Detecting Posture Quality with Real-time Pose Estimation

Michael Wu

J-ROBO-001

Table of Contents

Abstract

Nowadays people spend a significant amount of time sitting in front of screens for school, occupational, or leisure activities [1,2]. Prolonged sitting combined with poor posture can contribute to neck pain and back pain, the highest healthcare expenditure among 154 health conditions in the United States, e.g., a total cost of 1,345 billion USD in 2016 [3-6]. However, it is difficult to efficiently prevent poor posture with conventional methods focusing on exercise programs [7], educational programs [8], and wearable biofeedback sensors [9].

A low-cost real-time posture detection system based on computer vision is developed. The developed procedure has advantages over previous methods. First, it requires no extra equipment cost as it utilizes existing hardware such as the built-in camera in PCs and mobile devices. Second, it utilizes the free pose estimation program OpenPose that can be run both locally and on the cloud for human body keypoints detection based on real-time image and video input. Furthermore, it utilizes the free web-based platform Teachable Machine to train pose classification models based on Transfer Learning, a Machine Learning technique that allows a user to add their own data and retrain a model on top of a previously trained base model. A total of 740 posture images of 14 volunteers from different age and gender groups were collected. The developed model achieves more than 95% accuracy after 70 epoch trainings. This detection system can help both computer users and sedentary workers to maintain a healthy posture during a prolonged session through real-time biofeedback.

1. Introduction and Motivation

Nowadays people spend a significant amount of time sitting in front of screens for school, occupational, or leisure activities [1,2]. Sedentary time has been exacerbated by the COVID-19 pandemic with increased screen use in remote learning and working. Prolonged sitting combined with poor posture can contribute to neck pain and back pain, both recognized as the leading global cause of disability in most countries [3]. This results in an enormous economic burden for both individuals and societies due to healthcare costs, decreased productivity, work absenteeism, lost wages, and work compensation [4, 5]. The total cost of healthcare related to back pain and neck pain in 2016 was estimated to be 1,345 billion USD in the United States, which was the highest healthcare expenditure among 154 health conditions [6]. However, it is difficult to efficiently prevent poor posture, with methods focusing on exercise programs [7], educational programs [8], smart chairs with pressure sensors [9], and wearable biofeedback sensors [6].

Artificial Intelligence (AI) is emerging as an effective solution for computer vision based real-time posture detection. AI is the development of computer systems able to do tasks that normally require human intelligence. It has seen tremendous growth in recent years. Meanwhile, machine learning deals with computers being able to “adapt” to different inputs, and “learn” based on data, without specific instructions to do so. For example, deep learning, a kind of machine learning, can classify and identify faces in images, even surpassing human accuracy [10]. This application enables computer vision, a type of AI that derives information from images and other visual inputs.

This work is to build a real-time poor posture detection system using the free pose estimation program OpenPose, and the free web-based machine learning platform Teachable Machine. The developed system can provide a convenient and low-cost solution by utilizing existing hardware such as the built-in camera in laptops, PCs, and mobile phones.

2. Problem Statement and Research Questions

Problem Statement: There are increasing prolonged screentime related posture issues at school, work, and homes. Current existing posture detection methods can be costly with poor accuracy [7-9].

To develop a low-cost and effective solution to enable widespread deployment, the following **Research Questions** were formulated.

1) **Is there a low-cost way to detect poor posture when sitting in front of a screen?**

   Yes. We can use the existing hardware such as the computer camera to generate visual inputs for a locally installed poor posture detection system based on computer vision and a machine learning algorithm.

2) **What computer vision program will be used?**

   OpenPose, which is free for non-commercial use. It is based on an approach called Part Affinity Fields to efficiently generate maps, positions of keypoints, and skeleton images [11]. OpenPose can be run both locally and on the cloud and can detect human body keypoints using webcam, image, and video input.

3) **What machine learning platform will be used?**

   Teachable Machine, which is a free web-based platform developed by Google to train machine learning (ML) classification models. It uses Transfer Learning, an ML technique which allows a user to add their own data and retrain a model on top of a previously trained base model that has learned a specific domain from a large dataset [12].

4) **What input data will be collected and analyzed?**

   Using raw images of sitting postures, positions of keypoints of the human body such as eyes and neck will be generated using OpenPose, then classified by Teachable Machine. These positions can be written as images with the skeletons overlaid, or JSON files with textual representations of positions.

3. Procedure

A flowchart of the developed procedure is shown in Figure 1. It starts with raw images collected capturing a person's sitting postures. The images then feed into the computer vision module, OpenPose, to generate skeleton images containing pose keypoints. At the initial step, selected skeleton images are processed on the machine learning platform, Teachable Machine, to train the detection system model. The trained detection system can then analyze skeleton images for posture quality detection.

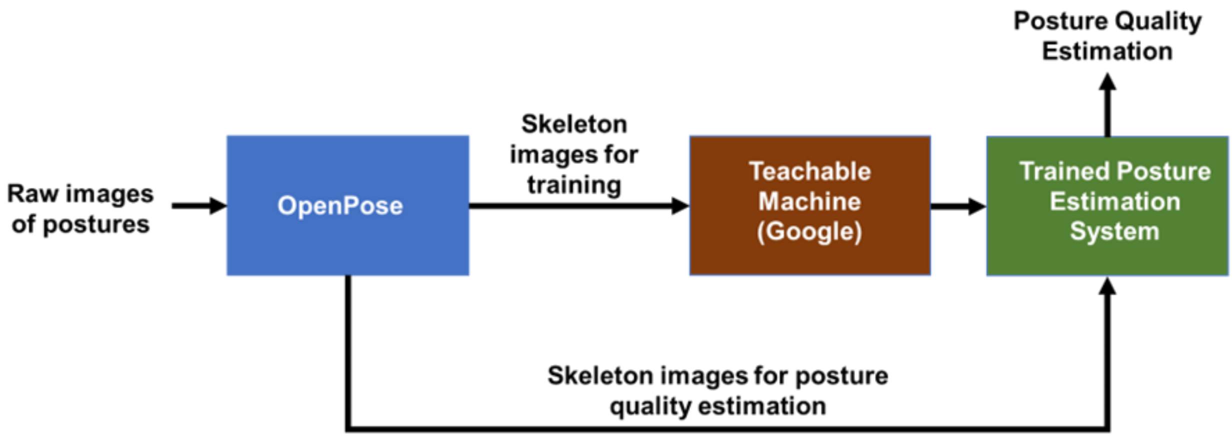


Figure 1. Procedure of the developed posture detection system

3.1 Human Body Keypoint Detection

OpenPose is a real-time human pose detection library developed by Carnegie Mellon University (CMU) to jointly detect the position of human body, foot, hand, and facial keypoints on single images [13]. It uses a bottom-up representation of association scores via Part Affinity Fields, a set of two-dimensional vector fields which record the location and orientation of body parts over the image domain. In addition, it uses CNNs (Convolutional Neural Networks), which have been shown to have a large increase in accuracy over other methods [14]. OpenPose is capable of detecting a total of 135 keypoints, which can be connected to describe an individual's pose.

An example of keypoints detected by OpenPose is shown in Figure 2. The positions can be written as images with the skeletons overlaid, JSON files with textual representations of positions, and skeleton images on a black background. OpenPose can be run both locally and on the cloud and can detect keypoints based on visual input including webcam streaming, still images, or other video input. It is a feedforward network, a type of neural network which predicts a set of 2D confidence maps of body part locations and 2D vector fields, which in turn show the association between parts. Finally, the maps and PAFs are parsed by a greedy algorithm (an algorithm that finds the locally optimal solution) to output the 2D locations of anatomical keypoints [11].

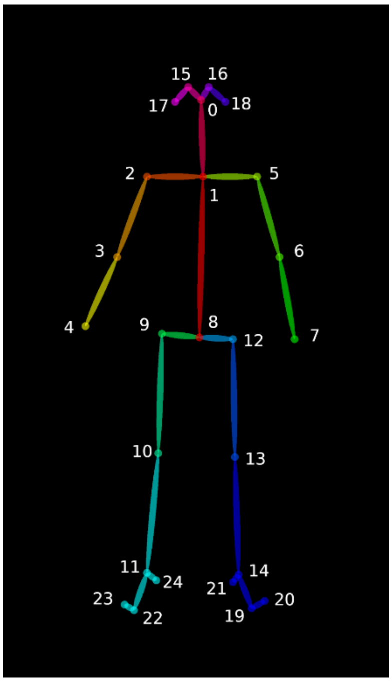


Figure 2. Keypoints detected by OpenPose [13]

3.2 Classification Models

Teachable Machine is Google's free web-based deep learning model creation platform where users can train a machine learning model for classification tasks. Inputs can comprise images, sounds, and poses. It was used for image posture quality detection with input consisting of the aforementioned skeleton pose images, which were sorted into two classes based on whether or not they showed good posture.

The Teachable Machine provides a neural network which is called Multilayer Perceptron (MLP). With the inputs, a neuron randomly assigns weights with which to multiply each input. Each actual input from the image is run through activation functions that determine the output, and there can be multiple "layers." Each layer feeds its results to the next, and this is known as a feedforward system. To improve the weights, at the very end the difference between the results predicted and the correct result is taken, and then "backpropagation" occurs with the hidden weights being updated with the error gradient. The model was set to go through 70 training epochs, and in the end, it was able to accurately determine and classify the posture quality of input skeleton images to an accuracy of over 95%.

3.3 Posture Quality Estimation Criteria

Common body angles such as neck flexion, upper cervical, etc. which are known to cause posture-related health issues were considered when manually classifying images by posture quality. The training images are labeled by their quality based on observing the following key parameters [6]: head tilt angle, neck flexion angle, upper cervical angle, side head tilt angle, and spine angle, as depicted in Figure 3.

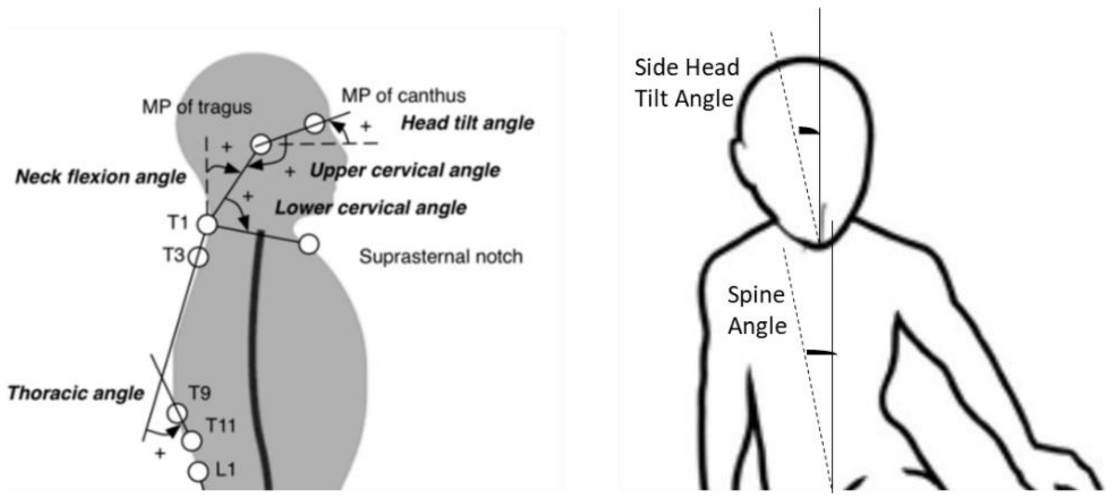


Figure 3. Angle definitions for posture quality estimation [6]

4. Data Collection and Processing

4.1 Participant Group Design

Volunteers were recruited from the school of the student researcher, considering different age and gender groups. The use of human participants was approved by the school's Scientific Review Committee/Institutional Review Board. Written informed consent was obtained from all participants.

In total, 14 volunteers participated, with 5 females and 9 males.

4.2. Raw Data Collection

Raw images showing sitting postures can be captured with any camera. The camera can be a webcam, cell phone camera, digital camera, or any such device. The images show a generally head-on view, similar to that of a webcam of a typical desktop or laptop. The images show the head, neck, shoulders, chest, and optionally more middle parts of the body. In total, 740 images comprised the data collected.

4.3. Create Skeleton Pose Images

A standard Dell laptop was used for OpenPose processing. Raw images of sitting postures were taken and fed into OpenPose to generate the "skeleton" pose images. For demonstration, an overlay of the raw image and its "skeleton" pose is shown in Figure 4. Two generated skeleton-only image outputs from OpenPose are shown in Figure 5 as examples. The skeleton images have 4 or 5 facial keypoints, 2 or 3 spinal keypoints, and 4 total arm keypoints with hands possibly included. The skeleton images are then taken as input to the developed posture quality estimation module running on the Teachable Machine platform.

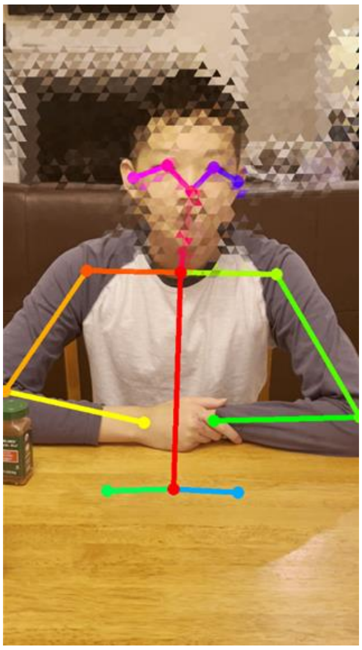


Figure 4. Demonstration of a sitting picture of the student researcher and its “skeleton” pose generated using OpenPose

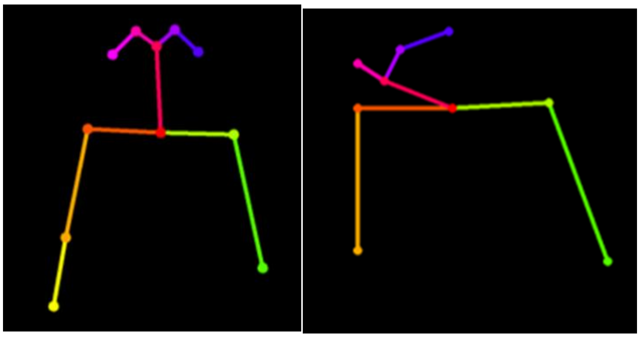


Figure 5. Example skeleton-only image output from OpenPose

5. Classification Model Training and Results

Teachable Machine uses Transfer Learning, which is the process of taking a model previously trained on a data set and applying it to a new data set for classification. It has many advantages over starting from a completely blank model, as it requires fewer data to train the new model and can be significantly faster by only retraining the final few layers of the model architecture instead of the whole network. A schematic comparison between the Transfer Learning and Traditional Machine Learning is shown in Figure 6. Transfer Learning well suits the developed low-cost real-time posture detection system for the reduced time and computation cost, as well as the much smaller training data requirement.

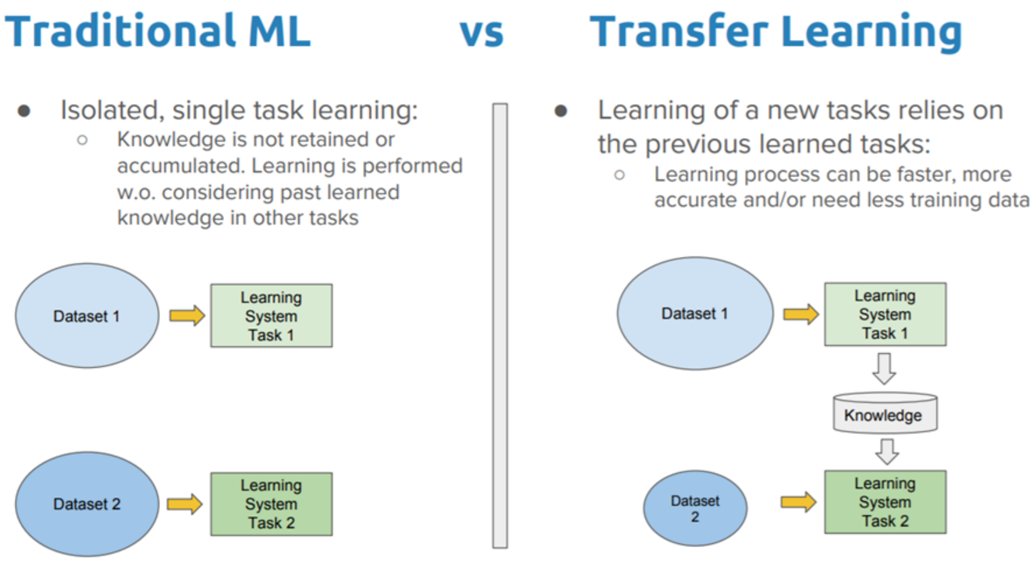


Figure 6. Traditional Machine Learning vs Transfer Learning [17]

5.1 Create the Classification Model

The first step to build a machine learning model is to create different classes for training. Here a Teachable Machine classification model was created with two classes, "Good Posture" and "Bad Posture." The classes take equal numbers of training images of 70 epochs, meaning that each training image was fed through the model 70 times. The rest of the model settings adopt the default recommended by Teachable Machine. More specifically, it consists of 8 layers with a training rate of 0.001 and a batch size of 16.

5.2 Model Training and Results

A total of 332 skeleton images for each class are used for actual training of the machine learning model with 50-50 ratio between good and poor postures. Teachable Machine isolated 15% of the given images for testing the loss and accuracy of the model. The accuracy and loss values are provided as part of the standard output from Teachable Machine. As shown in Figure 7, for the developed training model, the accuracy and loss curves converge rapidly with the number of epochs, even with some fluctuation. The training results converge with over 95% accuracy when the number of epochs is over 70.

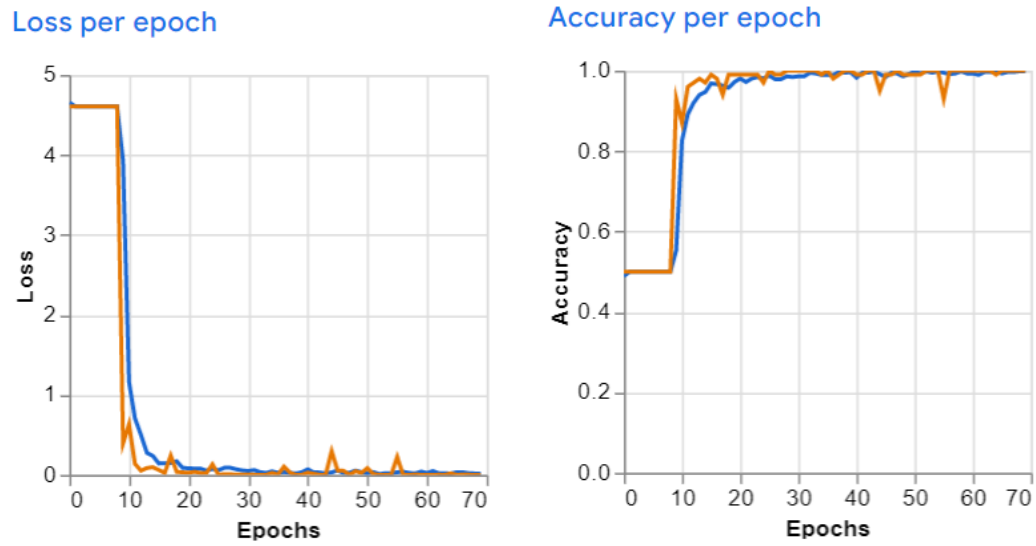
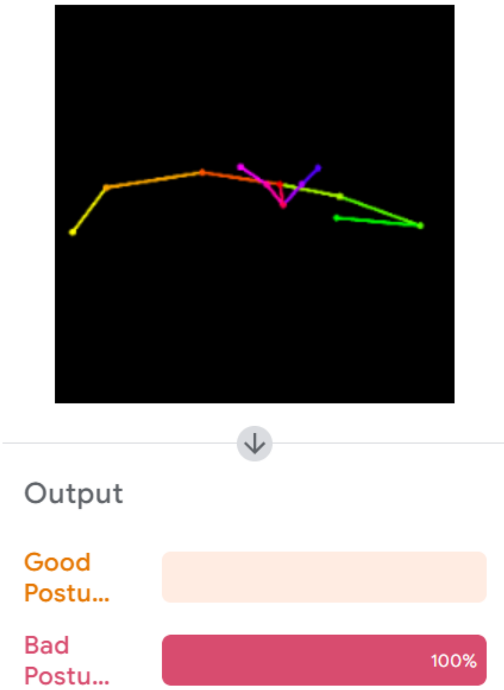
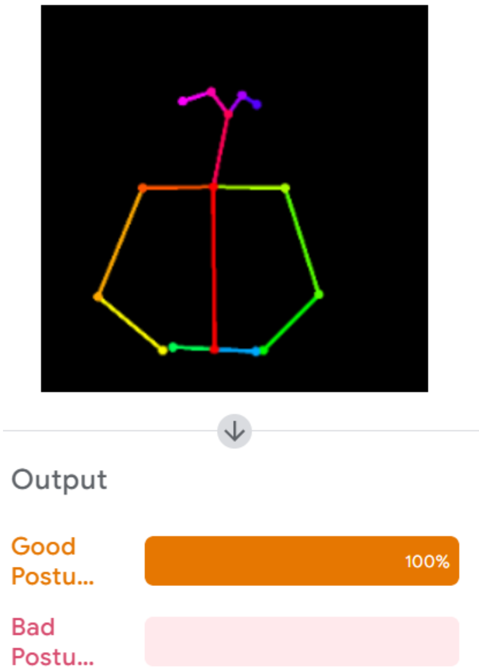


Figure 7. Accuracy and loss values per epoch

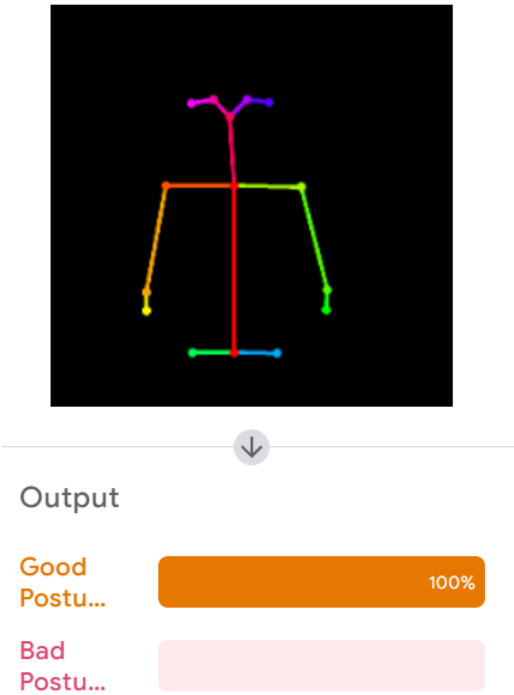
The trained model is further validated using the remaining 76 images with 50-50 ratio between good and poor postures. The detection system outputs "Good Posture" vs. "Bad Posture" classification. In addition, the output also consists of a slider that shows the percent match the input had with each of the two classes. Four typical results are shown in Figure 8 as examples.



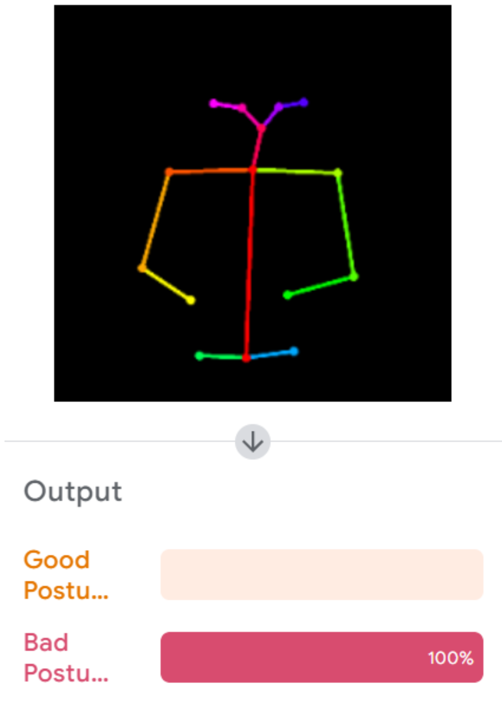
(a) Bad Posture, 100% matching



(b) Good Posure, 100% matching



(c) Good Posture, 100% matching



(d) Bad Posture, 100% matching

Figure 8. Example results from the developed posture detection system

6. Conclusion and Future Work

Nowadays, people spend longer periods in the sitting posture for occupational or leisure activities. Prolonged sitting combined with poor posture can contribute to neck pain and back pain, both recognized as the leading global cause of disability in most countries [3]. In addition, for children and young adults who use computer extensively for learning activities and recreation, it greatly increases their risk of developing spinal pain [18,19]. In a prospective study with 8-year follow-up, experiencing low back pain in youth was found to correlate with low back pain in adulthood [20]; therefore, it is especially important to identify effective approaches for maintain a healthy sitting posture in the younger population [6].

The developed posture detection system provides a low-cost and effective way that can enable a widespread deployment of screentime posture monitoring at school, work, and home. The detected bad posture can be further developed to trigger an alert such as screen flashing and restore the normal screen only after detecting a good posture. In addition, the detection system can be further developed as a mobile app to be completely portable for devices such as smartphones, tablets, laptops, etc. It can help both computer users and sedentary workers to maintain a healthy posture during a prolonged session through real-time biofeedback.

References

1. Mielke, G. I., Burton, N. W., Turrell, G., Brown, W. J. (2018). Temporal trends in sitting time by domain in a cohort of mid-age Australian men and women. Maturitas, 116, 108-115.
2. Matthews, C. E., Chen, K. Y., Freedson, P. S., Buchowski, M. S., Beech, B. M., Pate, R. R., & Troiano, R. P. (2008). Amount of time spent in sedentary behaviors in the United States, 2003–2004. American journal of epidemiology, 167(7), 875-881.
3. Hurwitz, E. L., Randhawa, K., Yu, H., Côté, P., Haldeman, S. (2018). The Global Spine Care Initiative: a summary of the global burden of low back and neck pain studies. European Spine Journal, 27(6), 796-801.
4. March, L., Smith, E. U., Hoy, D. G., Cross, M. J., Sanchez-Riera, L., Blyth, F., Woolf, A. D. (2014). Burden of disability due to musculoskeletal (MSK) disorders. Best practice & research Clinical rheumatology, 28(3), 353-366.
5. Dieleman, J. L., Cao, J., Chapin, A., Chen, C., Li, Z., Liu, A., Murray, C. J. (2020). US health care spending by payer and health condition, 1996-2016. Jama, 323(9), 863-884.
6. Kuo, Y. L., Huang, K. Y., Kao, C. Y., Tsai, Y. J. (2021). Sitting posture during prolonged computer typing with and without a wearable biofeedback sensor. International journal of environmental research and public health, 18(10), 5430.
7. Kim, D., Cho, M., Park, Y., Yang, Y. (2015). Effect of an exercise program for posture correction on musculoskeletal pain. Journal of physical therapy science, 27(6), 1791-1794.
8. Gil, I. (2018). A New Educational Program to Improve Posture and Quality of Life Among Students by Means of the Feldenkrais Method and a New Program: MAP-Motion and Posture. Journal of Education and Training Studies vol 6, 11. doi: 10.11114/jets.v6i11.3582
9. Gelaw, T. A., Hagos, M. T. (2022). Posture Prediction for Healthy Sitting Using a Smart Chair. In Advances of Science and Technology: 9th EAI International Conference, ICAST 2021, Hybrid Event, Bahir Dar, Ethiopia, August 27–29, 2021, Proceedings, Part I (pp. 401-411). Springer International Publishing.
10. Balaban, S. (2015). Deep learning and face recognition: the state of the art. Biometric and surveillance technology for human and activity identification XII, 9457, 68-75.
11. Cao, Z., Simon, T., Wei, S. E., Sheikh, Y. (2017). Realtime multi-person 2d pose estimation using part affinity fields. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 7291-7299).
12. Carney, M., Webster, B., Alvarado, I., Phillips, K., Howell, N., Griffith, J., Chen, A. (2020, April). Teachable machine: Approachable Web-based tool for exploring machine learning

classification. In Extended abstracts of the 2020 CHI conference on human factors in computing systems (pp. 1-8).

13. The Complete Guide to OpenPose in 2023. https://viso.ai/deep-learning/openpose/.
14. Cao, Z., Simon, T., Wei, S. E., Sheikh, Y. (2017). Realtime multi-person 2d pose estimation using part affinity fields. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 7291-7299).
15. Ariëns, G. A. M., Bongers, P. M., Douwes, M., Miedema, M. C., Hoogendoorn, W. E., van der Wal, G., van Mechelen, W. (2001). Are neck flexion, neck rotation, and sitting at work risk factors for neck pain? Results of a prospective cohort study. Occupational and environmental medicine, 58(3), 200-207.
16. Huang, M., Lee, T., Gibson, I., Hajizadeh, K. (2012). Effect of sitting posture on spine joint angles and forces. In Proceedings of the 6th International Conference on Rehabilitation Engineering & Assistive Technology (pp. 1-4).
17. A Newbie-Friendly Guide to Transfer Learning. https://www.v7labs.com/blog/transfer-learning-guide
18. Bhardwaj, Y., & Mahajan, R. (2017). Prevalence of neck pain and disability in computer users. Int J Sci Res, 6, 1288-1290.
19. Hafeez, K., Memon, A. A., Jawaid, M., Usman, S., Usman, S., & Haroon, S. (2013). Back pain–are health care undergraduates at risk?. Iranian journal of public health, 42(8), 819.
20. Hestbaek, L., Leboeuf-Yde, C., Kyvik, K. O., & Manniche, C. (2006). The course of low back pain from adolescence to adulthood: eight-year follow-up of 9600 twins. Spine, 31(4), 468-472.