joint as a learnable token attending to image features and to other joints, such as TokenPose [24] and ViTPose [25], combine global attention with moderate model size and are therefore a promising basis. Rehabilitation datasets such as UI-PRMD [26], KIMORE [27], IRDS [28] and UCO [1] make fine-tuning and exercise-quality assessment feasible [29], and practical systems such as ExerCam [30] show the value of camera-only deployment.

Limitations. The study relies on values published in a single benchmark [1] of healthy adults in a controlled room, so the findings may differ for patients, other exercises or home environments. Throughput figures depend on the hardware used in that study. The ~5° clinical reference is indicative rather than a formal acceptance criterion, and the 2D and 3D metrics are not directly comparable with each other. Broader surveys of HPE methods are available in [31, 4].

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

CONCLUSION

This posture-wise analysis of seven human pose estimation models on a rehabilitation benchmark shows that current general-purpose models are not yet sufficient for posture-agnostic rehabilitation monitoring. The most posture-robust 2D model is AlphaPose (PSI 1.14), the best 3D model for lying and sitting exercises is HybrIK (6.92° and 7.29°), and MediaPipe is the most stable and fastest 3D option. All models degrade on supine frames, the best 3D errors remain 1.9°–7.0° above a ~5° reference, and accuracy comes at the cost of speed. Future work will develop a joint-token transformer estimator with temporal lifting that addresses these requirements and will evaluate it with the protocol and indices introduced here.

REFERENCES

[1] Aguilar-Ortega, R., Berral-Soler, R., Jiménez-Velasco, I., Romero-Ramírez, F. J., García-Marín, M., Zafra-Palma, J., Muñoz-Salinas, R., Medina-Carnicer, R., & Marín-Jiménez, M. J. (2023). UCO Physical Rehabilitation: New dataset and study of human pose estimation methods on physical rehabilitation exercises. Sensors, 23(21), 8862. https://doi.org/10.3390/s23218862

[2] Debnath, B., O'Brien, M., Yamaguchi, M., & Behera, A. (2022). A review of computer vision-based approaches for physical rehabilitation and assessment. Multimedia Systems, 28(1), 209–239. https://doi.org/10.1007/s00530-021-00815-4

[3] Stenum, J., Cherry-Allen, K. M., Pyles, C. O., Reetzke, R. D., Vignos, M. F., & Roemmich, R. T. (2021). Applications of pose estimation in human health and performance across the lifespan. Sensors, 21(21), 7315. https://doi.org/10.3390/s21217315

[4] Avogaro, A., Cunico, F., Rosenhahn, B., & Setti, F. (2023). Markerless human pose estimation for biomedical applications: A survey. Frontiers in Computer Science, 5, 1153160. https://doi.org/10.3389/fcomp.2023.1153160

[5] Clark, R. A., Mentiplay, B. F., Hough, E., & Pua, Y. H. (2019). Three-dimensional cameras and skeleton pose tracking for physical function assessment: A review of uses, validity, current developments and Kinect alternatives. Gait & Posture, 68, 193–200. https://doi.org/10.1016/j.gaitpost.2018.11.029

[6] 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. arXiv:1812.08008

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

[8] Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., & Zitnick, C. L. (2014). Microsoft COCO: Common objects in context. In European Conference on Computer Vision (ECCV) (pp. 740–755). Springer. arXiv:1405.0312

[9] Ionescu, C., Papava, D., Olaru, V., & Sminchisescu, C. (2014). Human3.6M: Large scale datasets and predictive methods for 3D human sensing in natural environments. IEEE Transactions on Pattern Analysis and Machine Intelligence, 36(7), 1325–1339. https://doi.org/10.1109/TPAMI.2013.248

[10] Hernández, Ó. G., Morell, V., Ramon, J. L., & Jara, C. A. (2021). Human pose detection for robotic-assisted and rehabilitation environments. Applied Sciences, 11(9), 4183. https://doi.org/10.3390/app11094183

[11] Needham, L., Evans, M., Cosker, D. P., Wade, L., McGuigan, P. M., Bilzon, J. L., & Colyer, S. L. (2021). The accuracy of several pose estimation methods for 3D joint centre localisation. Scientific Reports, 11, 20673. https://doi.org/10.1038/s41598-021-00212-x

[12] Washabaugh, E. P., Shanmugam, T. A., Ranganathan, R., & Krishnan, C. (2022). Comparing the accuracy of open-source pose estimation methods for measuring gait kinematics. Gait & Posture, 97, 188–195. https://doi.org/10.1016/j.gaitpost.2022.08.008

[13] Kanazawa, A., Black, M. J., Jacobs, D. W., & Malik, J. (2018). End-to-end recovery of human shape and pose. In IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 7122–7131). https://doi.org/10.1109/CVPR.2018.00744

[14] Loper, M., Mahmood, N., Romero, J., Pons-Moll, G., & Black, M. J. (2015). SMPL: A skinned multi-person linear model. ACM Transactions on Graphics, 34(6), 248:1–248:16. https://doi.org/10.1145/2816795.2818013

[15] Fang, H.-S., Xie, S., Tai, Y.-W., & Lu, C. (2017). RMPE: Regional multi-person pose estimation. In IEEE International Conference on Computer Vision (ICCV) (pp. 2334–2343). https://doi.org/10.1109/ICCV.2017.256

[16] Fang, H.-S., Li, J., Tang, H., Xu, C., Zhu, H., Xiu, Y., Li, Y.-L., & Lu, C. (2023). AlphaPose: Whole-body regional multi-person pose estimation and tracking in real-time. IEEE Transactions on Pattern Analysis and Machine Intelligence, 45(6), 7157–7173. https://doi.org/10.1109/TPAMI.2022.3222784

[17] Lugaresi, C., Tang, J., Nash, H., McClanahan, C., Uboweja, E., Hays, M., Zhang, F., Chang, C.-L., Yong, M. G., Lee, J., Chang, W.-T., Hua, W., Georg, M., & Grundmann, M. (2019). MediaPipe: A framework for building perception pipelines. arXiv:1906.08172

[18] Bazarevsky, V., Grishchenko, I., Raveendran, K., Zhu, T., Zhang, F., & Grundmann, M. (2020). BlazePose: On-device real-time body pose tracking. arXiv:2006.10204

[19] McNally, W., Vats, K., Wong, A., & McPhee, J. (2022). Rethinking keypoint representations: Modeling keypoints and poses as objects for multi-person human pose estimation. In European Conference on Computer Vision (ECCV) (pp. 37–54). Springer. https://doi.org/10.1007/978-3-031-20068-7_3

[20] Li, W., Liu, H., Ding, R., Liu, M., Wang, P., & Yang, W. (2023). Exploiting temporal contexts with strided transformer for 3D human pose estimation. IEEE Transactions on Multimedia, 25, 1282–1293. arXiv:2103.14304

[21] Pavllo, D., Feichtenhofer, C., Grangier, D., & Auli, M. (2019). 3D human pose estimation in video with temporal convolutions and semi-supervised training. In IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 7753–7762). arXiv:1811.11742

[22] Li, J., Xu, C., Chen, Z., Bian, S., Yang, L., & Lu, C. (2021). HybrIK: A hybrid analytical-neural inverse kinematics solution for 3D human pose and shape estimation. In IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 3383–3393). arXiv:2011.14672

[23] Baradel, F., Groueix, T., Weinzaepfel, P., Brégier, R., Kalantidis, Y., & Rogez, G. (2021). Leveraging MoCap data for human mesh recovery. In International Conference on 3D Vision (3DV). IEEE. arXiv:2110.09243

[24] Li, Y., Zhang, S., Wang, Z., Yang, S., Yang, W., Xia, S.-T., & Zhou, E. (2021). TokenPose: Learning keypoint tokens for human pose estimation. In IEEE/CVF International Conference on Computer Vision (ICCV) (pp. 11313–11322). https://doi.org/10.1109/ICCV48922.2021.01112

[25] Xu, Y., Zhang, J., Zhang, Q., & Tao, D. (2022). ViTPose: Simple vision transformer baselines for human pose estimation. Advances in Neural Information Processing Systems (NeurIPS), 35, 38571–38584. arXiv:2204.12484

[26] Vakanski, A., Jun, H.-p., Paul, D., & Baker, R. (2018). A data set of human body movements for physical rehabilitation exercises. Data, 3(1), 2. https://doi.org/10.3390/data3010002

[27] Capecci, M., Ceravolo, M. G., Ferracuti, F., Iarlori, S., Monteriù, A., Romeo, L., & Verdini, F. (2019). The KIMORE dataset: KInematic assessment of MOvement and clinical scores for remote monitoring of physical REhabilitation. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 27(7), 1436–1448. https://doi.org/10.1109/TNSRE.2019.2923060

[28] Miron, A., Sadawi, N., Ismail, W., Hussain, H., & Grosan, C. (2021). IntelliRehabDS (IRDS)—A dataset of physical rehabilitation movements. Data, 6(5), 46. https://doi.org/10.3390/data6050046

[29] Liao, Y., Vakanski, A., & Xian, M. (2020). A deep learning framework for assessing physical rehabilitation exercises. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 28(2), 468–477. arXiv:1901.10435

[30] Rosique, F., Losilla, F., & Navarro, P. J. (2021). Applying vision-based pose estimation in a telerehabilitation application. Applied Sciences, 11(19), 9132. https://doi.org/10.3390/app11199132

[31] Zheng, C., Wu, W., Chen, C., Yang, T., Zhu, S., Shen, J., Kehtarnavaz, N., & Shah, M. (2023). Deep learning-based human pose estimation: A survey. ACM Computing Surveys, 56(1), 11:1–11:37. arXiv:2012.13392