*International Journal on Recent and Innovation Trends in Computing and Communication*
***ISSN: 2321-8169 Volume: 11 Issue: 11***
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 November 2023*

# A Statistical Examination of Pose Correction Strategies for Multidomain Applications

**Manas Ramteke [1] Dr. Vijayalaxmi Biradar[2]**
[1] Reasearch Scholar, Kalinga University, Naya Raipur,Chhattisgarh,India
[2] Associate Professor, Kalinga University, Naya Raipur,Chhattisgarh,India

***Abstract*:** Pose estimation and correction is a multidomain problem that includes identifying body key points, tracking these key points via multimodal analysis, making continuous recommendations, and improving poses. This review article presents an overview of recent improvements in posture correction estimating algorithms for human beings. Pose estimate is an important problem in many applications, including fitness tracking, motion analysis, and virtual reality scenarios. The study explores cutting-edge methods for estimating human poses, including deep learning-based approaches, multi-camera systems, and bioinspired models. Furthermore, the study discusses numerous applications of human posture estimation, such as gait analysis, rehabilitation, and sports analysis. The problems of human posture estimation, such as occlusions, limited training data, and differences in body form and size, are also examined for various circumstances.

***Keywords:*** *Human, Pose, Recommendation, Classification, Correction, Models, Accuracy, Delay, Complexity, Scenarios*

## 1. Introduction

The purpose of this review paper is to provide an overview of the recent developments that have been made in pose correction estimation methods, with a particular emphasis on human subjects. In this paper, we discuss the state-of-the-art methods that are used for human pose estimation and correction. These methods include traditional computer vision approaches as well as methods that are based on deep learning. In addition to that, it discusses several applications of human pose estimation and correction, such as gait analysis, rehabilitation, and sports analysis. In addition to this, the paper draws attention to the difficulties associated with human pose estimation and correction, as well as identifies possible research directions and applications for subsequent work. This review paper's objective is to provide a comprehensive overview of the current state of research in pose correction estimation for human subjects. More specifically, the paper will focus on highlighting the most important contributions to the field, as well as limitations and potential future directions for a field that is both exciting and rapidly evolving for different scenarios.

## 2. Literature Review of existing pose correction models

A wide variety of deep learning image processing models are proposed by researchers for estimation & correction of human poses. For instance, in [1], attempts were made to tackle the problem of determining a 3D location from a 2D image. Numerous innovative posture evaluation strategies have produced "quantitatively" beneficial outcomes. (Below 50 millimetres). However, they continue to be "perceptually" defective due to the fact that their effectiveness is measured solely by physical distance. A reliance on "quantitative" data may have delayed progress in 3D position assessment methods despite widespread awareness of this fact. Assuming that human perception (HP) is capable of extracting structural information from a signal, researchers propose a perceptual Pose SIMilarity (PSIM) measure to address this problem. Second, researchers present a model for perceptually accurate 3D pose prediction using networks with Temporal Propagating Long Short-Term Memory. (TP-LSTMs). Here, the interdependence, temporal coherence, and HP of spatio-temporal posture relationships are analyzed using information theory. The experimental results demonstrate that the proposed PSIM gauge correlates substantially better with users' intuitive assessments of their own posture than more conventional metrics. Quantitatively and qualitatively, researchers demonstrate that TP-LSTMs provide significantly superior performance to the current state-of-the-art methods.

Despite computer vision's great success in extracting meaning from images, [2] research explains how its algorithms are vulnerable in low-resolution settings or when working with limited data and label pairs. Monitoring a patient's position while they are lying in bed is one such task, which has numerous medical applications. Natural in-bed position monitoring necessitates body orientation determination even when all light is occluded. The absence of publicly available in-bed stance datasets prevents the application of many efficient human posture prediction algorithms to this task. In

---

this paper, researchers present their concurrently-collected multimodal Lying Pose (SLP) dataset, which consists of images of 109 individuals in bed positions acquired concurrently using various imaging modalities, such as RGB, long wave infrared (LWIR), depth, and pressure map. researchers also present a physical hyper parameter optimization strategy for generating ground truth pose labels under challenging visual conditions. This enables the training of state-of-the-art 2D posture estimation models with the SLP data, with optimistic results as high as 95% at PCKh@0.5 on a single scenario. The SLP structure is consistent with prominent databases of human posture. By incorporating novel modes using the proposed joint approaches, the accuracy of these models for predicting stance can be enhanced for different scenarios. Analysis of human behavior in still images and moving video necessitates human identification and stance analysis, according to studies cited in [3].

The majority of methods for estimating poses in groups of people begin with human identification and then input the resultant bounding boxes into a network for pose estimation. This top-down approach is hampered by the early commitment of initial detections, which leads to posture estimation failures in congested environments and other scenarios with uncertainties or occlusions. For multiple individual identification and posture evaluation, researchers propose the DetPoseNet (DPN), a three-stage network with an end-to-end architecture. A multi-scale pose enhancement sub-net, an extraction sub-net for coarse-scale pose proposals, and a filtering module based on these coarse-scale pose proposals constitute their method. In a single iteration, the coarse-pose proposal sub-net extracts the entire-body bounding frames and body key point proposals. Filtering on the basis of proposed individuals and key points eliminates false positives and accelerates processing. On each updated hypothesis space, the subnetwork responsible for refining poses employs a cascaded posture estimation technique. The objective of the posture modification sub-net's multi-scale guiding and multi-scale regression is to simultaneously enhance context feature learning. Using structure-aware loss and keypoint filtration, researchers can further improve stance enhancement's robustness. Due to the adaptability of their framework, most existing top-down pose estimators can be utilized as the subnet for pose modification in their method. Using the COCO and OCHuman datasets, experimental results demonstrate the efficacy of the proposed structure. The proposed method predicts multi-person poses with accurate boundary frames in less than a second with 5-6 times faster computing efficiency levels. According to research presented in [4], the importance of human position monitoring in the field of computer vision is growing. (CV). Many people are searching for alternatives to webcams to enhance the privacy of posture monitoring. Radio-frequency identification (RFID) bracelets are a low-cost, portable device that can be used to monitor the three-dimensional position of humans. RFID-stance is a CV-supported, deep learning-based method for predicting the three-dimensional posture of a person in real time. After measuring the RFID phase data, the high precision low rank tensor completion technique is used to complete the missing information and reduce the severe phase deformation. Using the estimated spatial rotation angles of individual human appendages, real-time forward kinematic reconstruction of human position is performed. The sample is created with standard RFID scanners. In their experiments with RFID-Pose using Kinect 2.0 as a set of standards, researchers discovered a high degree of pose prediction accuracy and real-time operations.

According to research published in [5], the field of human-computer interaction has always emphasized three-dimensional human position assessment or human monitoring. The following method for predicting human poses relies significantly on subject-specific modeling to validate human poses. It offers both historical context and a well-defined endpoint for monitoring scenarios. This article describes an entirely automatic subject modeling system for recreating human posture, body shape, and body material in challenging optimization scenarios. The proposed method integrates robust differentiable rendering into the domain-specific modeling procedure, thereby transforming the texture reconstruction problem into an analysis-by-synthesis reduction problem that can be efficiently solved by a gradient-based strategies. An efficient Novel Adaptive Covariance Matrix Annealing (ACMA) technique is proposed to further address the high-dimensional multimodal optimization issue. Adding domain expertise in the form of the human social order to the annealing optimization procedure is as simple as employing a correlation matrix-based filter. This combination of characteristics renders the novel algorithm immune to the temptation of local minima. Experiments conducted on the Human3.6 M dataset and the People-Snapshot dataset yield qualitatively and quantitatively comparable results to the current state of the art for a variety of use cases.

As discussed in [6], new computer vision and automated applications have increased the importance and interest in recovering 3D human poses. This endeavour is extremely difficult because stereoscopic images contain a wide variety of perspectives, views, occlusions, and basic geometry defects. To directly extract 3D human poses from corresponding 2D human pose-aware features or 2D pose predictions, the majority of current methods emphasize the construction of intricate priors or limits. Due to the paucity of 3D stance data for training and the fact that 2D space and 3D space are inherently distinct regions, the scalability of these methods is limited in all practical contexts. (e.g., outdoor scene). In an effort to address the issue, this paper presents a straightforward yet effective self-supervised correction technique for learning all intrinsic structures of human poses from a large number of images. The proposed technique incorporates two dual learning tasks, namely the 2D-to-3D pose transformation and the 3D-to-2D pose projection, in order to reconcile the gap between 3D and 2D human poses during a type of "free" self-supervision for accurate 3D human pose evaluation. The 3D-to-2D pose projection refines the intermediate 3D poses by maintaining geometric consistency between the 2D projections of 3D poses and the estimated 2D poses. Due to these two dual learning objectives (DLO), their model can adaptably learn from both large-scale external 3D human pose data and external 2D human pose data. Using their self-supervised correction method, researchers construct a 3D human pose machine that incorporates 2D spatial connections, temporal regularity of predictions, and 3D geometric information. Extensive comparisons to alternative methods using the Human3.6M and HumanEva-I benchmarks demonstrate that their technology is faster and less expensive.

The authors of [7] describe a method (Hybrid-stance) to enhance image-based human posture prediction. Stacked Hourglass Networks are utilized to create two variants of convolutional neural networks: RNet for enhancing stance and CNet for rectifying poses. Before generating the final stance, the CNet (Correction Network) instructs the pose revision RNet (revision Network) to modify the joint position. Each model in the combination consists of four hourglasses, with the output of each model being a set of joint recognition heatmaps. The RNet topology employs a circular standard form. However, the CNet type utilizes various forms of hourglasses to maintain equilibrium. Due to the environmental sensitivity of posture estimation in RGB images, their proposed method generates numerous detection heatmap outputs to expand the search space for the correct joint locations.

The RNet model is used to fine-tune the connections between each hourglass stage, and the heatmaps from each stage are then combined along the vertical axis with the heatmaps from each hourglass in the CNet model. On the gold-standard datasets for MPII and FLIC, their method achieves results that are comparable to the current state-of-the-art methodologies. researchers demonstrate how this improvement can be applied to multi-person pose estimation by first identifying a large number of people with SSD detection and then independently estimating the poses of each person in a variety of scenarios.

## 3. Statistical Analysis of reviewed techniques

Based on this evaluation, it can be observed that a wide variety of functionally variant models are proposed by researchers, and each of them vary in terms of their real-time performance for different application scenarios. A statistical analysis of these models is discussed in this section of the text, where each of these models are compared in terms of Accuracy (A), Precision (P), Computational Delay (D), Cost (C), and Scalability (S) measures. Values of accuracy & precision are directly referred from the text, while delay, complexity & scalability were converted into fuzzy ranges of Low (L=1), Medium (M=2), High (H=3), and Very High (VH=4), which will assist readers to compare these models on similarly quantized scales. Based on this evaluation strategy, the performance of these models is tabulated in table 1 as follows,

| **Model** | **A** | **C** | **D** | **P** | **S** |
|---|---|---|---|---|---|
| PSIM LSTM [1] | 91.5 | H | H | 87.64 | VH |
| SLP [2] | 85.4 | M | H | 82.39 | H |
| DPN [3] | 90.4 | H | H | 80.10 | H |
| RFID [4] | 75.5 | VH | M | 76.52 | L |
| ACMA [5] | 78.4 | H | H | 82.43 | H |
| DLO [6] | 79.5 | H | H | 87.70 | H |
| RNet CNet [7] | 93.5 | M | M | 91.61 | VH |
| DGNet [8] | 94.5 | M | H | 88.92 | H |
| CDF [9] | 91.4 | H | VH | 89.25 | H |
| RPS TN [10] | 85.3 | H | H | 86.66 | H |
| PPNet [11] | 95.5 | H | H | 89.61 | H |

*International Journal on Recent and Innovation Trends in Computing and Communication*
***ISSN: 2321-8169 Volume: 11 Issue: 11***
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 November 2023*

| | | | | | |
|---|---|---|---|---|---|
| Wi Mose [12] | 83.5 | VH | H | 89.28 | L |
| DCNN [13] | 94.3 | M | L | 92.20 | H |
| BRNN [14] | 94.5 | L | L | 90.95 | VH |
| PQPB [15] | 92.4 | H | L | 87.34 | H |
| ARH PE [16] | 90.5 | L | M | 84.98 | H |
| PLAM [17] | 83.5 | H | H | 84.33 | H |
| MeTRo [18] | 85.2 | VH | H | 86.59 | H |
| MTM [19] | 88.5 | VH | H | 88.66 | H |
| CNN [20] | 90.4 | H | M | 89.28 | M |
| LSTM [21] | 91.5 | H | H | 86.98 | H |
| HE Mlets PoSH [22] | 90.4 | VH | H | 85.02 | H |
| MRDN [23] | 83.4 | H | H | 81.44 | H |
| Pose Net & Depth Net [24] | 85.5 | VH | L | 85.41 | VH |
| BAO DP [25] | 79.5 | L | H | 88.00 | H |
| GRU [26] | 95.5 | H | H | 92.36 | VH |
| VRNet [27] | 93.4 | H | H | 91.05 | H |
| GAL [28] | 92.8 | M | H | 87.77 | H |
| GPA RF [29] | 91.5 | L | M | 82.43 | H |
| Uni Pose [30] | 83.4 | H | H | 82.43 | H |
| UWB [31] | 76.5 | H | H | 85.51 | H |
| LG Pose FCN [32] | 91.5 | H | H | 88.46 | H |
| Pose GCN [33] | 92.8 | H | VH | 85.84 | H |
| KM [34] | 85.5 | H | H | 82.75 | H |
| 6DoF [35] | 83.5 | H | H | 84.39 | H |
| PRCM [36] | 83.4 | H | H | 86.92 | H |
| PAM [37] | 90.5 | VH | H | 87.74 | H |
| VIHC [38] | 91.2 | M | H | 85.31 | VH |
| CAD [39] | 85.9 | H | H | 84.98 | H |
| RTK [40] | 83.1 | H | H | 86.16 | H |
| SLAM [41] | 90.2 | L | H | 86.92 | M |
| PF [42] | 89.5 | M | VH | 87.44 | H |
| CSPC [43] | 85.4 | H | H | 87.77 | H |
| SLAM BoW [44] | 91.8 | L | H | 91.08 | H |
| DCO GIS LAM [45] | 90.5 | H | H | 93.48 | M |
| I2Net [46] | 95.5 | H | VH | 94.43 | H |
| PC [47] | 99.1 | H | H | 94.70 | H |
| AAL [48] | 93.4 | H | L | 86.95 | VH |
| KF [50] | 83.9 | VH | H | 84.64 | H |

Table 1. Comparative analysis of the Reviewed Models

*International Journal on Recent and Innovation Trends in Computing and Communication*
***ISSN: 2321-8169 Volume: 11 Issue: 11***
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 November 2023*

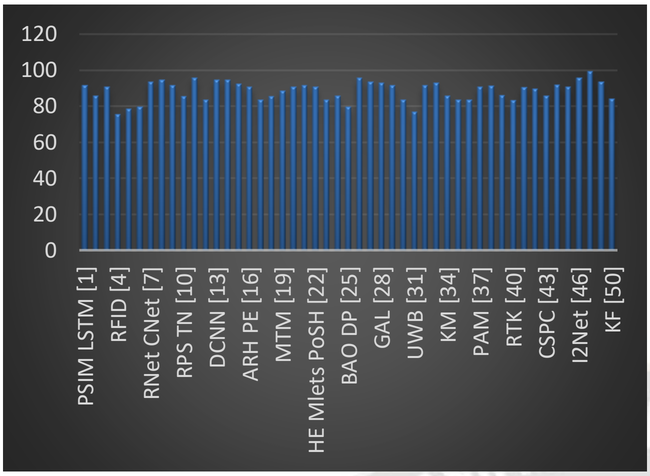


Figure 1. Accuracy of pose correction for different models

Based on this comparison, it can be observed that PC [47], PPNet [11], GRU [26], I2Net [46], DGNet [8], BRNN [14], and DCNN [13] showcase higher accuracy, thus can be used for high-efficiency pose correction applications.

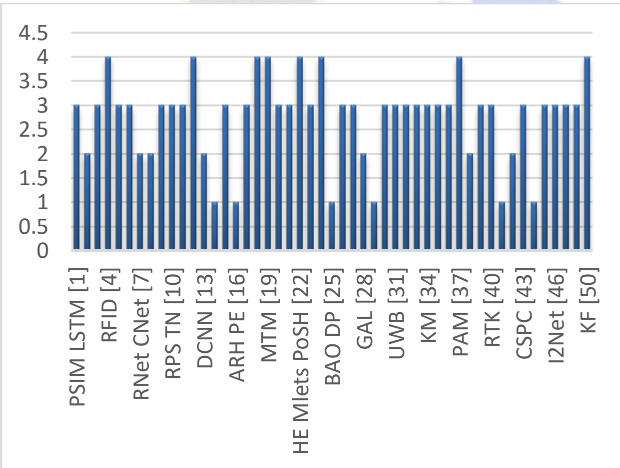


Figure 2. Precision of pose correction for different models

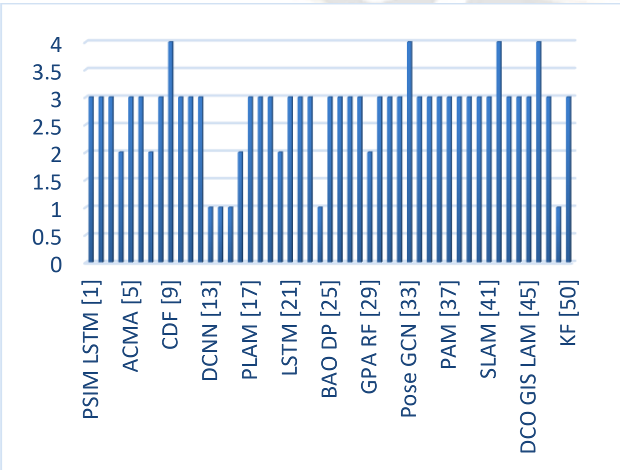


Figure 3. Delay of pose correction for different models

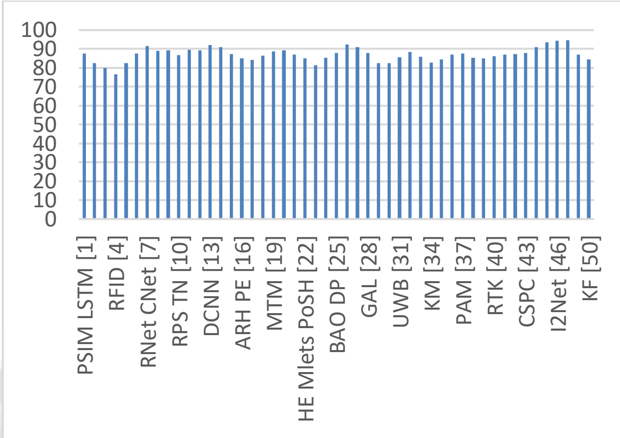


Figure 4. Precision of pose correction for different models

Work in PC [47], I2Net [46], DCO GIS LAM [45], GRU [26], DCNN [13], RNet CNet [7], SLAM BoW [44], and VRNet [27] is able to achieve better precision, thus can be used when highly consistent performance is needed during pose correction deployments.

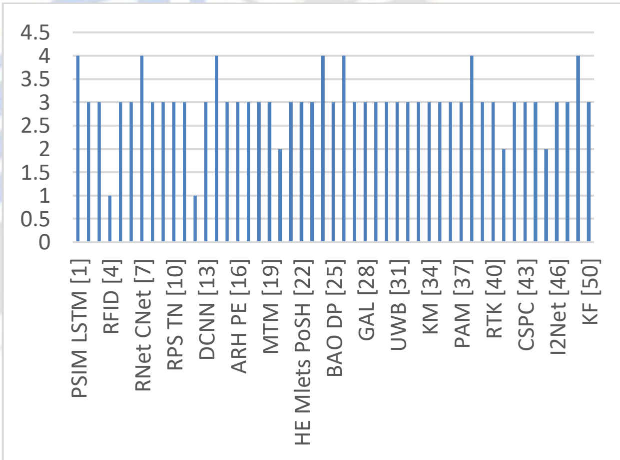


Figure 5. Scalability of pose correction for different models

Work in BRNN [14], ARH PE [16], BAO DP [25], GPA RF [29], SLAM [41], and SLAM BoW [44] is capable of outperforming in the field of pose correction in terms of cost, thus can be used for cost-aware deployments. Work in DCNN [13], BRNN [14], PQPB [15], Pose Net & Depth Net [24], and AAL [48] showcased lower delay, thus can be used for high-speed application scenarios.

Work proposed in PSIM LSTM [1], RNet CNet [7], BRNN [14], Pose Net & Depth Net [24], GRU [26], VIHC [38], and AAL [48] is able to improve scalability levels, thus can be used for multiple application deployments. A fusion metric from these parameters is calculated via equation 1,

$$PCRM = \frac{A+P}{200} + \frac{1}{C} + \frac{1}{D} + \frac{S}{4} \ldots (1)$$

Where, $PCRM$ is the fused Pose Correction Rank Metric, which is evaluated for individual models. Based on this evaluation, and its visualization in figure 6, it can be observed that BRNN [14], DCNN [13], PC [47], I2Net [46], SLAM BoW [44], DCO GIS LAM [45], RNet CNet [7], GRU [26], VRNet [27], BAO DP [25], PQPB [15], and CNN [20] can be used for high accuracy, low complexity, low delay, high precision, and high scalability scenarios.

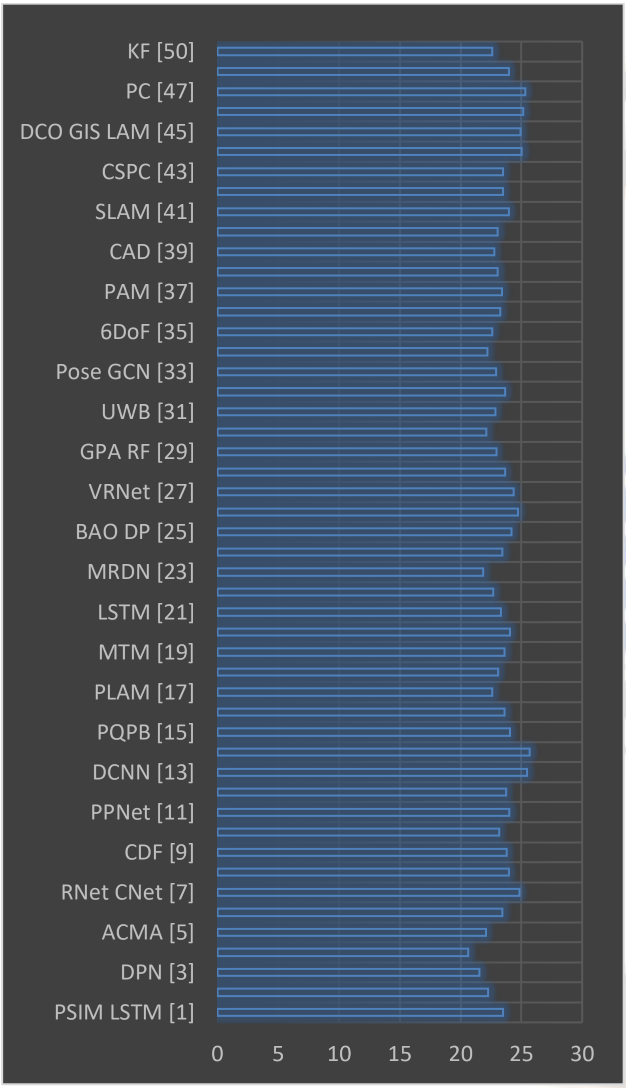


Figure 6. PCRM of pose correction for different models

Thus, researchers can use these models to enhance performance of human pose correction under real-time scenarios.

## 4. Conclusion and future scope

Pose estimation is critical for a wide range of applications, including fitness tracking, motion analysis, and virtual reality situations. The study explored the most modern strategies for predicting human postures, such as deep learning-based algorithms, multi-camera systems, and bio-inspired models. The research also covered a variety of human posture estimation applications, including gait analysis, rehabilitation, and sports analysis. The problems in judging human postures are also highlighted in relation to different events. These challenges include occlusions, a lack of training data, and differences in body size and form. Work in BRNN, ARH PE, BAO DP, GPA RF, and SLAM. BoW has the potential to beat existing posture correction systems in terms of cost, making it suitable for low-cost deployments. Work in DCNN, BRNN, PQPB, Pose Net, Depth Net, and AAL revealed lower latency, indicating that they may be implemented in high-speed application settings.

## References

[1] K. Lee, W. Kim and S. Lee, "From Human Pose Similarity Metric to 3D Human Pose Estimator: Temporal Propagating LSTM Networks," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 45, no. 2, pp. 1781-1797, 1 Feb. 2023, doi: 10.1109/TPAMI.2022.3164344.

[2] S. Liu, X. Huang, N. Fu, C. Li, Z. Su and S. Ostadabbas, "Simultaneously-Collected Multimodal Lying Pose Dataset: Enabling In-Bed Human Pose Monitoring," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 45, no. 1, pp. 1106-1118, 1 Jan. 2023, doi: 10.1109/TPAMI.2022.3155712.

[3] L. Ke, M. -C. Chang, H. Qi and S. Lyu, "DetPoseNet: Improving Multi-Person Pose Estimation via Coarse-Pose Filtering," in IEEE Transactions on Image Processing, vol. 31, pp. 2782-2795, 2022, doi: 10.1109/TIP.2022.3161081.

[4] C. Yang, X. Wang and S. Mao, "RFID-Pose: Vision-Aided Three-Dimensional Human Pose Estimation With Radio-Frequency Identification," in IEEE Transactions on Reliability, vol. 70, no. 3, pp. 1218-1231, Sept. 2021, doi: 10.1109/TR.2020.3030952.

[5] Y. Lu, G. Chen, C. Pang, H. Zhang and B. Zhang, "Subject-Specific Human Modeling for Human Pose Estimation," in IEEE Transactions on Human-Machine Systems, vol. 53, no. 1, pp. 54-64, Feb. 2023, doi: 10.1109/THMS.2022.3195952.

[6] K. Wang, L. Lin, C. Jiang, C. Qian and P. Wei, "3D Human Pose Machines with Self-Supervised Learning," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 42, no. 5, pp. 1069-1082, 1 May 2020, doi: 10.1109/TPAMI.2019.2892452.

[7] A. Kamel, B. Sheng, P. Li, J. Kim and D. D. Feng, "Hybrid Refinement-Correction Heatmaps for Human Pose Estimation," in IEEE Transactions on Multimedia,

vol. 23, pp. 1330-1342, 2021, doi: 10.1109/TMM.2020.2999181.

[8] J. Zhang, Y. Wang, Z. Zhou, T. Luan, Z. Wang and Y. Qiao, "Learning Dynamical Human-Joint Affinity for 3D Pose Estimation in Videos," in IEEE Transactions on Image Processing, vol. 30, pp. 7914-7925, 2021, doi: 10.1109/TIP.2021.3109517.

[9] Y. Zhao, Z. Yuan and B. Chen, "Accurate Pedestrian Detection by Human Pose Regression," in IEEE Transactions on Image Processing, vol. 29, pp. 1591-1605, 2020, doi: 10.1109/TIP.2019.2942686.

[10] Y. Dang, J. Yin and S. Zhang, "Relation-Based Associative Joint Location for Human Pose Estimation in Videos," in IEEE Transactions on Image Processing, vol. 31, pp. 3973-3986, 2022, doi: 10.1109/TIP.2022.3177959.

[11] L. Zhao, N. Wang, C. Gong, J. Yang and X. Gao, "Estimating Human Pose Efficiently by Parallel Pyramid Networks," in IEEE Transactions on Image Processing, vol. 30, pp. 6785-6800, 2021, doi: 10.1109/TIP.2021.3097836.

[12] Y. Wang, L. Guo, Z. Lu, X. Wen, S. Zhou and W. Meng, "From Point to Space: 3D Moving Human Pose Estimation Using Commodity WiFi," in IEEE Communications Letters, vol. 25, no. 7, pp. 2235-2239, July 2021, doi: 10.1109/LCOMM.2021.3073271.

[13] S. Feng, X. Li and H. Hu, "Combining Parsing Information With Joint Structure for Human Pose Estimation," in IEEE Access, vol. 8, pp. 123408-123418, 2020, doi: 10.1109/ACCESS.2020.3004937.

[14] Y. Xu, W. Wang, T. Liu, X. Liu, J. Xie and S. -C. Zhu, "Monocular 3D Pose Estimation via Pose Grammar and Data Augmentation," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 44, no. 10, pp. 6327-6344, 1 Oct. 2022, doi: 10.1109/TPAMI.2021.3087695.

[15] L. Zhao, J. Xu, C. Gong, J. Yang, W. Zuo and X. Gao, "Learning to Acquire the Quality of Human Pose Estimation," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 31, no. 4, pp. 1555-1568, April 2021, doi: 10.1109/TCSVT.2020.3005522.

[16] H. Liu, T. Liu, Z. Zhang, A. K. Sangaiah, B. Yang and Y. Li, "ARHPE: Asymmetric Relation-Aware Representation Learning for Head Pose Estimation in Industrial Human–Computer Interaction," in IEEE Transactions on Industrial Informatics, vol. 18, no. 10, pp. 7107-7117, Oct. 2022, doi: 10.1109/TII.2022.3143605.

[17] Y. Xue, J. Chen, X. Gu, H. Ma and H. Ma, "Boosting Monocular 3D Human Pose Estimation With Part Aware Attention," in IEEE Transactions on Image Processing, vol. 31, pp. 4278-4291, 2022, doi: 10.1109/TIP.2022.3182269.

[18] I. Sárándi, T. Linder, K. O. Arras and B. Leibe, "MeTRAbs: Metric-Scale Truncation-Robust Heatmaps for Absolute 3D Human Pose Estimation," in IEEE Transactions on Biometrics, Behavior, and Identity Science, vol. 3, no. 1, pp. 16-30, Jan. 2021, doi: 10.1109/TBIOM.2020.3037257.

[19] D. C. Luvizon, D. Picard and H. Tabia, "Multi-Task Deep Learning for Real-Time 3D Human Pose Estimation and Action Recognition," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 43, no. 8, pp. 2752-2764, 1 Aug. 2021, doi: 10.1109/TPAMI.2020.2976014.

[20] C. Xie, D. Zhang, Y. Hu and Y. Chen, "Hierarchical Dynamic Programming Module for Human Pose Refinement," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 32, no. 9, pp. 6214-6226, Sept. 2022, doi: 10.1109/TCSVT.2022.3164663.