Review Article
Academic Platform Journal of Engineering and Smart Systems (APJESS) 13(3), 94–107, 2025
https://dergipark.org.tr/en/pub/apjess
Received: 20-11-2024, Accepted: 28-06-2025
https://doi.org/10.21541/apjess.1588025

Advancements in Human Pose Estimation: A Review of Key Studies and Findings till 2025

[1] Uğur ÖZBALKAN, *[2] Özgür Can TURNA

[1] Department of Computer Engineering, Fenerbahçe University, Türkiye, ugur.ozbalkan@fbu.edu.tr
[1] Department of Computer Engineering, İstanbul University-Cerrahpaşa, Türkiye, ozbalkanugur95@gmail.com
*[2] Corresponding Author, Department of Computer Engineering, İstanbul University-Cerrahpaşa, Türkiye, ozgurcan.turna@iuc.edu.tr

Abstract

This paper presents an in-depth literature review that comprehensively covers the major developments, methods, architectures and datasets used in the field of human pose prediction up to 2025. The review covers a broad spectrum, starting with traditional methods, deep learning-based techniques, convolutional neural networks, graph-based approaches and more recently prominent transformer-based models. In addition to two-dimensional (2D) and three-dimensional (3D) human pose estimation methods, the paper analyses in detail the diversity of data sets, applications of Microsoft Kinect technology, real-time pose estimation systems and related architectural designs. Overall, the review of more than 120 papers shows that existing systems have made significant progress in terms of accuracy, computational efficiency and practical applications, but that there are still some challenges to overcome in complex scenarios such as multiple person detection, occlusion problems and outdoor environments. This in-depth analysis highlights current trends in the field, future research directions and potential applications.

Keywords: Human Pose Estimation, Deep Learning, Microsoft Kinect, Real-time Applications

1. INTRODUCTION

Human pose estimation has evolved significantly over time, transitioning from conventional methods to advanced deep learning techniques. Early approaches utilized pictorial structures and traditional convolutional neural networks [1]. Since then, the field has progressed to more sophisticated methods such as DeepPose, Adversarial PoseNet, and OpenPose [2]. Recent advancements include the integration of stacked hourglass networks and part proximity fields [3]. Researchers have concentrated on enhancing network architectures, optimizing training, and developing post-processing techniques to enhance performance [4]. The field has expanded to encompass both 2D and 3D pose estimation, with applications ranging from human-computer interaction to sports analysis [5]. Notwithstanding the significant progress that has been made, challenges remain in pose estimation for multiple individuals, especially in outdoor environments [6]. Ongoing research continues to address these limitations and explore new approaches to improve accuracy and robustness [7,8].

This study aims to provide a comprehensive overview of technological and methodological developments in the field of human pose estimation. In this context, seven main research questions have been identified and articles retrieved from scientific databases have been meticulously analyzed in line with these questions. A broad perspective is presented, from traditional methods to modern deep learning approaches, from model architecture to the diversity of datasets, as well as real-time applications and the integration of sensor technologies. The results evaluate the current application areas and performance criteria of human pose estimation and provide important clues to the potential for future research and application.

2. MATERIALS AND METHODS

Firstly, the research questions were determined, and the articles related to the research questions were analyzed by accessing the relevant databases through Google Scholar. A total of 120 articles were analyzed throughout the study. The research questions prepared for the study are presented in Table 1.

Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/)

Table 1. Research questions of this study

	Research questions (RQ)
RQ1	What are the major developments in human pose estimation?
RQ2	Which models are used for 2-dimensional and 3-dimensional human pose estimation?
RQ3	Which data sets are used on human pose estimation?
RQ4	What work is done with Microsoft Kinect technology?
RQ5	Which architectures are related to human pose estimation?
RQ6	Which methods are used in real-time human pose estimation studies and what are the results?
RQ7	What are the recent advances in human pose prediction and how do they compare with past developments?

3. RESULTS AND DISCUSSION

Articles related to the identified research questions were reviewed.

3.1. Major Developments in Human Pose Estimation

The field of Human Pose Estimation (HPE) has evolved significantly over the years, driven by advances in computer vision and deep learning. Table 2 below provides a chronological summary of key developments in Human Pose Estimation, highlighting key contributions that have shaped the landscape of this field.

3.2. 2-Dimentional and 3-Dimensional Human Pose Estimation Models

Human Posture Prediction: Articulated human postures are often represented by a combination of singular terms and graphical models that are combinations of body parts or visual structures [21,22,23]. Recently, significant progress has been made with the introduction of ConvNets to learn better feature representation [11,22,24,25]. Figure 1 includes images of applications related to recent technological advancements and new datasets, reflecting the growing interest in this field within the scientific community.

The human body, with its lots of limbs and joints, forms a sophisticated system, and ascertaining their 3D spatial positions accurately can be a demanding task, not only for artificial systems but also for humans. Marinoiu and the team have investigated how humans grasp the 3D visual position space and how this understanding can be related to the physical 3D space in which we live [27]. In this study, the researchers created a dataset that not only encompassed 2D and 3D positions but also included synchronized eye movement recordings. They displayed various human body configurations and measured how accurately the 3D positions were reconstructed. On average, they discovered that people's ability to reconstruct 3D positions based on visual stimuli provided in laboratory environments was not significantly better than existing computer vision algorithms [27].

Despite the encountered challenges, automatic methods provide valuable solutions for executing a specific task. Approaches based on modelling aim to surmount these obstacles by utilizing human body models grounded on prior knowledge. Figure 2 represents the most common 3D human body models, which are skeletal models encompassing both shape and structural features, and both types of models delineate kinematic details.

Conventional approaches typically address 2D human pose estimation using tree-based models. These models include two components: one for identifying body joints and another for characterizing binary relationships between pairs of body joints [28].

In the past ten years, various methods, many of which employ deep learning techniques, have been devised and have significantly enhanced performance on established benchmarks [29].

Figure 1. A comprehensive review of real-world applications related to human motion analysis and pose estimation [26]

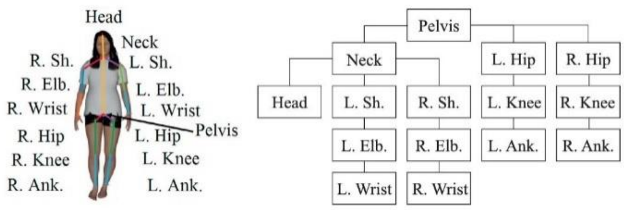


Figure 2. Body model with 15 joints and tree structure [26]

Table 2. Major developments in human pose estimation

Year	Method / Approach	Contributions / Explanations
2014	DeepPose: Deep Neural Network Based Pose Estimation	Utilizing deep neural networks (DNNs), the author approached pose estimation as a regression problem. Considerable advancement has been made in the realm of predicting joint positions through the utilization of cascade regression [9].
2016	DeeperCut: Multi-Person Pose Prediction	The integration of large-scale datasets and deep learning architectures has been demonstrated to present an effective approach to overcoming overlap and proximity problems in crowded scenes [10].
2016	Stacked Hourglass Networks	The innovative convolutional architecture of the model enabled the capture of spatial relationships at different scales, with the result that more accurate pose estimation was achieved by preserving spatial hierarchies [11].
2016	Convolutional Pose Machines	The introduction of a sequential prediction framework has been demonstrated to indirectly model long-range dependencies between body joints, thereby enhancing the accuracy of articulated pose prediction [12].
2017	Weakly-Supervised 2D & 3D Pose Estimation	Large-scale 2D annotation data was utilized to develop a unified framework integrating 2D and 3D pose estimation tasks [13].
2017	3D Pose Estimation from 2D Data	As demonstrated in [14], a significant reduction in error rates has been achieved in the estimation of 3D human poses from a single image, through the implementation of a volumetric approach that progresses from coarse to fine.
2018	Simple Baselines	It is evident that the adoption of simplified and optimized architectures has enabled COCO to attain competitive outcomes on the benchmark, thereby substantiating the hypothesis that attaining high performance is indeed feasible through the utilization of model simplicity [15].
2019	High-Resolution Networks	Sun et al. presented a detailed method that improves the accuracy of key point heat maps by preserving high-resolution representations in the pose estimation process [16].
2020	Graph Convolutional Networks	In the field of video-based pose estimation, the study employed graph convolutional networks to model the structural relationships between body joints, thereby facilitating integrated pose prediction with realistic frame reproduction [17].
2021	Decoupled Representations for Motion Forecasting	Parsaeifard et al. developed a novel and effective approach to human motion prediction by utilizing decoupled representations to capture complex spatiotemporal interactions [18].
2022	Bilateral Pose Transformers	Yen et al. presented a pioneering model that facilitates more precise detection of human anatomical key points by integrating advanced transformer architectures [19].
2023	Transformers in Pose Estimation	The employment of transformer architectures in both 2D and 3D pose estimation has facilitated the development of efficient solutions to increasingly complex and challenging pose estimation tasks [20]. In recent times, this approach has been adopted in numerous novel studies.

Table 3. Datasets used in the literature on human pose estimation

Name of study	Year	Datasets used
DeepPose: Human Pose Estimation via Deep Neural Networks [9]	2014	Frames Labeled In Cinema (FLIC) - Leeds Sports Pose (LSP)
2D Human Pose Estimation: New Benchmark and State of the Art Analysis [30]	2014	MPII Human Pose
Multi-source Deep Learning for Human Pose Estimation [31]	2014	Leeds Sports Pose (LSP) - PARSE UIUC People
3D Pictorial Structures for Multiple Human Pose Estimation [32]	2014	HumanEva-I - KTH Multiview Football Dataset II
Flowing ConvNets for Human Pose Estimation in Videos [33]	2015	BBC Pose - Extended BBC Pose - ChaLearn - Poses in the wild (PiW) - Frames Labeled in Cinema (FLIC)
Stacked Hourglass Networks for Human Pose Estimation [11]	2016	Frames Labeled in Cinema (FLIC) - MPII Human Pose
Structured Feature Learning for Pose Estimation [34]	2016	Frames Labeled in Cinema (FLIC) - Leeds Sports Pose (LSP)
Convolutional Pose Machines [35]	2016	MPII Human Pose - Leeds Sports Pose (LSP) - Frames Labeled in Cinema (FLIC)

Name of study	Year	Datasets used
End-to-End Learning of Deformable Mixture of Parts and Deep Convolutional Neural Networks for Human Pose Estimation [36]	2016	Leeds Sports Poses (LSP) - Frames Labeled in Cinema (FLIC) - The Image Parse (PARSE)
Human Pose Estimation with Iterative Error Feedback [37]	2016	MPII Human Pose - Leeds Sports Pose (LSP)
Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields [38]	2017	MPII human multi-person - COCO 2016
Multi-Context Attention for Human Pose Estimation [39]	2017	MPII Human Pose - Extended Leeds Sports Poses (LSP)
Learning Feature Pyramids for Human Pose Estimation [40]	2017	MPII Human Pose - Leeds Sports Pose (LSP)
Recurrent Human Pose Estimation [41]	2017	MPII Human Pose - Leeds Sports Pose (LSP)
Vnect: Real-time 3D Human Pose Estimation with a Single RGB Camera [42]	2017	H3.6M dataset - MPI-INF-3DHP dataset
DensePose: Dense Human Pose Estimation In The Wild [43]	2018	Dense-Pose-COCO (Common Objects in Context)
PoseTrack: A Benchmark for Human Pose Estimation and Tracking [44]	2018	MPII Human Pose - Posetrack: Joint multi-person pose estimation and tracking.
Simple Baselines for Human Pose Estimation and Tracking [45]	2018	Dense-Pose-COCO(Common Objects in Context) Posetrack: Joint multi-person pose estimation and tracking.
3D Human Pose Estimation in the Wild by Adversarial Learning [46]	2018	Human3.6M - MPI-INF3DHP - MPII Human Pose
Exploiting temporal information for 3D human pose estimation [47]	2018	Human3.6M
Fast Human Pose Estimation [48]	2019	MPII Human Pose - Leeds Sports Pose (LSP)
Deep High-Resolution Representation Learning for Human Pose Estimation [16]	2019	MPII Human Pose - Dense-Pose-COCO (Common Objects in Context) - Posetrack: Joint multi-person pose estimation and tracking
Rethinking on Multi-Stage Networks for Human Pose Estimation [49]	2019	COCO - MPII Human Pose
Distribution-Aware Coordinate Representation for Human Pose Estimation [50]	2020	COCO - MPII Human Pose
SimPoE: Simulated Character Control for 3D Human Pose Estimation [51]	2021	Human3.6M - In-house human motion
Deep Dual Consecutive Network for Human Pose Estimation [52]	2021	PoseTrack2017 - PoseTrack2018
3D Human Pose Estimation with Spatial and Temporal Transformers [53]	2021	Human3.6M - MPI-INF-3DHP
MHFormer: Multi-Hypothesis Transformer for 3D Human Pose Estimation [54]	2022	Human3.6M - MPI-INF-3DHP
ViTPose: Simple Vision Transformer Baselines for Human Pose Estimation [55]	2022	ViTPose
EHPE: Skeleton Cues-based Gaussian Coordinate Encoding for Efficient Human Pose Estimation [56]	2022	MS COCO - MPII Human Pose
SportsPose - A Dynamic 3D sports pose dataset [57]	2023	SportsPose
DiffPose: Toward More Reliable 3D Pose Estimation [58]	2023	DiffPose (Compared with Human3.6M and MPIINF-3DHP datasets)
HEViTPose: High-Efficiency Vision Transformer for Human Pose Estimation [59]	2023	HEViTPose (Compared with MPII Human Pose and COCO datasets)
MDST: 2-D Human Pose Estimation for SISO UWB Radar Based on Micro-Doppler Signature via Cascade and Parallel Swin Transformer [60]	2024	HPSUR
End-to-End 3D Human Pose Estimation Network With Multi-Layer Feature Fusion [61]	2024	3DPW, AGORA and MPII Human Pose
Enhancement and optimization of human pose estimation with multi-scale spatial attention and adversarial data augmentation [62]	2024	MS COCO - MPII Human Pose

3.3. Datasets used on Human Pose Estimation

Data sets come to the forefront in studies on Human Pose Estimation. The data sets used in the reviewed studies are given in Table 3.

3.4. An Overview of The Most Widely Employed Datasets

An overview of the most widely employed datasets in the research can be outlined as follows:

3.4.1. MPII Human Pose

The MPII Human Pose Dataset consists of approximately 25,000 images, including 15,000 training, 3,000 validation and 7,000 test samples [63]. The use of this dataset in articles according to years is shown in Figure 3.

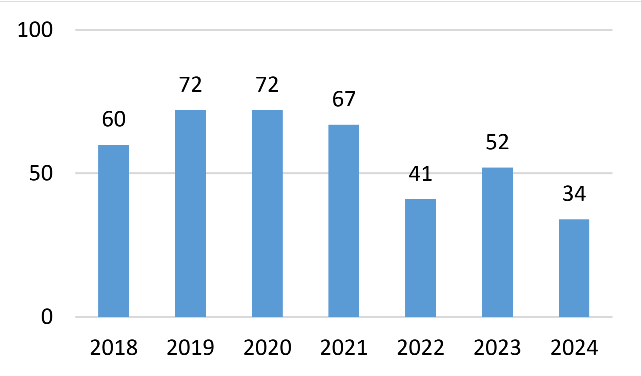


Figure 3. MPII dataset number of articles by year [63]

3.4.2. The Leeds Sports Poses (LSP)

The LSP dataset consists of 2000 athlete images, 1000 for testing and 1000 for training, with 14 joint positions in each image [64]. The use of this dataset in articles according to years is shown in Figure 4.

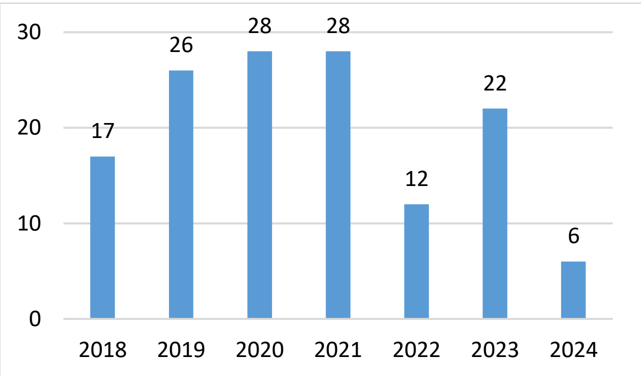


Figure 4. LSP dataset number of articles by year [64]

3.4.3. MPI-INF-3DHP

The MPI-INF-3DHP dataset is designed for 3D human body pose estimation, consisting of more than 1.3 million frames taken from 14 camera perspectives [65]. The use of this dataset in articles according to years is shown in Figure 5.

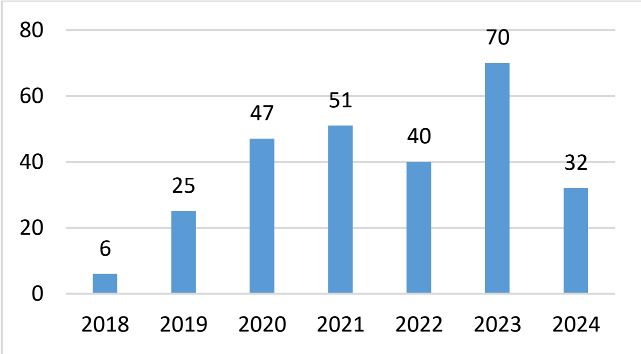


Figure 5. MPI-INF-3DHP dataset number of articles by year [65]

3.4.4. Human3.6M

The Human3.6M is a dataset of 3.6 million human poses covering the movements of 11 professional actors in 17 situations [66]. The use of this dataset in articles according to years is shown in Figure 6.

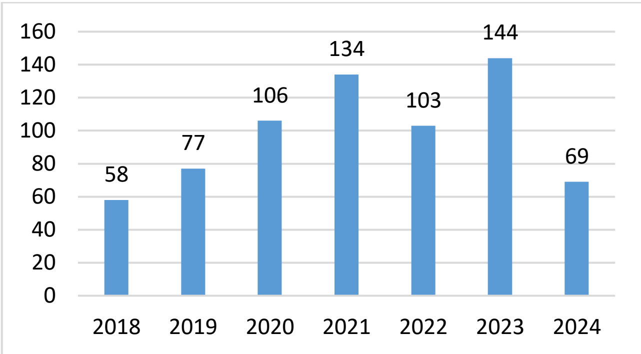


Figure 6. Human3.6M dataset number of articles by year [66]

3.4.5. PoseTrack

PoseTrack is a dataset of 514 videos containing 66,374 frames from 300 training, 50 validation and 208 test sets [67]. The use of this dataset in articles according to years is shown in Figure 7.

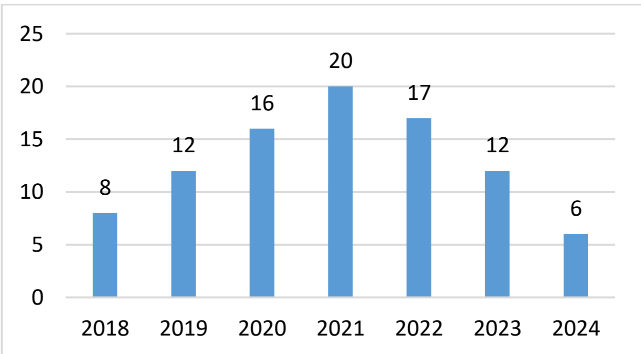


Figure 7. PoseTrack dataset number of articles by year [67]

3.4.6. Coco

The MS COCO (Microsoft Common Objects in Context) dataset contains 328,000 images for the detection of key points, objects [68]. The use of this dataset in articles according to years is shown in Figure 8.

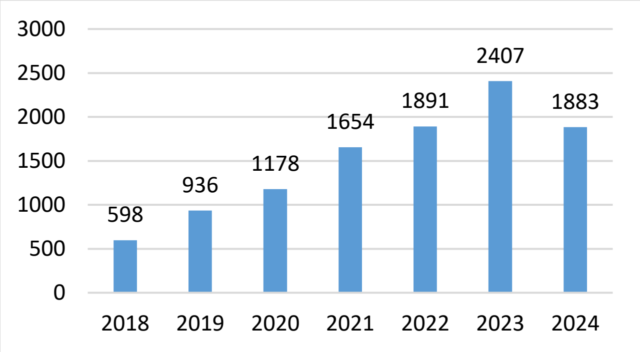


Figure 8. Coco dataset number of articles by year [68]

3.4.7. DensePose (DensePose-COCO)

The DensePose-COCO dataset contains 50,000 COCO images with manual annotations added to it [69]. The use of this dataset in articles according to years is shown in Figure 9.

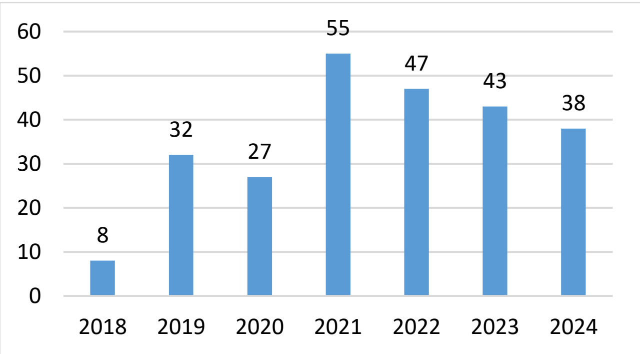


Figure 9. DensePose dataset number of articles by year [69]

3.4.8. Frames Labelled in Cinema (FLIC)

The FLIC dataset consists of 5,003 Hollywood movie images obtained using a person detector on selected frames [70]. The use of this dataset in articles according to years is shown in Figure 10.

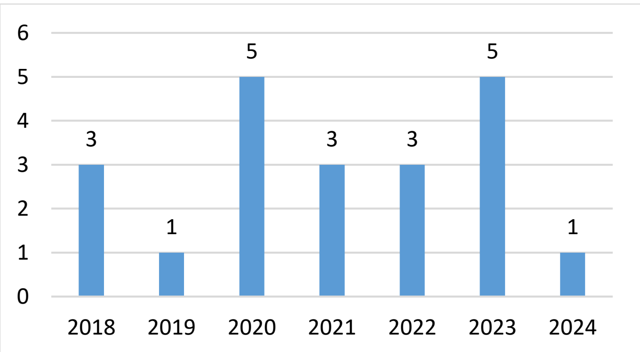
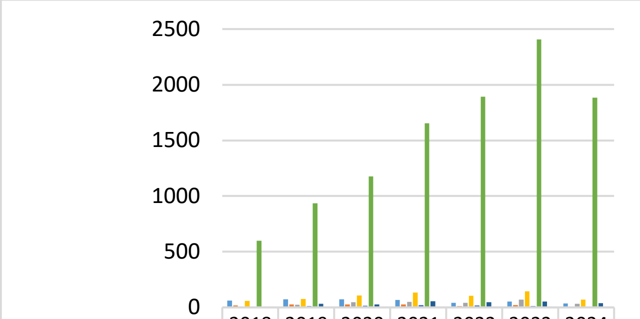


Figure 10. FLIC dataset number of articles by year [70]

The number of articles in which the analyzed data sets were used in the last 5 years is given in Figure 11.



	2018	2019	2020	2021	2022	2023	2024
MPII	60	72	72	67	41	52	34
LSP	17	26	28	28	12	22	6
MPI-INF-3DHP	6	25	47	51	40	70	32
Human3.6M	58	77	106	134	103	144	69
Posetrack	8	12	16	20	17	12	6
Coco	598	936	1178	1654	1891	2407	1883
DensePose-Coco	8	32	27	55	47	53	38
FLIC	3	1	5	3	3	5	1

Figure 11. Number of articles of datasets by year

The visualizations of the MPII, LSP, FLIC, Dense Pose and Dense Pose-Coco datasets are shown in Figures 12,13,14,15, and 16.

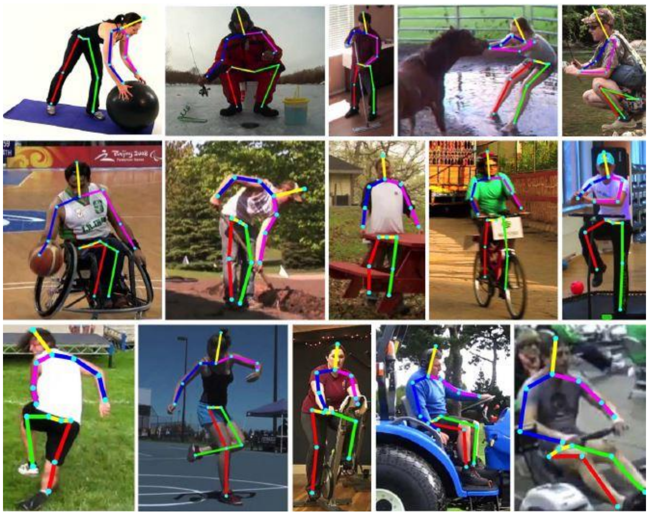


Figure 12. MPII [41]

Figure 13. LSP [48]

Figure 14. Frames Labeled In Cinema (FLIC) [9]

Figure 15. DensePose [43]

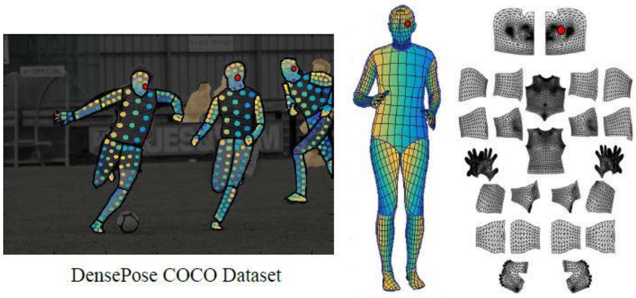


Figure 16. DensePose-COCO [43]

Recent research in human pose estimation (HPE) has focused on addressing dataset limitations and improving model performance. Existing datasets often lack diversity in subjects, poses, and environments, hindering generalization to real-world scenarios [71]. To address this, researchers have proposed methods like joint harmonization and scale normalization to improve cross-dataset performance [71]. The field has seen advancements in both 2D and 3D pose estimation, with deep learning-based approaches achieving high accuracy [4, 72]. However, challenges remain, including insufficient training data, depth ambiguities, and occlusion [72]. New datasets, such as COCO-WholeBody, have been introduced to provide more comprehensive annotations for whole-body pose estimation [73]. Additionally, self-supervised approaches have been developed to address the lack of labeled data for certain activities [74]. HPE has found applications in various fields, including sports and physical exercise [75].

3.5. Works Done with Microsoft Kinect Technology

Microsoft Kinect technology has been widely applied in various fields, including multimedia computing, robotics, and healthcare. Zeng and Lun both highlight its potential for human-computer interaction and motion recognition [76, 77]. Chow and Ellaithy discuss its use in improving the quality of 3D camera technology and for indoor navigation in robotics [78, 79]. Ismail and Chang focus on its applications in human motion capture and rehabilitation, with the latter comparing its performance to a high-fidelity optical system [80, 81]. Molyneaux and Bueno explore its use in 3D reconstruction and metrological evaluation, respectively [82, 83]. These studies collectively demonstrate the versatility and potential of Microsoft Kinect technology in various domains.

Information on other studies based on Kinect is given in Table 4.

3.6. Architectures related to Human Pose Estimation

A range of user-friendly solutions for specific pose detection have been developed. Bao introduced PoseNAS, a network architecture that simultaneously designs a pose encoder and decoder, achieving state-of-the-art performance [91]. Rogez proposed the LCR-Net++ architecture for 2D and 3D human pose estimation, which shows better results than the methods used [92]. Ning presented LightTrack, a framework for online human pose tracking that outperforms other online methods [93]. Regarding Human Pose Estimation, Dyce on the depth of the Kinect camera and Oleinikov on task-based methods through a control system and both focusing on user-friendly interfaces [94, 95]. Dimitrijevic took an approach to detecting people's silhouettes, while Buys took an RGB-D based approach to human pose estimation [96, 97]. Cao used the OpenPose system for real-time human pose estimation [98].

The names and architectural Information of some of the articles analyzed regarding the architectures analyzed or used are given in Table 5

Table 4. Information on other studies based on Kinect

Name of study	Year	A brief summary
Performance Analysis of Body Tracking with the Microsoft Azure Kinect [84]	2021	This paper investigates experimentally by performing body tracking with a Kinect camera.
Human Following Robot using Kinect in Embedded Platform [85]	2022	It tries to detect an individual target by tracking the movements in the environment with Kinect technology.
Virtual Yoga System Using Kinect Sensor [86]	2022	Microsoft Kinect sensor was used to detect, monitor and inform the user about the user's movements related to yoga.
Classification of human movements by using Kinect sensor [87]	2023	Different methods were proposed by comparing with the methods in the literature.
A real-time multi view gait-based automatic gender classification system using Kinect sensor [88]	2023	An automatic gender classification system was proposed using Kinect technology.
Human–machine interaction and implementation on the upper extremities of a humanoid robot [89]	2024	The paper uses Kinect depth sensor and MediaPipe framework to obtain 3D joint positions for controlling the movement of a humanoid robot's upper extremities.
Engineering User Interfaces for Tailored and Monitored Movement Rehabilitation Programs [90]	2024	The Microsoft Kinect can be used to support tailored and monitored movement rehabilitation programs.

Table 5. Architectures used and analyzed

Name of study	Architectures used and analyzed	Year
Pose-native Network Architecture Search for Multi-person Human Pose Estimation [91]	Pose-native Network Architecture Search (PoseNAS)	2020
PoseTED: A Novel Regression-Based Technique for Recognizing Multiple Pose Instances [99]	PoseTED	2021
Yoga Pose Detection Using Deep Learning Techniques [100]	OpenPose	2021
Analysis of Deep Learning Based Pose Estimation Techniques for Locating Landmarks on Human Body Parts [101]	BlazePose	2021
Human Pose Estimation using Deep Learning Techniques [102]	MoveNet, Convolutional Pose Machines, Hourglass Network	2022
ConvPose: A modern pure ConvNet for human pose estimation [103]	ConvPose	2023
AnatPose: Bidirectionally learning anatomy-aware heatmaps for human pose estimation [104]	AnatPose	2024

Human pose estimation from videos is crucial in a variety of applications, including physical exercise measurement, sign language recognition, and whole-body control. It can serve as the foundation for yoga, dance, and fitness applications. Moreover, it can enable the superimposition of digital content and information on the physical world through augmented reality [105].

3.7. Methods used in Real-Time Human Pose Estimation Studies

BlazePose is used in areas such as fitness tracking in real-time human pose estimation over 33 body key points and is shown BlazePose results in yoga and fitness poses in Figure 17 [105].

Figure 17. BlazePose results in yoga and fitness poses [105]

BlazePose Full has been found to outperform OpenPose in Yoga/Fitness use cases, although BlazePose models perform slightly worse than OpenPose when working with AR data [105]. Additionally, BlazePose can run 25-75 times faster on a single midrange phone CPU than OpenPose on a 20-core desktop CPU, depending on the level of quality required [105].

The MediaPipe library is used to detect key points of the human body in an input image. The output of this process is a list of coordinates in the X, Y, and Z axes for 33 main key points of the human body. This list of coordinates identifies the position of each major body part in the input image, enabling the creation of an accurate skeletal orientation of the user. Figure 18 shows 33 landmarks, indexed from 0 to 32, indicating the major joints and locations in the human body using the MediaPipe library [105].

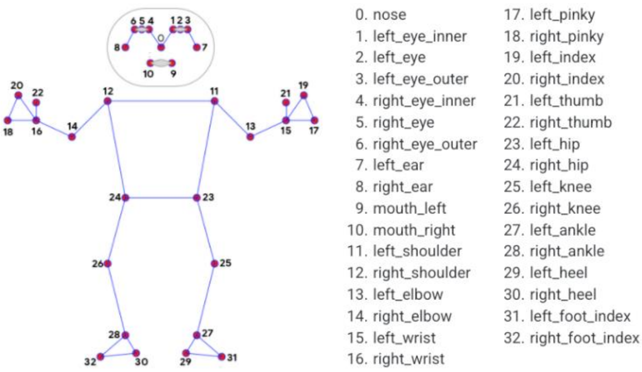


Figure 18. Location Signs in MediaPipe [105]

Figure 19 illustrates the MediaPipe graph with transparent boxes representing the computational nodes or calculators, solid boxes representing the external input/output to the graph, and lines entering from above and exiting from below the nodes representing the input and output flows, respectively. Ports on the left side of some nodes indicate ingress-side packets [106].

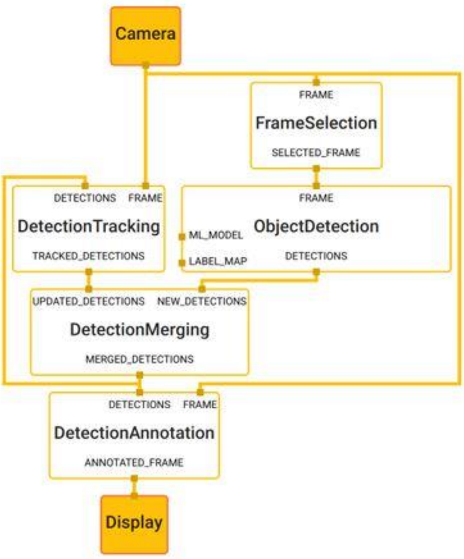


Figure 19. Object detection using MediaPipe [106]

Figure 20. True/False Pose Demo [107]

Providing real-time feedback to a person performing a sports movement based on an image can help ensure that the movement is performed correctly and prevent injuries resulting from incorrect movements. As shown in Figure 20, Radha Tawar et al. proposed a Yoga Pose Detection system that utilizes OpenCV and MediaPipe with RGB to achieve this goal [107].

3.8. Recent Advances in Human Pose Estimation and Comparisons with Past Developments

Advances in human pose estimation have significantly improved accuracy and efficiency in both 2D and 3D domains. Deep learning techniques, particularly convolutional neural networks, have revolutionized the field, enabling more accurate key point localization and body part detection [2, 72]. Researchers have focused on optimizing network architectures, refining training methods, and developing post-processing strategies to improve performance [4]. Significant progress has been made in monocular 3D pose estimation, with average joint errors reduced to around 20 mm [108]. However, challenges remain, including depth ambiguity, occlusion, and insufficient training data [72]. Recent studies have explored solutions for complex scenarios, such as multi-person pose estimation and motion capture [109]. The field has expanded to include 3D mesh reconstruction and body shape estimation, where both parametric and non-parametric approaches have been investigated [110]. These advances have broad applications in robotics, entertainment, healthcare, and human-computer interaction [111, 112].

Recent advances in human pose estimation have been driven by deep learning techniques, particularly convolutional and recurrent neural networks [113]. Research has progressed from single-person to multi-person pose estimation, with top-down and bottom-up approaches being explored [14]. Studies have compared different algorithms and methods in both 2D and 3D domains [115], evaluating the robustness of models such as Posenet, MediaPipe and BlazePose [116]. Improvements in accuracy and computational efficiency have been achieved through framework optimizations, such as those implemented in MediaPipe [117]. Advanced techniques incorporating context-aware features and machine learning classifiers such as XGBoost have shown promising results in pose estimation and event classification tasks [118]. Challenges remain in dealing with occlusions and complex scenes, with future research directions focusing on improving model architectures and integrating multimodal information [119, 120].

4. CONCLUSION

A comprehensive literature review in the field of human pose estimation reveals a broad perspective ranging from traditional methods to today's deep learning and transformer-based approaches. The reviewed studies have shown that there has been a significant paradigm shift in the field, with conventional methods being replaced by advanced neural network architectures that provide high accuracy and efficiency. In particular, the superior performance of convolutional neural networks, graph-based models and recently popularized transformer architectures in 2D and 3D pose estimation has attracted the attention of researchers and practitioners.

The literature reviewed shows that the diversity and enrichment of datasets is critical for model training and generalization. The use of MPII, LSP, Human3.6M, COCO and other datasets plays an important role in testing and optimizing models in real-world scenarios. In addition, the integration of advanced sensor technologies such as Microsoft Kinect is another important factor that improves the accuracy and practical applicability of human pose estimation, especially in real-time applications.

However, several challenges highlighted in the literature show that the field is still evolving. In particular, accurate multi-person detection in complex scenes, occlusion problems, depth uncertainty, and sustainability of model performance under harsh environmental conditions such as outdoor environments stand out as the main problems to be solved. To overcome such problems, more scalable and computationally efficient models, multimodal data integration and hybrid methods should be investigated.

Future research is expected to focus on developing further innovative solutions from both a theoretical and practical perspective. In particular, the integration of deep learning methods as well as modern approaches such as data augmentation techniques, transfer learning and self-supervised learning strategies can play an important role in overcoming current limitations in human pose prediction. Furthermore, the development of algorithms optimized for real-time applications and low-power mobile platforms will enable wider adoption of advances in the field.

In conclusion, this study provides a comprehensive overview of current advances in human pose prediction, while providing guidance for future research. The new models to be developed are expected to provide more integrated, flexible and robust frameworks to address the challenges faced in both academic and industrial applications. In this context, advances in human pose estimation will continue to provide innovative solutions in a wide range of disciplines such as human-computer interaction, health, sports analysis, robotics and augmented reality.

ACKNOWLEDGEMENT
This work is a part of the Ph.D. thesis titled "Simultaneous Feedback Generation with Artificial Intelligence Modeling via Video Tracking in Physical Exercises" at the Institute of Graduate Studies, Istanbul University-Cerrahpasa, Istanbul, Turkey.

Author contributions: Uğur Özbalkan conducted the literature review, data collection, and original draft preparation. Özgür Can Turna supervised the study, contributed to the methodology, and performed review and editing of the manuscript.
Conflict of Interest: No conflict of interest was declared by the authors.

REFERENCES

[1] Khan, Naimat Ullah, and Wanggen Wan. "A review of human pose estimation from single image." 2018 international conference on audio, language and image processing (ICALIP). IEEE, 2018.

[2] Sen, Siddharth, Somya Maheshwari, and Ankit Kumar. "Review of Recent Developments in Human Pose Estimation." Proceedings of the International Conference on Innovative Computing & Communication (ICICC). 2021.

[3] Munea, Tewodros Legesse, et al. "The progress of human pose estimation: A survey and taxonomy of models applied in 2D human pose estimation." IEEE Access 8 (2020): 133330-133348.

[4] Chen, Haoming, et al. "2D Human pose estimation: A survey." Multimedia systems 29.5 (2023): 3115-3138.

[5] Sarafianos, Nikolaos, et al. "3d human pose estimation: A review of the literature and analysis of covariates." Computer Vision and Image Understanding 152 (2016): 1-20.

[6] Gong, Wenjuan, et al. "Human pose estimation from monocular images: A comprehensive survey." Sensors 16.12 (2016): 1966.

[7] Liu, Zhao, et al. "A survey of human pose estimation: the body parts parsing based methods." Journal of Visual Communication and Image Representation 32 (2015): 10-19.

[8] Kharb, Utkarsh, and Tanay Gautam. "Review and analysis of various human pose estimation models." Proceedings of the Advancement in Electronics & Communication Engineering (2022).

[9] Toshev, Alexander, and Christian Szegedy. "Deeppose: Human pose estimation via deep neural networks." Proceedings of the IEEE conference on computer vision and pattern recognition. 2014.

[10] Insafutdinov, Eldar, et al. "Deepercut: A deeper, stronger, and faster multi-person pose estimation model." Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part VI 14. Springer International Publishing, 2016.

[11] Newell, A., Yang, K., Deng, J. "Stacked Hourglass Networks for Human Pose Estimation". In: European Conference on Computer Vision, 2016.

[12] Wei, Shih-En, et al. "Convolutional pose machines." Proceedings of the IEEE conference on Computer Vision and Pattern Recognition. 2016.

[13] Zhu, Yi, et al. "Soft proposal networks for weakly supervised object localization." Proceedings of the IEEE international conference on computer vision. 2017.

[14] Pavlakos, Georgios, et al. " Coarse-to-fine volumetric prediction for single-image 3D human pose." Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.

[15] Xiao, Bin, Haiping Wu, and Yichen Wei. "Simple baselines for human pose estimation and tracking." Proceedings of the European conference on computer vision (ECCV). 2018.

[16] Sun, Ke, et al. "Deep high-resolution representation learning for human pose estimation." Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2019.

[17] Zhao, Yang, and Yong Dou. "Pose-forecasting aided human video prediction with graph convolutional networks." IEEE Access 8 (2020): 147256-147264.

[18] Parsaeifard, Behnam, et al. "Learning decoupled representations for human pose forecasting." Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021.

[19] Yen, Chia Chen, Tao Pin, and Hongmin Xu. "Bilateral pose transformer for human pose estimation." Proceedings of the 4th International Symposium on Signal Processing Systems. 2022.

[20] Yin, He, Chang Lv, and Yeqin Shao. "3D Human Pose Estimation Based on Transformer." Journal of Physics: Conference Series. Vol. 2562. No. 1. IOP Publishing, 2023.

[21] Yang, Y., Ramanan, D., "Articulated pose estimation with f lexible mixtures-of-parts", In CVPR, 2011. 2.

[22] Chen, X., Yuille, A. L., "Articulated pose estimation by a graphical model with image dependent pairwise relations", In NIPS, 2014. 2.

[23] Pishchulin, L., Andriluka, M., Gehler, P., Schiele, B., "Poselet conditioned pictorial structures", In CVPR, 2013.

[24] Tompson, J., Goroshin, R., Jain, A., LeCun, Y., Bregler, C., "Efficient object localization using convolutional networks", In CVPR, 2015.

[25] Wei, S. E., Ramakrishna, V., Kanade, T., Sheikh, Y., "Convolutional pose machines", In CVPR, 2016.

[26] Sarafianosa, N., Boteanub, B., Lonescub, B., Kakadiarisa, L. A, "3D Human Pose Estimation: A Review of the Literature and Analysis of Covariates", Computer Vision and Image Understanding 152 (2016) 1–20.

[27] Marinoiu, E. , Papava, D. , Sminchisescu, C., "Pictorial human spaces: how well do humans perceive a 3D articulated pose?", In Proc. IEEE International Conference on Computer Vision, Sydney, Australia, 2013, pp. 1289–1296.

[28] Yang, W., Ouyang, W., Wang, X., Ren, J., Li, H. and Wang, X, "3d human pose estimation in the wild by adversarial learning", Proceedings of the IEEE conference on computer vision and pattern recognition, 2018.

[29] Wang, J., Tan, S., Zhen, X., Xu, S., Zheng, F., He, Z., Shao, L., "Deep 3D human pose estimation: A review", Computer Vision and Image Understanding, 210, 2021, 103225.

[30] Andriluka, M., Pishchulin, L., Gehler, P., Schiele, B., "2D Human Pose Estimation: New Benchmark and State of the Art Analysis", IEEE Conference on Computer Vision and Pattern Recognition, 2014.

[31] Ouyang, W., Chu, X., Wang, X., "Multi-source Deep Learning for Human Pose Estimation", IEEE Conference on Computer Vision and Pattern Recognition, 2014.

[32] Belagiannis V., Amin, S., Andriluka, M., Schiele, B., Navab, N., Ilic, S., "3D Pictorial Structures for Multiple Human Pose Estimation", IEEE Conference on Computer Vision and Pattern Recognition, 2014.

[33] Pfister, T., James, C. and Andrew Zisserman, "Flowing convnets for human pose estimation in videos", Proceedings of the IEEE international conference on computer vision, 2015.

[34] Chu, Xiao, et al. "Structured feature learning for pose estimation" Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.

[35] Wei, S., Ramakrishna, V., Kanade, T., Sheikh, Y., "Convolutional pose machines", Proceedings of the IEEE conference on Computer Vision and Pattern Recognition, 2016.

[36] Yang, W., Ouyang, W., Li, H., Wang, X., "End-to-End Learning of Deformable Mixture of Parts and Deep Convolutional Neural Networks for Human Pose Estimation", IEEE Conference on Computer Vision and Pattern Recognition, 2016.

[37] Carreira, J., Agrawal, P., Fragkiadaki, K., Malik, J., "Human pose estimation with iterative error feedback", Proceedings of the IEEE conference on computer vision and pattern recognition, 2016.

[38] Chao, Z., Simon, T., Wei, S., Sheikh, Y., "Realtime multi-person 2d pose estimation using part affinity fields", Proceedings of the IEEE conference on computer vision and pattern recognition, 2017.

[39] Chu, X., Yang, W., Ouyang, W., Ma, C., Yuille, A. L., Wang, X., "Multi-context attention for human pose estimation", Proceedings of the IEEE conference on computer vision and pattern recognition, 2017.

[40] Yang, W., Li, S., Ouyang, W., Li, H., Wang, X., "Learning feature pyramids for human pose estimation", proceedings of the IEEE international conference on computer vision, 2017.

[41] Belagiannis, V., Zisserman, A., "Recurrent Human Pose Estimation", IEEE 12th International Conference on Automatic Face & Gesture Recognition, 2017.

[42] Mehta, D., Sridhar, S., Sotnychenko, O., Rhodin, H., Shafiei, H., Seidel, H., Xu, W., Casas, D., Theobalt, C., "Vnect: Real-time 3d human pose estimation with a single rgb camera", Acm transactions on graphics (tog) 36.4 (2017): 1-14.

[43] Guler, R. A., Neverova, N., Kokkinos, I., "Densepose: Dense human pose estimation in the wild", Proceedings of the IEEE conference on computer vision and pattern recognition, 2018.

[44] Andriluka, M., Iqbal, U., Insafutdinov, E., Pishchulin, L., Milan, A., Gall, J., Schiele, B., "Posetrack: A benchmark for human pose estimation and tracking", Proceedings of the IEEE conference on computer vision and pattern recognition, 2018.

[45] Xiao, B., Wu, H., Wei, Y., "Simple baselines for human pose estimation and tracking", Proceedings of the European conference on computer vision (ECCV), 2018.

[46] Yang, W., Ouyang, W., Wang, X., Ren, J., Li, H., Wang, X., "3d human pose estimation in the wild by adversarial learning", Proceedings of the IEEE conference on computer vision and pattern recognition, 2018.

[47] Hossain, M. R. I., Little, J. J., "Exploiting temporal information for 3d human pose estimation", Proceedings of the European conference on computer vision (ECCV), 2018.

[48] Zhang, F., Zhu, X., Ye, M., "Fast human pose estimation", Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2019.

[49] Li, W., Wang, Z., Yin, B., Peng, Q., Du, Y., Xiao, T., Yu, G., Lu, H., Wei, Y., Sun, J., "Rethinking on multi-stage networks for human pose estimation", arXiv preprint arXiv:1901.00148 (2019).

[50] Zhang, F., Zhu, X., Dai, H., Ye, M., Zhu, C., "Distribution-aware coordinate representation for human pose estimation", Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2020.

[51] Yuan, Y., Wei, S., Simon, T., Kitani, K., Saragih, J., "Simpoe: Simulated character control for 3d human pose estimation", Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2021.

[52] Liu, Z., Chen, H., Feng, R., Wu, S., Ji, S., Yang, B., Wang, X., "Deep dual consecutive network for human pose estimation", Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2021.

[53] Zheng, C., Zhu, S., Mendieta, M., Yang, T., Chen, C., Ding, Z., "3d human pose estimation with spatial and temporal transformers", Proceedings of the IEEE/CVF international conference on computer vision, 2021.

[54] Li, W., Liu, H., Tang, H., Wang, P., Gool, L. V., "Mhformer: Multi-hypothesis transformer for 3d human pose estimation", Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022.

[55] Xu, Y., Zhang, J., Zhang, Q., Tao, D., "ViTPose: Simple Vision Transformer Baselines for Human Pose Estimation", 36th Conference on Neural Information Processing Systems (NeurIPS 2022), ISBN: 9781713871088.

[56] Liu, H., Liu, T., Chen, Y., Zhang, Z., Li, Y. F., "EHPE: Skeleton Cues-based Gaussian Coordinate Encoding for Efficient Human Pose Estimation", IEEE Transactions on Multimedia, 2022, Electronic ISSN: 1941-0077.

[57] Ingwersen, C., K., Mikkelstrup, C. M., Jensen, J. N., Hannemose, M. R., Dahl, A. B., "SportsPose-A Dynamic 3D sports pose dataset", Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023.

[58] Gong, J., Foo, L. G., Fan, Z., Ke, Q., Rahmani, H., Liu, J., "DiffPose: Toward More Reliable 3D Pose Estimation", Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2023, pp. 13041-13051.

[59] Wu, C., Tan, G., Li, C., "HEViTPose: High-Efficiency Vision Transformer for Human Pose Estimation", arXiv preprint arXiv:2311.13615 (2023).

[60] Zhou, Xiaolong, et al. "MDST: 2D Human Pose Estimation for SISO UWB Radar based on Micro-Doppler Signature via Cascade and Parallel Swin Transformer." IEEE Sensors Journal (2024).

[61] Cai, Guoci, et al. "End-to-End 3D Human Pose Estimation Network with Multi-Layer Feature Fusion." IEEE Access (2024).

[62] Zhang, Tong, et al. "Enhancement and optimisation of human pose estimation with multi-scale spatial attention and adversarial data augmentation." Information Fusion (2024): 102522.

[63] Andriluka, M., Pishchulin, L., Gehler, P. and Schiele, B., "2D Human Pose Estimation: New Benchmark and State of the Art Analysis", 2014. Accessed on: 17.12.2023.[Online].Available: https://paperswithcode.com/dataset/mpii

[64] Yu, X., Zhou F. and Chandraker, M., "Deep Deformation Network for Object Landmark Localization", 2016. Accessed on: 17.12.2023.[Online].Available: https://paperswithcode.com/dataset/lsp

[65] Chen, T., Fang, C., Shen, X., Zhu, Y., Chen, Z. and Luo, J., "Anatomy-aware 3D Human Pose Estimation with Bone-based Pose Decomposition", 2020. Accessed on: 17.12.2023. [Online]. Available: https://paperswithcode.com/dataset/mpi-inf-3dhp

[66] Han, F., Reily, B., Hoff, W. and Zhang, H., "Space-Time Representation of People Based on 3D Skeletal Data: A Review", 2017. Accessed on: 17.12.2023. [Online].Available: https://paperswithcode.com/dataset/human3-6m

[67] Xiao, B., Wu, H. and Wei, Y., "Simple Baselines for Human Pose Estimation and Tracking", 2018. Accessed on: 17.12.2023. [Online]. Available: https://paperswithcode.com/dataset/posetrack

[68] Lin, T., Maire, M., Belongie, S., Bourdev, L., Girshick, R., Hays, J., Perona, P., Ramanan, D., Zitnick, C. L. and Dollar, P., "Microsoft COCO: Common Objects in Context", 2014. Accessed on: 17.12.2023. [Online]. Available: https://paperswithcode.com/dataset/coco

[69] Güler, R. A., Neverova, N. and Kokkinos, I., "DensePose: Dense Human Pose Estimation In The Wild", 2018. Accessed on: 17.12.2023. [Online]. Available: https://paperswithcode.com/dataset/densepose

[70] Sapp, B. and Taskar, B., "MODEC: Multimodal Decomposable Models for Human Pose Estimation", 2013. Accessed on: 17.12.2023. [Online]. Available: https://paperswithcode.com/dataset/flic

[71] Rapczyński, Michał, et al. "A baseline for cross-database 3d human pose estimation." Sensors 21.11 (2021): 3769.

[72] Zheng, Ce, et al. "Deep learning-based human pose estimation: A survey." ACM Computing Surveys 56.1 (2023): 1-37.

[73] Jin, Sheng, et al. "Whole-body human pose estimation in the wild." Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part IX 16. Springer International Publishing, 2020.

[74] Wandt, Bastian, et al. "Canonpose: Self-supervised monocular 3d human pose estimation in the wild." Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2021.

[75] Badiola-Bengoa, Aritz, and Amaia Mendez-Zorrilla. "A systematic review of the application of camera-based human pose estimation in the field of sport and physical exercise." Sensors 21.18 (2021): 5996.

[76] Zhang, Z, "Microsoft Kinect Sensor and Its Effect", IEEE MultiMedia, Volume 19, Issue 2, 2011, pp 4–10.

[77] Lun, R., Zhao, W., "A Survey Of Applications and Human Motion Recognition with Microsoft Kinect", International Journal of Pattern Recognition and Artificial Intelligence, 29(05):150330235202000, 2015.

[78] Chow, J., Lichti, D., "Photogrammetric Bundle Adjustment With Self-Calibration of the PrimeSense 3D Camera Technology: Microsoft Kinect", IEEE Access, Volume: 1, Electronic ISSN: 2169-3536, 2013, pp. 465 – 474.

[79] ElIaithy, R.A., Huang, J., Yeh, M., "Study on the Use of Microsoft Kinect for Robotics Applications", Proceedings of the 2012 IEEE/ION Position, Location and Navigation Symposium, ISBN:978-1-4673-0387-3, 2012.

[80] Ismail, N. H. B., Basah, S., "The applications of Microsoft Kinect for human motion capture and analysis : A review", 2nd International Conference on Biomedical Engineering (ICoBE), ISBN:978-1-4799-1749-5, 2015.

[81] Chang, C. Y., Langel, B., Zhang, M., Koenig, S., Requejo, P., Somboon, N., Sawchuk, A. A., Rizzol, A. A., "Towards Pervasive Physical Rehabilitation Using Microsoft Kinect", 6th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth) and Workshops, ISBN:978-1-936968-43-5, 2012.

[82] Molyneaux, D., "KinectFusion rapid 3D reconstruction and interaction with Microsoft Kinect", International Conference on the Foundations of Digital Games, FDG '12, Raleigh, NC, USA, 2012.

[83] Bueno, M., Díaz-Vilariño, L., Martínez-Sánchez, J., González-Jorge, H., Lorenzo, H., Arias, P., "Metrological evaluation of KinectFusion and its comparison with Microsoft Kinect sensor", Measurement 73, 2015, pp. 137–145.

[84] Romeo, L., Marani, R., Malosio, M., Perri, A. G., D'Orazio, T., "Performance Analysis of Body Tracking with the Microsoft Azure Kinect", 29th Mediterranean Conference on Control and Automation (MED), ISBN:978-1-6654-4660-0, 2021.

[85] Kapgate, S., Sahu, P., Das, M., Gupta, D., "Human Following Robot using Kinect in Embedded Platform", 1st International Conference on the Paradigm Shifts in Communication, Embedded Systems, Machine Learning and Signal Processing (PCEMS), ISBN:978-1-6654-5905-1, 2022.

[86] Wagh, P., Patil, S., Shrivastav, R., Bachhav, S., "Virtual Yoga System Using Kinect Sensor. International Research Journal of Modernization in Engineering Technology and Science", Volume:04, Issue:04, April-2022.

[87] Açış, B., Güney, S., "Classification of human movements by using Kinect sensor", Biomedical Signal Processing and Control, Volume 81, March 2023, 104417.

[88] Azhar, M., Ullah, S., Raees, M., Rahman, K. U., Rehman, I. U., "A real-time multi view gait-based automatic gender classification system using kinect sensor", Multimedia Tools and Applications, 2023, 82:11993–12016.

[89] Jha, Panchanand, et al. "Human–machine interaction and implementation on the upper extremities of a humanoid robot." Discover Applied Sciences 6.4 (2024): 152.

[90] Reuther, Corentin, et al. "Engineering User Interfaces for Tailored and Monitored Movement Rehabilitation Programs." Companion Proceedings of the 16th ACM SIGCHI Symposium on Engineering Interactive Computing Systems. 2024.

[91] Bao, Q, Liu, W., Hong, J. H., Duan, L., Mei, T., "Pose-native Network Architecture Search for Multi-person Human Pose Estimation", MM '20: Proceedings of the 28th ACM International Conference on Multimedia, 2020, pp. 592–600.

[92] Rogez, G., Weinzaepfel, P., Schmid, C., "LCR-Net++: Multi-person 2D and 3D Pose Detection in Natural Images", arXiv:1803.00455, 2018.

[93] Ning, G., Pei, J., Huang, H., "LightTrack: A Generic Framework for Online Top-Down Human Pose Tracking", 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2020, ISBN:978-1-7281-9360-1.

[94] Dyce, William, et al. "Tabu search for human pose recognition." Three-Dimensional Image Processing, Measurement (3DIPM), and Applications 2014. Vol. 9013. SPIE, 2014.

[95] Oleinikov, Georgii, et al. "Task-based control of articulated human pose detection for openvl." IEEE Winter Conference on Applications of Computer Vision. IEEE, 2014.

[96] Dimitrijevic, Miodrag, Vincent Lepetit, and Pascal Fua. "Human body pose detection using Bayesian spatio-temporal templates." Computer vision and image understanding 104.2-3 (2006): 127-139.

[97] Buys, Koen, et al. "An adaptable system for RGB-D based human body detection and pose estimation." Journal of visual communication and image representation 25.1 (2014): 39-52.

[98] Cao, Zhe, et al. "Realtime multi-person 2d pose estimation using part affinity fields." Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.

[99] Jeny, Afsana Ahsan, Masum Shah Junayed, and Md Baharul Islam. "PoseTED: A Novel Regression-Based Technique for Recognizing Multiple Pose Instances." International Symposium on Visual Computing. Cham: Springer International Publishing, 2021.

[100] Narayanan, S. Sankara, et al. "Yoga pose detection using deep learning techniques." Proceedings of the International Conference on Innovative Computing & Communication (ICICC). 2021.

[101] Gadhiya, Ravindra, and Nilesh Kalani. "Analysis of deep learning based pose estimation techniques for locating landmarks on human body parts." 2021 International Conference on Circuits, Controls and Communications (CCUBE). IEEE, 2021.

[102] Grover, Anant, Deepak Arora, and Anuj Grover. "Human pose estimation using deep learning techniques." Proceedings of the 4th International Conference on Information Management & Machine Intelligence. 2022.

[103] Niu, Yue, et al. "Convpose: A modern pure convnet for human pose estimation." Neurocomputing 544 (2023): 126301.

[104] Du, Songlin, Zhiwen Zhang, and Takeshi Ikenaga. "AnatPose: Bidirectionally learning anatomy-aware heatmaps for human pose estimation." Pattern Recognition (2024): 110654.

[105] Bazarevsky, V. "BlazePose: On-device Real-time Body Pose tracking." arXiv preprint arXiv:2006.10204 (2020).

[106] Lugaresi, Camillo, et al. "Mediapipe: A framework for building perception pipelines." arXiv preprint arXiv:1906.08172 (2019).

[107] Tawar, R., Jagtap, S., Hirve, D., Gundgal, T., Kale, N., "Real-Time Yoga Pose Detection", International Research Journal of Modernization in Engineering Technology and Science, Volume:04/Issue:05/May-2022.

[108] Desmarais, Yann, et al. "A review of 3D human pose estimation algorithms for markerless motion capture." Computer Vision and Image Understanding 212 (2021): 103275.

[109] Liu, Wu, et al. "Recent advances of monocular 2d and 3d human pose estimation: A deep learning perspective." ACM Computing Surveys 55.4 (2022): 1-41.

[110] Huang, Zhangjin, and Rashid Khan. "A review of 3D human body pose estimation and mesh recovery." Digital Signal Processing 128 (2022): 103628.

[111] Chen, Yucheng, Yingli Tian, and Mingyi He. "Monocular human pose estimation: A survey of deep learning-based methods." Computer vision and image understanding 192 (2020): 102897.

[112] Yan, Jielu, et al. "Recent advances in 3D human pose estimation: From optimization to implementation and beyond." International Journal of Pattern Recognition and Artificial Intelligence 36.02 (2022): 2255003.

[113] Gupta, Prince. “Review on Human Pose Estimation using AI Fitness Tracker.” International Journal Of Scientific Research In Engineering And Management, (2024).

[114] Song, Xuzhe. "Researches Advanced in Human Pose Estimation Based on Deep Learning." Journal of Physics: Conference Series. Vol. 2404. No. 1. IOP Publishing, 2022.

[115] Knap, Pawel. "Human modelling and pose estimation overview." arXiv preprint arXiv:2406.19290 (2024).

[116] Singh, Priyanshu, et al. "Evaluating the Robustness of Human Pose Estimation Models: A Comparative Analysis." 2024 11th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO). IEEE, 2024.

[117] Sengar, Sandeep Singh, Abhishek Kumar, and Owen Singh. "Efficient Human Pose Estimation: Leveraging Advanced Techniques with MediaPipe." arXiv preprint arXiv:2406.15649 (2024).

[118] Wahid, Wasim, et al. "Advanced Human Pose Estimation and Event Classification using Context-Aware Features and XGBoost Classifier." IEEE Access (2024).

[119] Elshami, Nada E., Ahmad Salah, and Heba Mohsen. "A Comparative Study of Recent 2D Human Pose Estimation Methods." 2024 6th International Conference on Computing and Informatics (ICCI). IEEE, 2024.

[120] Chen, Wenhao. “A Review on Human Pose Estimation Based on Deep Learning” Applied and Computational Engineering 80(1):67-73, 2024.