Automated Body Postures Assessment From Still Images Using Mediapipe

Mazin H. Aziz [1], Hamed A. Mahmood [2]

[1] Computer Engineering Department, University of Mosul, Mosul
ORCID No: https://orcid.org/0000-0003-0124-7841
[2] Computer Engineering Department, University of Mosul, Mosul
ORCID No: https://orcid.org/0000-0001-6143-330X

Keywords
body postures, pose, landmarks, Mediapipe, Machine Learning.

Abstract
Human poses assessment was an exciting research trend in the last decades. It was used in sports, health care, and many other fields, to help people get better performance. Machine learning and artificial intelligence techniques are used for this purpose. This paper used Google Mediapipe as a part of a framework for automatic Human-body pose assessment in real time. The proposed framework is based on detecting reference image poses, finding pose landmarks, and extracting discriminative features for each pose. These same process stages are applied to each image frame taken for the trainee using a web camera. The last stage of the framework compares the extracted features for the learner pose image with the saved features of the reference. The reference image was proposed to enable the system to be used for various applications. This system acts like a smart mirror that detects differences between the user pose and the reference still image then gives correction information in real time. Experiments were performed on side view cases like standing and sitting activities and gave promising results. This system could be very helpful for automatically self-pose assessment at home, or as an auxiliary tool for a certain learning program.

Research Article
Submission Date : 10.04.2023
Accepted Date : 12.06.2023

1. INTRODUCTION

Human body posture is a very interesting research line in the last two decades at least. It is important in our daily life activities (Stenum et al., 2021). There are correct postures for almost every state of our bodies. Incorrect postures could be very harmful and should be detected and avoided. Automated pose assessment attracts a lot of attention using different technical methods like using wearable and unwearable sensors (Diraco, Leone, & Siciliano, 2013; Lawanont, Inoue, Mongkolnam, & Nukoolkit, 2018; Leightley & Yap, 2018), computer vision, and machine learning (Munea et al., 2020; Nicolae-Adrian, Claudiu, Ana-Maria, & Ciprian, 2021). Google's MediaPipe Pose is a part of a framework which offers a pre-trained machine learning model that detects body joints in an image and video. It is based on the convolutional neural network BlazePose. From an RGB image, MediaPipe predicts 33 body important locations termed posture landmarks (each of which consists of x, y, and z referring to the hip-point) (Bazarevsky et al., 2020).

Resp author; e-mail: mazin.haziz@uomosul.edu.iq

In this paper, we present a method for real time body pose assessment based on comparing the features of the trainee or learner video-frames (frame by frame) with the features of a reference standstill image and give proper text advice to the user to reach the best match with the reference pose in real time. This method could be application-independent and be applied to various activities. A side view posture was taken as an example for standing and sitting states. The reference image means the perfect posture that the trainee wants to reach and could be given by the trainer or selected from a confident reference.

2. PREVIOUS WORKS

(Jesmeen et al., 2023) have developed a real-time CNN-based system to recognize three sleep postures (right, left, and lying on back). They collected their own dataset to train the CNN, and provided 99.59% classification accuracy. (Chiang et al., 2022) proposed a posture recognition system for bedridden elderly patients using an RGB-Depth camera, where only four-3D joints were adopted. The system used OpenPose for 2D skeleton-joint detection. The recognition accuracy was about 95%. If the elderly person doesn't change his posture for a pre-given period, a warning message is sent to the medical staff. MediaPipe was used by (Kwon & Kim, 2022) to get the body posture landmarks for an exerciser doing (squat, push-up) exercise at home, where Open CV was adopted for video capturing via a webcam, posture estimation, body angle calculation, incorrect pose detection, and giving text messages for pose correction. An adaptive generative adversarial network (GAN) approach was produced by (Xu et al., 2022) for landmarks detection. DenseNet and KNN were combined by (Dittakavi et al., 2022) to produce a Pose Tutor for pose detection and correction. Human body posture detection and assessment can have a variety of applications. One of them is Yoga posture assessment (George, Dcouth, Jaimy, Daji, & Antony, 2022) (Long, Jo, & Nam, 2022) (Wu et al., 2022) (Kothari, 2020). Another application is baby sleep monitoring (Khan, 2021). Head pose was analyzed by (Hammadi, Grondin, Ferland, & Lebel, 2022), while (Johnston et al., 2022) focused on Human motion. Body balance (Nguyen, Woo, Huynh, & Jeong, 2022), and fall detection(Han, Yang, & Huang, 2020) were investigated too. Sitting pose (Chen, 2019) and Workout assessment (Dawange, Chavan, & Dusane, 2021) (Kwon & Kim, 2022) were studied also. Body language attracted (Piñero-Fuentes et al., 2021), while Islamic Prayer activity monitoring was implemented by (Gupta, 2020) (Koubaa et al., 2020). A Pose Tutor (Dittakavi et al., 2022) was suggested too. It is well noted that each one of the previous works was tailored to a certain use, here we introduce a method that could be the first step to cross over many applications and could be applicable for various usages.

3. MATERIALS AND METHODS

We proposed a way that helps people to make self-assessment for their body posture using a selected reference pose image. It can be used for various types of postures and not limited to a certain use. The proposed setup consists of a webcam connected to a PC, as depicted in **Error! Reference source not found.**, that executes the proposed algorithm using Python program. The user needs to feed the PC with the reference pose image then the system will acquire the user pose-images via the webcam, and he will get the assessment of his posture and some advice for improvement for each frame of the video. The assessment and the correction advice are given as text messages along with some sketches added on the images.

Figure 1. The proposed setup

The suggested framework consists of two identical processing pipelines as shown in **Error! Reference source not found.**; one for the reference image and the other for the user video. The common last stage compares the extracted features from both lines and give the proper assessment in real time.

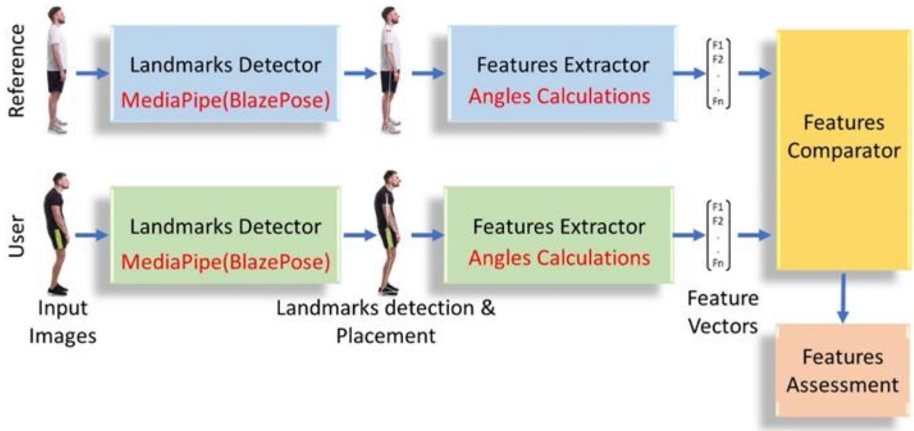


Figure 1. The suggested framework

3.1. Landmarks Detection

The input images are subjected to MediaPipe for landmarks detection. Google MediaPipe uses PlazePose to find 33 body landmarks with high accuracy. These landmarks can be projected on the original image. Each landmark is defined by three coordinates x, y and z and. The landmarks can also be connected with line-segments to form a skeleton shape for the original image as shown in **Error! Reference source not found.** which reflects the body posture in that image.

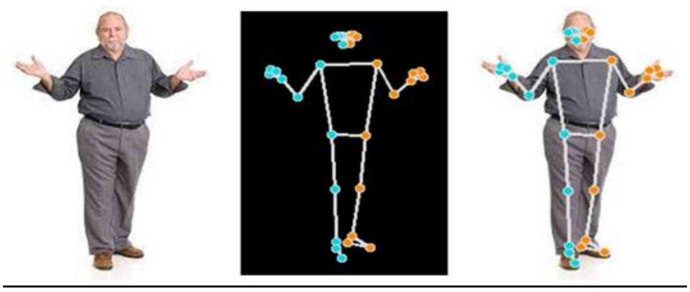


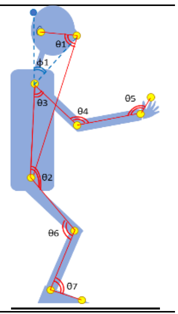
Figure 2. Image skeleton extracted using Mediapipe key points

3.2. Features Extraction

Seven handcrafted features were extracted from the detected landmarks for the right side of the body. These features were empirically chosen to produce distinguishable differences between

Table 1. Postures' features for the right side of the body

	Features	2-Vertix- Angle	3-Vertix-Angle
θ1	Head Tilt	Shoulder-Nose (φ1)	Hip-Nose-Ear
θ2	Torso Alignment	-	Shoulder-Hip-Knee
θ3	Arm Angle	-	Elbow- Shoulder-Hip
θ4	Elbow Angle	-	Shoulder-Elbow-Wrist
θ5	Wrist Angle	-	Elbow-Wrist-Index Finger
θ6	Knee Angle	-	Ankle-Knee-Hip
θ7	Foot angle	-	Knee-Ankle-Foot Index



the reference and the user poses. We adopted angle calculations between two vectors of the skeleton shape as presented in Table 1. The angles can be calculated from three vertices (landmarks) or from two landmarks using a third vertex lying on one of the cartesian axes of the image. Angle calculation from three vertices is not affected by image orientation since it is only related to the predicted landmarks of the body. The angle-dependent features may describe body posture with acceptable accuracy for many cases, and they aren't affected by body parts' length, width, or shape. We found these features are suitable to detect posture misleading from side view images for standing and sitting.

3.3. Features Comparison and Decision Making

The pre-last stage of the framework is the comparator, which compares the extracted features from each frame of the user video with those of the reference image. The comparison is done by subtracting each feature of the reference from its corresponding feature from the user. The result gave an indication of the deviation of each body part of the user body from its equivalent of the reference and direction of that deviation. The smaller the deviation the better performing. The last stage is the pose assessment which was accomplished by overlaying text and line sketches on the user video-frames. The sketches indicated the angle vectors for each feature in addition to line-arrows pointing to the direction of the required corrections. While the text messages were of two types, the difference angles that should be minimized in addition to an advice sentence as depicted in **Error! Reference source not found.**.

Figure 3. Sample of the trainee image overlayered with the pose assessment notice

4. RESULTS AND DISCUSSIONS

The proposed method was applied on side view posture for two main body situations as shown in **Error! Reference source not found.**. The first application is used to correct body posture and help people to avoid harmful standing postures. The second application is the sitting state which takes a lot of time of our lives and could lead to many unwanted consequences on our bodies. The results that were given by the presented framework were compared with the Man-made decisions and gave comparable results with some notes. The first note is that the torso and neck bowing cannot be detected precisely by using landmarks. The second note is the need to align the camera coordinates with the reference image coordinates to extract more useful features. The third note is that the selection of the reference image is crucial.

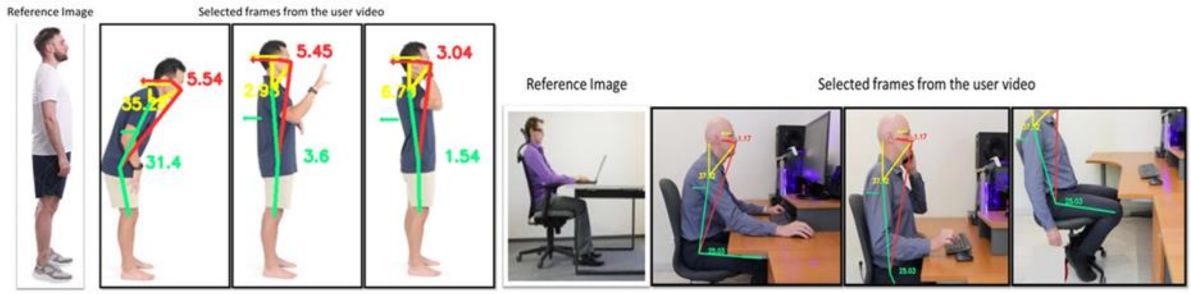


Figure 4. An example for standing and sitting postures for three selected frames of the user video

Body rotation was not considered in this research, and it is left to future works. Anyway, the system works well regardless of image and video backgrounds, image size, age, or gender difference between the user and the reference as demonstrated in **Error! Reference source not found.**.

Figure 5. Age, gender, background, and image size differences has no effects on the operation

5. CONCLUSION

In this paper, we presented a human body posture assessment method using the convolutional neural network model BlazePose via MediaPipe in the detection of 33 body landmarks. These landmarks were used to extract discriminative handcrafted features. The assessment method was based on comparing the features of a reference image with each frame of the user video in real-time using a webcam. This method is applicable for many usages and doesn't depend on the opinion of the designer, it almost looks like the traditional training ways. A side view was experimented with the standing and sitting postures and showed comparable results. This method can be developed to be general by making the system more intelligent by detecting the body side and selecting the most suitable features, from pre-saved features, depending on the reference image body state.

Conflict of Interest

The authors declare that they have no conflicts of interest related to this research. The study was conducted without any external financial support, and there are no personal relationships or financial interests that could influence the research findings or interpretation.

Author Contributions

Mazin H. Aziz: Conceptualization, Methodology, Writing, Visualization, and Editing.

Hamed A. Mahmood: Methodology and Programming.

REFERENCES

Bazarevsky, V., Grishchenko, I., Raveendran, K., Zhu, T., Zhang, F., & Grundmann, M. (2020). *BlazePose: On device Real-time Body Pose tracking*. Retrieved from doi: https://doi.org/10.48550/arXiv.2006.10204

Chen, K. (2019). Sitting Posture Recognition Based on OpenPose. *IOP Conference Series: Materials Science and Engineering*, *677*(3). doi:https://doi.org/10.1088/1757-899X/677/3/032057

Chiang, J.-C., Lie, W.-N., Huang, H.-C., Chen, K.-T., Liang, J.-Y., Lo, Y.-C., & Huang, W.-H. (2022). Posture Monitoring for Health Care of Bedridden Elderly Patients Using 3D Human Skeleton Analysis via Machine Learning Approach. *Applied Sciences*, *12*(6), 3087. doi: https://doi.org/10.3390/app12063087

Dawange, S., Chavan, A., & Dusane, A. (2021). Workout Analysis Using Mediapipe BlazePose and Machine Learning. *International Journal of Creative Research Thoughts (IJCRT)*, *9*(12), 294–297.

Diraco, G., Leone, A., & Siciliano, P. (2013). Human posture recognition with a time-of-flight 3D sensor for in-home applications. *Expert Systems with Applications*, *40*(2), 744–751. doi: https://doi.org/10.1016/j.eswa.2012.08.007

Dittakavi, B., Bavikadi, D., Desai, S. V., Chakraborty, S., Reddy, N., Balasubramanian, V. N., … Sharma, A. (2022). Pose Tutor: An Explainable System for Pose Correction in the Wild. *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, *2022-June*, 3539–3548. doi: https://doi.org/10.1109/CVPRW56347.2022.00398

George, S., Dcouth, J., Jaimy, S., Daji, E., & Antony, A. (2022). COMPUTER-ASSISTED TRAINING SYSTEM FOR YOGA AND PILATES USING COMPUTER VISION METHODS. *Journal of Tianjin University Science and Technology*, *55*(06), 176–186. doi: https://doi.org/10.17605/OSF.IO/FXSQE

Gupta, R. (2020). Activity Monitoring of Islamic Prayer (Salat) Postures using Deep Learning. *Roceedings - 2020 6th Conference on Data Science and Machine Learning Applications, CDMA*, (October), 106–111. doi: https://doi.org/10.1109/CDMA47397.2020.00024

Hammadi, Y., Grondin, F., Ferland, F., & Lebel, K. (2022). Evaluation of Various State of the Art Head Pose Estimation Algorithms for Clinical Scenarios. *Sensors*, *22*(18), 1–12. doi: https://doi.org/10.3390/s22186850

Han, K., Yang, Q., & Huang, Z. (2020). A two-stage fall recognition algorithm based on human posture features. *Sensors (Switzerland)*, *20*(23), 1–21. doi: https://doi.org/10.3390/s20236966

Jesmeen, M. Z. H., Bhuvaneswari, T., Mazbah, A. H., Chin, Y. B., Siong, L. H., & Aziz, N. H. A. (2023). SleepCon: Sleeping Posture Recognition Model using Convolutional Neural Network. *Emerging Science Journal*, *7*(1), 50–59. doi: https://doi.org/10.28991/ESJ-2023-07-01-04

Johnston, S. H., Berg, M. F., Eikevåg, S. W., Ege, D. N., Kohtala, S., & Steinert, M. (2022). Pure Vision-Based Motion Tracking for Data-Driven Design - A Simple, Flexible, and Cost-Effective Approach for Capturing Static and Dynamic Interactions. *Proceedings of the Design Society*, *2*(2015), 485–494. doi: https://doi.org/10.1017/pds.2022.50

Khan, T. (2021). An Intelligent Baby Monitor with Automatic Sleeping Posture Detection and Notification. *Ai*, *2*(2), 290–306. doi: https://doi.org/10.3390/ai2020018

Kothari, S. (2020). Yoga Pose Detection and Classification Using Deep Learning (San Jose State University). doi: https://doi.org/10.32628/cseit206623

Koubaa, A., Ammar, A., Benjdira, B., Al-Hadid, A., Kawaf, B., Al-Yahri, S. A., … Ba Ras, M. (2020). Activity Monitoring of Islamic Prayer (Salat) Postures using Deep Learning. *Proceedings - 2020 6th Conference on Data Science and Machine Learning Applications, CDMA 2020*, 106–111. doi: https://doi.org/10.1109/CDMA47397.2020.00024

Kwon, Y., & Kim, D. (2022). Real-Time Workout Posture Correction using OpenCV and MediaPipe. *The Journal of Korean Institute of Information Technology*, *20*(1), 199–208. doi: https://doi.org/10.14801/jkiit.2022.20.1.199

Lawanont, W., Inoue, M., Mongkolnam, P., & Nukoolkit, C. (2018). Neck posture monitoring system based on image detection and smartphone sensors using the prolonged usage classification concept. *IEEJ Transactions on Electrical and Electronic Engineering*, *13*(10), 1501–1510. doi: https://doi.org/10.1002/tee.22778

Leightley, D., & Yap, M. H. (2018). Digital analysis of sit-to-stand in masters athletes, healthy old people, and young adults using a depth sensor. *Healthcare (Switzerland)*, *6*(1). doi: https://doi.org/10.3390/healthcare6010021

Long, C., Jo, E., & Nam, Y. (2022). Development of a yoga posture coaching system using an interactive display based on transfer learning. *Journal of Supercomputing*, *78*(4), 5269–5284. doi: https://doi.org/10.1007/s11227-021-04076-w

Munea, T. L., Jembre, Y. Z., Weldegebriel, H. T., Chen, L., Huang, C., & Yang, C. (2020). The Progress of Human Pose Estimation: A Survey and Taxonomy of Models Applied in 2D Human Pose Estimation. *IEEE Access*, *8*, 133330–133348. doi: https://doi.org/10.1109/ACCESS.2020.3010248

Nguyen, H. T. P., Woo, Y., Huynh, N. N., & Jeong, H. (2022). Scoring of Human Body-Balance Ability on Wobble Board Based on the Geometric Solution. *Applied Sciences (Switzerland)*, *12*(12). doi: https://doi.org/10.3390/app12125967

Nicolae-Adrian, J., Claudiu, A., Ana-Maria, V., & Ciprian, G. (2021). A systematic review of integrated machine learning in posture recognition. *Timisoara Physical Education and Rehabilitation Journal*, *14*(27), 15–20. doi: https://doi.org/10.2478/tperj-2021-0009

Piñero-Fuentes, E., Canas-Moreno, S., Rios-Navarro, A., Domínguez-Morales, M., Sevillano, J. L., & Linares-Barranco, A. (2021). A Deep-Learning Based Posture Detection System for Preventing Telework-Related Musculoskeletal Disorders. *Sensors*, *21*(15), 5236. doi: https://doi.org/10.3390/s21155236

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). doi: https://doi.org/10.3390/s21217315

Wu, Y., Lin, Q., Yang, M., Liu, J., Tian, J., Kapil, D., & Vanderbloemen, L. (2022). A computer vision-based yoga pose grading approach using contrastive skeleton feature representations. *Healthcare (Switzerland)*, *10*(1), 1–12. doi: https://doi.org/10.3390/healthcare10010036

Xu, Z., Qu, W., Cao, H., Dong, M., Li, D., & Qiu, Z. (2022). An Adaptive Human Posture Detection Algorithm Based on Generative Adversarial Network. *Computational Intelligence* *and Neuroscience, 2022*. doi: https://doi.org/10.1155/2022/7193234