Light-Weight Seated Posture Guidance System with Machine Learning and Computer Vision

Rithik Kapoor
rithik.kapoor@mavs.uta.edu
The University of Texas at Arlington
TX, USA

Ashish Jaiswal
ashish.jaiswal@mavs.uta.edu
The University of Texas at Arlington
TX, USA

Fillia Makedon
makedon@uta.edu
The University of Texas at Arlington
TX, USA

ABSTRACT

In today's world, the increased time people spend in front of their computers has been one of the main causes for neck and back pains. Especially, since the pandemic, it has been quite evident that slouching at home for long hours on hand-held devices and computers has led many people towards spinal pains and injuries. Backed with scientific research, it has been proven that these pains can be prevented with proper monitoring of the seated posture and taking breaks in between. This paper focuses on building a light-weight end-to-end system that monitors the user's posture and provides feedback whenever it is necessary for them to fix their posture or take a break. Our system utilizes the most common devices: a webcam or a smartphone camera to capture input frames and a machine learning model to differentiate between good and bad postures with 98% accuracy while the user is seated. The newly developed pipeline helps the users in improving their posture without any additional cost or hardware.

CCS CONCEPTS

• **Computing methodologies** → *Machine learning approaches*; • **Applied computing** → **Consumer health**.

KEYWORDS

pose estimation, machine learning, deep learning, neural networks, computer vision, classification, pose

ACM Reference Format:
Rithik Kapoor, Ashish Jaiswal, and Fillia Makedon. 2022. Light-Weight Seated Posture Guidance System with Machine Learning and Computer Vision. In *The15th International Conference on PErvasive Technologies Related to Assistive Environments (PETRA '22), June 29-July 1, 2022, Corfu, Greece.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3529190.3535341

1 INTRODUCTION

Medical studies [13, 25] have proven that bad posture practice for longer periods results in major back problems. The rise in neck and back pains can result in serious spinal injuries and disabilities. Lower back pain has been identified as one of the leading causes of work-related disability and loss of productivity in industrialized settings [21]. With COVID impacting our daily lives, companies and organizations have diverted to home-based remote working culture to limit the spread of the virus. This has increased in people sitting for a longer period slouching on their chairs. According to an e-survey [8], 48% of the participants worked at home at least sometime during the COVID pandemic and 33.7% worked exclusively from home. This is a huge jump from the one in twenty people that teleworked regularly in 2018 [8]. Among these employees, many of them are teleworking for the first time and do not have the experience of spending long hours in front of the screen or how to maintain proper posture. It, therefore, becomes highly important for us to develop techniques to analyze posture and utilize them to prevent health issues among the millions in the current generation adopting this new lifestyle.

In this work, we categorized a posture as "good posture" based on the guidelines presented in [16]:

- The spine is approximately perpendicular to the thighs and the ears, shoulders, and hips are in a straight line.
- The body's weight is distributed equally and not to just one side.
- The neck is not hunched forward too much.
- The shoulders are not protracted forward and should be relaxed.

On the other hand, any posture that does not match with the above criteria is considered to be a "bad posture". Bad posture for regular prolonged periods can result in numerous health issues like severe low/middle back pain, moderate discomfort in eyes/neck/head, discomfort in the upper back/shoulders, and elevated stress levels as highlighted in [2, 11] as well as cause long terms issues like injuries and spine curvature. Identifying bad posture early can prevent a lot of these health problems, that in the long term, may result in extreme discomfort and permanent disability.

In the recent years, great progress has been made in pose estimation models that can predict various keypoints on the body with a high accuracy. Most of the popular models like OpenPose [5] and BlazePose [4] use machine learning (ML) and deep learning (DL) techniques to detect parts on the human body like eyes, shoulders, hips, etc. Most of the times, pose estimation models are generally used to classify human activities in the scene [26, 28]. But these models can be leveraged to further classify various poses such as good and bad posture as done in this paper. Furthermore, advancements in self-supervised learning methods [15] have enabled the training of these pose estimation models to be much more efficient. Researchers in the past have also tried to use gesture-based hand rehabilitation techniques [7] using computer vision and machine learning, but limited work can be found that deals with posture correction systems.

PETRA '22, June 29-July 1, 2022, Corfu, Greece

ACM ISBN 978-1-4503-9631-8/22/06.
https://doi.org/10.1145/3529190.3535341

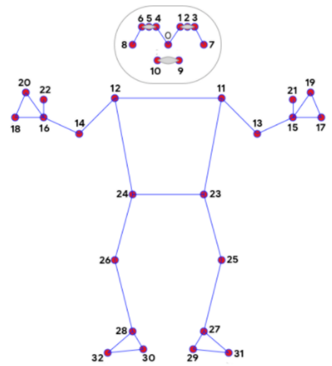
Although there are other posture correction apps on the market, almost all of them involve using some kind of expensive tracker that is placed on the body (like Upright [1]) and do not appear to use machine learning and computer vision. In this paper, we aim to create a system that can be incorporated in a variety of devices like mobile phones and laptops to detect bad posture. This system can then be used to correct the posture of employees who are teleworking by giving them a notification whenever they get into bad posture.

2 RELATED WORK

Researchers have made valiant efforts in building systems that can detect bad postures and provide feedback to the user in order to correct the posture. Most of the works have focused on analysing postures based on data from sensors that are usually attached or worn by the user. In [20], the authors embedded six force sensors into a chair that communicate with a mobile application to notify the user about their bad posture. In addition, the user receives statistical data along with wrong sitting positions. On the other hand, instead of using the force sensors, [24] uses four load cells mounted only on the seat plate to categorize among six pre-defined postures. The idea is to measure the change in load while a person sits in multiple positions. The drawback to these systems is that they are embedded into the chair and not portable for other use.

Furthermore, Manju et al. [12] proposed a posture monitoring system based on flex sensors and load cells. The flex sensor is attached to the user's back in order to monitor the body's bend and determine whether the posture is good or bad. However, the flex sensor cannot provide precise information about the trunk flexion and focuses only on the spine shape. A different modality to measure posture's correctness was found in this work [19] where the authors placed a 3-axis accelerometer on the user's neck in order to define the spine shape. Another use of intertial sensor was found in Low et al. work [18] where they focused on improving posture for opthalmologists. Although the posture monitoring systems based on the intertial sensors (accelerometers) introduced superior advantage in measuring the trunk deviation angle, they suffer from the need to position them optimally for subjects with different shape and size.

The introduction of visual sensors for posture detection was seen in [17] where the authors mounted a 3D depth camera on the ceiling to detect a person falling on the floor. They used a K-Nearest Neighbors (KNN) classifier to distinguish been accidental falls and other generic activities. In another work [14], the authors focused on triggering an alarm once a fall is detected using an RGB-camera sensor. Numerous research have been conducted that uses vision sensors along with some form of features extracted such as skeletal information, optical flow and so on for activity analysis [3, 6, 9, 10, 22, 26–28]. The vision sensors seem to work well for this purpose, with Roberts et al. [23] introducing a vision-based activity analysis framework that estimates and tracks 2D worker pose and outputs activities such as bricklaying and plastering that were being performed by the workers. Even with such advancements in computer vision and deep learning, systems designed to encourage good seating postures that can be implemented in day-to-day life have not been researched extensively. This is where our system comes into play which only requires a simple webcam or a smartphone camera that can track bad posture and notify the user whenever necessary.



0. nose	17. left_pinky
1. left_eye_inner	18. right_pinky
2. left_eye	19. left_index
3. left_eye_outer	20. right_index
4. right_eye_inner	21. left_thumb
5. right_eye	22. right_thumb
6. right_eye_outer	23. left_hip
7. left_ear	24. right_hip
8. right_ear	25. left_knee
9. mouth_left	26. right_knee
10. mouth_right	27. left_ankle
11. left_shoulder	28. right_ankle
12. right_shoulder	29. left_heel
13. left_elbow	30. right_heel
14. right_elbow	31. left_foot_index
15. left_wrist	32. right_foot_index
16. right_wrist	

Figure 1: Keypoints detected by BlazePose [4]

3 METHODOLOGY

Instead of working directly on the input frames recorded by a webcam or a smartphone camera to classify the posture, we first use a pose estimation model called BlazePose [4] to extract 33 different body keypoints (landmarks) from each frame as shown in fig 2. It helps to overcome the complexity of feature extraction from images and increases the overall processing speed of the system. The system is composed of two main components: the pose estimation system (extracting body keypoints) and the posture classification system (good/bad).

For our posture classification system, we use a machine learning model instead of having a simple rule based system that uses joint angles to classify posture. This is because although the latter may be faster and perform well in classifying posture when the camera is placed to the side of the user, it won't be able to calculate the joint angles in the same way when placed in front and is highly sensitive to the position in which the camera is placed. In the case of machine learning models, the camera can be placed to the front or side of the user without much change in accuracy since the models have been trained on frames from a variety of angles in both positions. Our system collects frames continuously and for each frame in a video input stream, it performs the following steps to categorize the type of posture in the frame:

(1) Extract 33 body keypoints using BlazePose from each input frame.
(2) Normalize keypoints by translation and scale as shown in equations 1, 2, 3, and 4.
(3) Check the confidence scores of keypoints 25 (left knee) and 26 (right knee). If they are below 50% in either knee then the legs are probably not visible from where the camera is placed and the model trained with "no limbs" (as described in section 4) is used. If the confidence scores are greater than 50% in both knees, then the model trained with "all keypoints" is used. This enables the model to classify posture accurately even with occlusions in front of the camera, by only using the keypoints from the upper half of the body.

(4) Classify the frame as good or bad posture using normalized keypoints and the selected machine learning model.

Finally, if ten consecutive frames are classified as bad posture, the system issues a sound notification to the user reminding them to fix their posture.

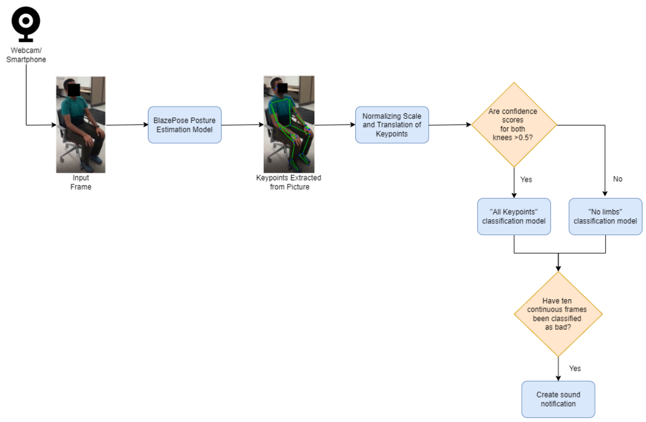


Figure 2: System Overview Diagram

The following subsections explain each of the steps in detail.

3.1 Posture Estimation via BlazePose

The model used to extract the body keypoints from each input frame is BlazePose [4]. BlazePose is a lightweight convolutional neural network architecture for human pose estimation that is tailored for real-time inference on mobile devices. It detects 33 different keypoints on the human body as shown in fig 1. Aside from detecting the x and y coordinate values for each of the keypoints from an RGB image, it also estimates an experimental z value that represents the distance of the keypoint from the camera. Since BlazePose is a fast-face detector that acts as a proxy for a person detector, we do not need to worry about other objects in the scene except for the human. The face is used as the primary feature for the pose detector because of the observation that the face is the strongest signal for the neural network to predict the position of the torso. By detecting the face first, BlazePose is able to run at real-time speeds since it is then able to predict information about the pose's alignment parameters like the middle point between hips and the size of the circle circumscribing the whole person before moving on to detecting the keypoints.

The pose tracker further uses the alignment parameters predicted before to detect the 33 keypoint coordinates in the picture and a refined area of interest to search in the next frame to increase efficiency. When compared to other pose estimation models, BlazePose has been found to be 25-75 times faster on a single mid-tier phone CPU as compared to commonly used models like OpenPose. It can run at over 30 frames per second on a Google Pixel 2 phone [4]. Leveraging BlazePose is critical to make our posture classification system highly accessible and allow users to use it on any device that doesn't have specialized hardware like dedicated GPUs.

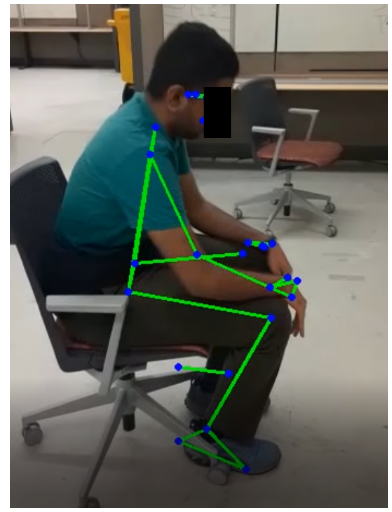


Figure 3: Keypoints detected by BlazePose on Real Person

3.2 Data collection

We conducted our data collection under IRB Protocol #2016-0693. For the data collection, our setup just required a chair and a smartphone. Videos of eleven different people were taken, out of which six were used for the training set and the other five were used for the testing purposes. The participants were between 19 to 46 years of age and consisted of 9 males and 2 females. For each participant, the following steps were carried out to collect data of their good/bad postures:

(1) The participant was instructed to be seated in a good posture on the chair according to guidelines given in section 1.
(2) Videos were recorded at multiple angles from the front and side views. From the front position, a part of the video was taken in a way such that only the upper half of the body was visible. This is to replicate a scenario in which the camera is placed on top of a table such that the legs cannot be seen due to occlusion.
(3) The participant was next instructed to be seated in a bad posture violating one or more of the guidelines presented in section 1 (for example hunching neck too much).
(4) Similar videos were recorded as described in step 2.

These videos were further sliced frame by frame and analyzed by BlazePose to extract the coordinates of the keypoints from the frame. Next, the coordinates were normalized and were saved in a CSV file to be later used for training ML/DL models. A total of 7666 image frames were extracted to make up the training set and 5898 frames to make up the test set. In addition, there was no overlap in data between train and test set from the same participant. This is to ensure that the models are not biased on the test set.

3.3 Normalizing the Keypoint Coordinates

To detect a bad posture, the system has to work from any given position and angle. Since it is not necessary that the subject always aligns in a particular position with the camera, it becomes very important to factor in the changes and variations in the seating positions relative to the camera. On top of that, BlazePose only outputs raw coordinates of the body keypoints relative to each other and are not standardized to any particular coordinate system. Thus, to make the system more robust to human orientation changes, pose normalization methods need to be implemented before passing it to the ML models.

In this paper, we use translation and scale to normalize all the landmark coordinates. To account for translational variations, we first find the coordinates of the middle point M that lies between the left hip (L) and the right hip (R). It can be easily obtained by taking the average of both the coordinates as $M = \begin{pmatrix} m_1 & m_2 & m_3 \end{pmatrix} = \frac{1}{2}(L + H)$. Next, all 33 body keypoint coordinates are subtracted by the coordinates of M to remove the translational variations. It also moves the pose center closer to the origin in the cartesian coordinate system. The above steps can be achieved by the following matrix operations as shown in equations 1 and 2:

$$T = X - C \tag{1}$$

$$C_{33\times 3} = \begin{pmatrix} m_1 & m_2 & m_3 \\ m_1 & m_2 & m_3 \\ \vdots & \vdots & \vdots \\ m_1 & m_2 & m_3 \end{pmatrix} \tag{2}$$

where $T, X, C \in \mathbb{R}^{33\times 3}$, X is the original matrix containing the x, y, and experimental z coordinates of each body keypoint, and T is the matrix normalized by translation.

Finally, to normalize the scale of the pose, first, the maximum distance between any keypoint to the pose center M is calculated. This distance is considered as the scaling factor for that particular human subject. Hence, all the body keypoint coordinates are divided by the calculated maximum distance. It normalizes the scale of each pose resulting in all coordinate values between 0 to 1. The above steps can be represented by the following matrix operations as shown in equations 3 and 4:

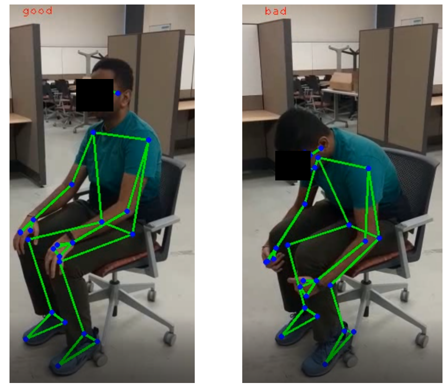
$$S_{33\times 1} = \begin{pmatrix} ||\boldsymbol{t_1}|| \\ ||\boldsymbol{t_2}|| \\ \vdots \\ ||\boldsymbol{t_{33}}|| \end{pmatrix}, \text{where } T = \begin{pmatrix} \boldsymbol{t_1} \\ \boldsymbol{t_2} \\ \vdots \\ \boldsymbol{t_{33}} \end{pmatrix} \text{and } ||t|| = \sqrt{\sum_{i=1}^{3} |\boldsymbol{t_i}|^2} \text{ (L2-norm of } t\text{)} \tag{3}$$

$$N_{33\times 3} = \frac{T}{Max(S)} \tag{4}$$

Here, T is the matrix normalized by translation and N is the final matrix normalized with respect to both translation and scale.

4 RESULTS AND OBSERVATIONS

From the first phase, all body keypoints extracted from the image frames using BlazePose are normalized and stored in separate CSV files for each subjects. The CSV files containing normalized data are only for training and testing purposes. Once the model is deployed, the extracted keypoints are classified in real-time and do not need to be stored anywhere. The pre-processed normalized body keypoints were used to train and compare three different sets of ML models: K-Nearest Neighbors (KNN), XGBoost, and Neural Networks (NN).



(a) Posture classified as good by the NN model.

(b) Posture classified as bad by the NN model.

Figure 4: NN Model Classification Results

For the KNN model, we used the 10 nearest neighbors with uniform weights for classification. For the XGBoost model, we used the tree booster with an eta (learning rate) of 0.3, and max depth of 6. Finally, for the Neural Network model, we used three hidden layers with 100 perceptrons each, a learning rate of 0.001, Relu as the activation function, Adam as the optimizer, and trained the model over a 100 epochs.

In addition to using different models, the combination of keypoints used for training was also varied. For the first case, all 33 body keypoints (All Keypoints) were used for training. Next, the models were trained excluding the keypoints from the limbs (No Limbs). Finally, the experimental depth values (z-values) were dropped and only x and y coordinates were used from all 33 keypoints to train the third set of models (Remove Z). The results from each set of training for the three models can be found in table 1.

We can see that we get the best accuracy and precision in the Neural Network model using all keypoints at 98% accuracy. We also observe that we get a decent accuracy of 93% with the no limbs Neural Network model. This means that we can speed up our posture detection algorithm by just detecting 15 keypoints instead of the original 33. Finally, we observe that removing the Z features doesn't have too much of an effect on the performance of the models and was in fact performing better in case of the KNN model. This suggests that adding a depth sensor would not have benefited the models much and an RGB camera should suffice for our purposes.

All the above models appear to maintain a constant real time performance above 20 FPS without any drop in accuracy on an i7-8750H laptop. The KNN model was further integrated in an Android app and achieved a real time performance of 7-9 FPS on a Samsung A10. On a laptop with an i7-8750H, the model's performance improved to 25-28 FPS. This demonstrates that the posture system explained in this paper is accessible on everyday devices and doesn't require specialized hardware like GPUs.

Finally for comparison, a basic Convolutional Neural Network was trained using the same image frames from the training set used by the above models. The model architecture consisted of three convolutional layers followed by 2 fully connected layers. The model was trained for 20 epochs, and used the Adam optimizer. The trained model was tested on the same testing set as the above models and greatly under performed with an accuracy of just 64% highlighting the effectiveness of our proposed models.

Table 1: Performance comparison among the different models: XGBoost, Neural Networks (NN), and K-Nearest Neighbors (KNN). Acc. represents the accuracy of each model on three different combination set of keypoints: All Keypoints, No Limbs, and Removed-Z. P_G and P_B represent the precisions for classification of Good and Bad postures respectively.

Keypoints	XGBoost			NN			KNN		
	Acc.	P_G	P_B	Acc.	P_G	P_B	Acc.	P_G	P_B
All Keypoints	0.92	0.87	0.98	**0.98**	0.97	0.99	0.92	0.87	0.98
No Limbs	0.89	0.82	0.98	0.93	0.88	0.98	0.92	0.86	0.99
Removed-Z	0.92	0.87	0.99	0.94	0.92	0.96	0.96	0.92	0.99

5 CONCLUSION

In this paper, we present an end-to-end posture corrector system that utilizes computer vision and machine learning techniques to notify the user about their bad posture. The system can take in input frames from any regular webcam or smartphone camera, extract body keypoints (landmarks) from the images, and detect whether the posture is good or bad. Based on experimentation with multiple ML models and different combination sets of body keypoints, we found out that our system works best with a Neural Network model with all 33 body keypoints used for training. With people spending most of their time in front of a computer, it becomes highly important to maintain proper posture to avoid any long-term health issues. This is where our system comes into play being fast, efficient, and being able to provide real-time feedback. For future extensions, using a similar pipeline, the system can assist people with stretching workouts after prolonged periods of being seated in a chair. We are also planning to add detailed feedback report that can notify about the posture behavior of each individual and how it can be improved. Finally, the performance of the model can be further optimized to reach high FPS with any low-end smartphone available in the market.

REFERENCES

[1] 2022. Everyday posture coaching. https://www.uprightpose.com/

[2] Reem S AlOmar, Nouf A AlShamlan, Saad Alawashiz, Yaser Badawood, Badr A Ghwoidi, and Hassan Abugad. 2021. Musculoskeletal symptoms and their associated risk factors among Saudi office workers: a cross-sectional study. *BMC Musculoskeletal Disorders* 22, 1 (2021), 1–9.

[3] Ashwin Ramesh Babu, Mohammad Zakizadeh, James Robert Brady, Diane Calderon, and Fillia Makedon. 2019. An intelligent action recognition system to assess cognitive behavior for executive function disorder. In *2019 IEEE 15th International Conference on Automation Science and Engineering (CASE)*. IEEE, 164–169.

[4] Valentin Bazarevsky, Ivan Grishchenko, Karthik Raveendran, Tyler Zhu, Fan Zhang, and Matthias Grundmann. 2020. Blazepose: On-device real-time body pose tracking. *arXiv preprint arXiv:2006.10204* (2020).

[5] Zhe Cao, Tomas Simon, Shih-En Wei, and Yaser Sheikh. 2017. Realtime multi-person 2d pose estimation using part affinity fields. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 7291–7299.

[6] Alex Dillhoff, Konstantinos Tsiakas, Ashwin Ramesh Babu, Mohammad Zakizadehghariehali, Benjamin Buchanan, Morris Bell, Vassilis Athitsos, and Fillia Makedon. 2019. An automated assessment system for embodied cognition in children: from motion data to executive functioning. In *Proceedings of the 6th international Workshop on Sensor-based Activity Recognition and Interaction*. 1–6.

[7] Farnaz Farahanipad, Harish Ram Nambiappan, Ashish Jaiswal, Maria Kyrarini, and Fillia Makedon. 2020. HAND-REHA: dynamic hand gesture recognition for game-based wrist rehabilitation. In *Proceedings of the 13th ACM International Conference on PErvasive Technologies Related to Assistive Environments*. 1–9.

[8] European Foundation for the Improvement of Living, Working Conditions, et al. 2020. Living, Working and COVID-19—First Findings—April 2020. (2020).

[9] Preetham Ganesh, Reza Etemadi Idgahi, Chinmaya Basavanahally Venkatesh, Ashwin Ramesh Babu, and Maria Kyrarini. 2020. Personalized system for human gym activity recognition using an RGB camera. In *Proceedings of the 13th ACM International Conference on PErvasive Technologies Related to Assistive Environments*. 1–7.

[10] Srujana Gattupalli, Ashwin Ramesh Babu, James Robert Brady, Fillia Makedon, and Vassilis Athitsos. 2018. Towards deep learning based hand keypoints detection for rapid sequential movements from rgb images. In *Proceedings of the 11th PErvasive Technologies Related to Assistive Environments Conference*. 31–37.

[11] Thomas Gerding, Megan Syck, Denise Daniel, Jennifer Naylor, Susan E Kotowski, Gordon L Gillespie, Andrew M Freeman, Thomas R Huston, and Kermit G Davis. 2021. An assessment of ergonomic issues in the home offices of university employees sent home due to the COVID-19 pandemic. *Work* Preprint (2021), 1–12.

[12] Manju Gopinath and Angeline Kirubha. 2015. Real time monitoring of posture to improve ergonomics. *Journal of Biomedical Engineering and Medical Imaging* 2, 2 (2015), 22.

[13] Wilhelmina E Hoogendoorn, Paulien M Bongers, Henrica CW De Vet, Marjolein Douwes, Bart W Koes, Mathilde C Miedema, Geertje AM Ariëns, and Lex M Bouter. 2000. Flexion and rotation of the trunk and lifting at work are risk factors for low back pain: results of a prospective cohort study. *Spine* 25, 23 (2000), 3087–3092.

[14] Ya-Wen Hsu, Jau-Woei Perng, and Hui-Li Liu. 2015. Development of a vision based pedestrian fall detection system with back propagation neural network. In *2015 IEEE/SICE International Symposium on System Integration (SII)*. IEEE, 433–437.

[15] Ashish Jaiswal, Ashwin Ramesh Babu, Mohammad Zaki Zadeh, Debapriya Banerjee, and Fillia Makedon. 2020. A survey on contrastive self-supervised learning. *Technologies* 9, 1 (2020), 2.

[16] Byung Yong Jeong and Ara Yoon. 2014. Ergonomics of office seating and postures. *Journal of the Ergonomics Society of Korea* 33, 2 (2014), 167–174.

[17] Michal Kepski and Bogdan Kwolek. 2014. Fall detection using ceiling-mounted 3d depth camera. In *2014 International conference on computer vision theory and applications (VISAPP)*, Vol. 2. IEEE, 640–647.

[18] Eugene Low, Toong Hai Sam, Kian Sek Tee, Ruzairi Abdul Rahim, Hashim Saim, Wan Nurshazwani Wan Zakaria, Safinaz Mohd Khialdin, Hazlita Isa, and Chin Fhong Soon. 2020. Development of a wireless and ambulatory posture monitoring system. *International Journal of Integrated Engineering* 12, 2 (2020), 170–176.

[19] Sangyong Ma, Woo-Hyeong Cho, Cheng-Hao Quan, and Sangmin Lee. 2016. A sitting posture recognition system based on 3 axis accelerometer. In *2016 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*. IEEE, 1–3.

[20] Slavomir Matuska, Martin Paralic, and Robert Hudec. 2020. A smart system for sitting posture detection based on force sensors and mobile application. *Mobile Information Systems* 2020 (2020).

[21] Nur SA Ramdan, AYB Hasyim, Seri R Kamat, Mohd NA Mokhtar, and Siti A Asmai. 2014. On lower-back pain and its consequence to productivity. *Journal of Industrial and Intelligent Information* 2, 2 (2014).

[22] Ashwin Ramesh Babu, Mohammad Zaki Zadeh, Ashish Jaiswal, Alexis Lueckenhoff, Maria Kyrarini, and Fillia Makedon. 2020. A multi-modal system to assess cognition in children from their physical movements. In *Proceedings of the 2020 International Conference on Multimodal Interaction*. 6–14.

[23] Dominic Roberts, Wilfredo Torres Calderon, Shuai Tang, and Mani Golparvar-Fard. 2020. Vision-based construction worker activity analysis informed by body posture. *Journal of Computing in Civil Engineering* 34, 4 (2020), 04020017.

[24] Jongryun Roh, Hyeong-jun Park, Kwang Jin Lee, Joonho Hyeong, Sayup Kim, and Boreom Lee. 2018. Sitting posture monitoring system based on a low-cost load cell using machine learning. *Sensors* 18, 1 (2018), 208.

[25] Joshua Scott Will, David C Bury, and John A Miller. 2018. Mechanical low back pain. *American family physician* 98, 7 (2018), 421–428.

[26] Mohammad Zaki Zadeh, Ashwin Ramesh Babu, Ashish Jaiswal, Maria Kyrarini, and Fillia Makedon. 2021. Self-supervised human activity recognition by augmenting generative adversarial networks. In *The 14th PErvasive Technologies Related to Assistive Environments Conference*. 171–176.

[27] Mohammad Zaki Zadeh, Ashwin Ramesh Babu, Ashish Jaiswal, Maria Kyrarini, Morris Bell, and Fillia Makedon. 2021. Automated system to measure Tandem Gait to assess executive functions in children. In *The 14th PErvasive Technologies Related to Assistive Environments Conference*. 167–170.

[28] Mohammad Zaki Zadeh, Ashwin Ramesh Babu, Ashish Jaiswal, and Fillia Makedon. 2022. Self-Supervised Human Activity Representation for Embodied Cognition Assessment. *Technologies* 10, 1 (2022), 33.