WebPOSE: A Novel Computer Vision Approach To Posture Correction and Potential Osteoarthritis Prevention

Shivam Syal

Middlesex County Academy for Science, Mathematics and Engineering Technologies, Edison, NJ

Table of Contents

Abstract

Title: WebPOSE: A Novel Computer Vision Approach To Posture Correction and Potential Osteoarthritis Prevention

Name: Shivam Syal

School: Middlesex County Academy of Science, Mathematics and Engineering Technologies, Edison, NJ

Teacher: Dr. Latha Nair

The COVID-19 pandemic has created an indefinite dependency on computers, forcing students and the workforce to remain seated for long periods of time. The development of bad habits such as slouching and inactivity leads to poor posture, and eventually osteoarthritis. By 2025, nearly 60 million people will suffer from osteoarthritis, 40% of them being children, and COVID-19 is exacerbating this issue. Current approaches to posture correction involve expensive physical mechanisms, making scalability for these out-dated solutions poor and impossible for large, diverse populations. To address these shortcomings, I have exploited existing computer vision technologies to develop a zero-cost, highly scalable solution, WebPOSE. Using facial recognition with Haar feature-based cascade classifiers, WebPOSE utilizes the user's webcam to actively measure the current state of their posture and recommends a position change when needed. These changes are calculated with a newly developed POSE-CV Algorithm, which determines the user's distance from their computer's screen. With the implementation of Faster R-CNNs, WebPOSE results in low CPU usage and exponentially fast object detection. The applications of WebPOSE are various, besides just osteoarthritis prevention: student engagement detection, long-term disease prevention, worker productivity quantification, and more. This research paper details the design, development, and future applications of WebPOSE.

1 Introduction

Living in a technological age, there is an ever rising amount of desktop users, with 56.7% of the world's Internet traffic coming from various desktops, including immobile computers, laptops, and other versions of PCs in 2019 [2]. Recently, due to the COVID-19 pandemic, most states and localities have already issued school and public business closures in an effort to slow the spread of the virus. Hence, millions of Americans are now turning to technology for communication, entertainment, and work, causing an unprecedented spike in online data traffic. The first weeks of March 2020 saw an 18% increase in in-home data usage compared to the same period in 2019 [4]. This unique situation that students and the workforce are facing has coupled an indefinite dependency on computers with prolonged hours of sitting, building up habits such as poor posture and hunched sitting [5], which concurrently cause chronic musculoskeletal pain disorders such as osteoarthritis [3][6]. Current approaches to posture correction are exclusively physical mechanisms, including, but not limited to, straps, bands, and/or sensors. Effectiveness of these techniques are questionable, as electromyography done on upper extremities of the body, when the physical mechanisms were in use, was quite variable and proved that the aforementioned techniques provided little to no improvement in posture [1].

To efficiently, inexpensively, and effectively correct the posture of millions, to possibly billions, of desktop users around the world, I present WebPOSE: a real-time desktop application which actively monitors a user's current posture through his/her local webcam video stream and suggests a position change when necessary. All calculations and predictions are done through a newly developed POSE-CV Algorithm which uses facial recognition and trigonometric functions to determine the user's distance from his/her screen. After passing set thresholds, the system alerts the user to change his/her position by either moving "front," "back," "right," or "left." POSE-CV exploits existing computer vision technologies to accurately measure and correct posture, such as the OpenCV library, TensorFlow library, and the Keras API. To make recognitions efficiently and make the program inexpensive on CPU usage, Faster R-CNNs (Recurrent Convolutional Neural Networks) are used.

2 Materials and Methods

All training in this paper was performed on the same laptop with the main configuration of Intel Core i5@3.1 GHzP, 1.5GB GPU Intel Iris Plus Graphics 650 and 8GB random access memory. All programs are written in Python 3 and call OpenCV, TensorFlow, Keras, Numpy and run under a macOS High Sierra 10.13.5 system. The CPU and GPU are used concurrently in training and testing the model. Client-side laptops contained similar configurations.

2.1 POSE-CV Faster R-CNN Model Generation

The face detection framework based on a Faster R-CNN (Ren et al., 2018), with a VGG16 CNN (Simonyan, 2015), shown in Fig. 1, uses human face images in the training sets as inputs and outputs classification results and bounding box coordinates of electrical components in the images [7].

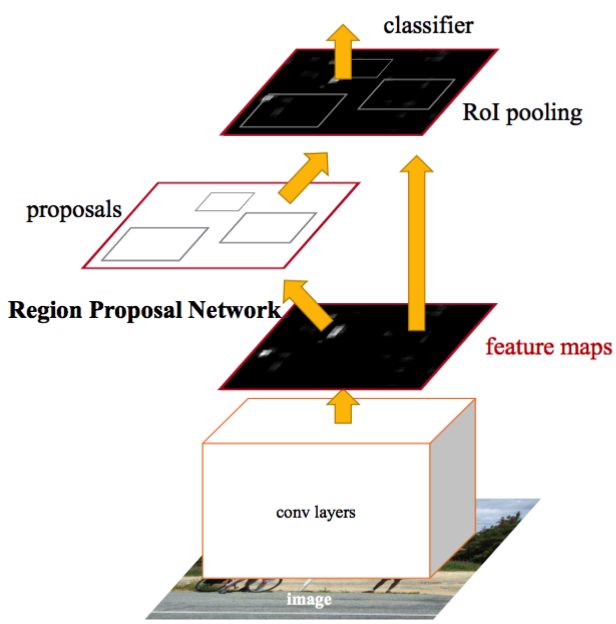

Figure 1: Faster R-CNN is a single, unified network for object detection. The RPN module serves as the 'attention' of this unified network [8].

POSE-CV Model uses the traditional Faster R-CNN model characteristics, utilizing the Regional Proposal Networks (RPNs), sliding windows/anchors, and bounding-box regression. At the end of the day, this implementation uses Caffe to generate "deploy.prototxt.text" and "model.caffemodel" files. Model accuracy plateaus at ~63% while loss is a bit more variable, as shown in Fig. 2. The UTKFace dataset [13] was used for training, a corpus of 20k+ face images with age, gender and race annotations, resulting in the high aforementioned accuracy.

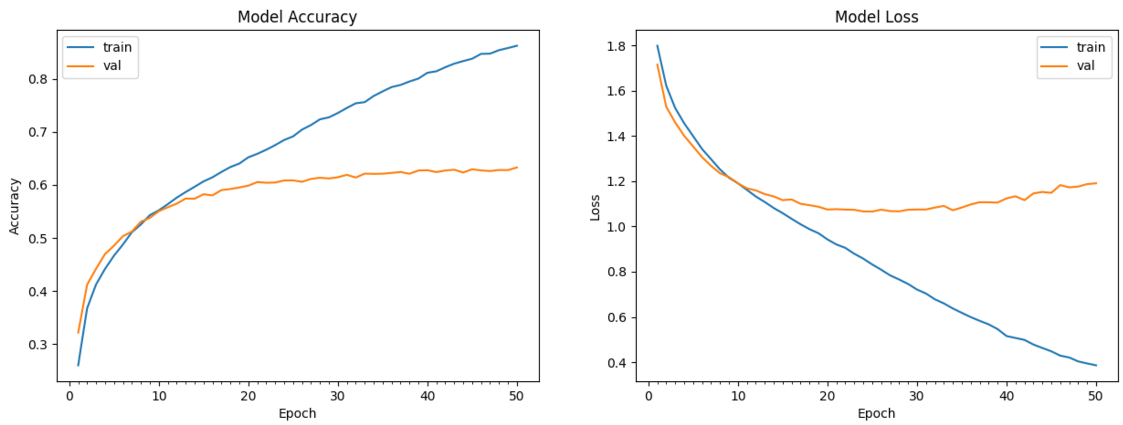

Figure 2: POSE-CV Model's accuracy performs quite well on testing values alongside comparable loss.

2.2 POSE-CV Distance Calculation

To detect the posture of a user, POSE-CV uses triangle similarity. Initially, we need to calibrate the webcam in order to find the true distance of a user. To do this, the following equation is used.

$$Distance = \frac{(Object\ Width \cdot Webcam\ Focal\ Length)}{Perceived\ Object\ Width}$$

POSE-CV can determine focal length of the webcam using OpenCV by creating a camera matrix, shown in Fig. 3. To determine the $Perceived\ Object\ Width$, POSE-CV uses facial recognition. With the use of Haar feature-based classifier cascades and the Faster R-CNN model, POSE-CV finds the user's face in the video stream's current frame. To do this, POSE-CV calculates and returns the minimal up-right bounding rectangle for the specified point set or non-zero pixels of gray-scale image. The width of this rectangle is then set as the $Perceived\ Object\ Width$. Using automatic image processing with OpenCV, the $Perceived\ Object\ Width$ pixel value is converted to inches, satisfying the $Object\ Width$ variable.

$$camera\ matrix = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}$$

Figure 3: Intrinsic and extrinsic parameters of the webcam are stored in the 3x3 camera matrix.

After the distance is calculated, it is evaluated at a 'POSE value' of 0-200, which represents the degree of the user's posture. Optimal posture as shown in Fig. 4 (~135 degrees [11]) is set between the threshold of 50-70, with 0 representing the user sitting at a 180 degree angle, and 200 representing the user sitting at a 0 degree angle (all respective to the sagittal axis of the body). Sitting up straight (parallel to a straight wall) would be a 90 degree angle posture.

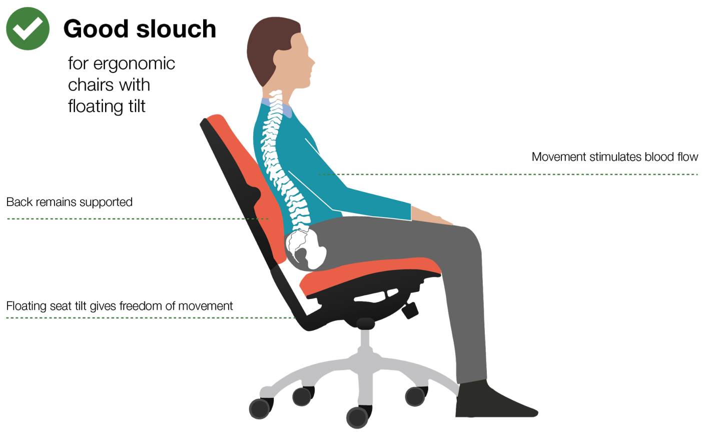

Figure 4: Optimal seating position is ~135 degree angle to ensure lumbar support [11][12]

2.4 POSE-CV Architecture

The diagram shown below, Fig. 5, to detail the process of a refined WebPOSE.

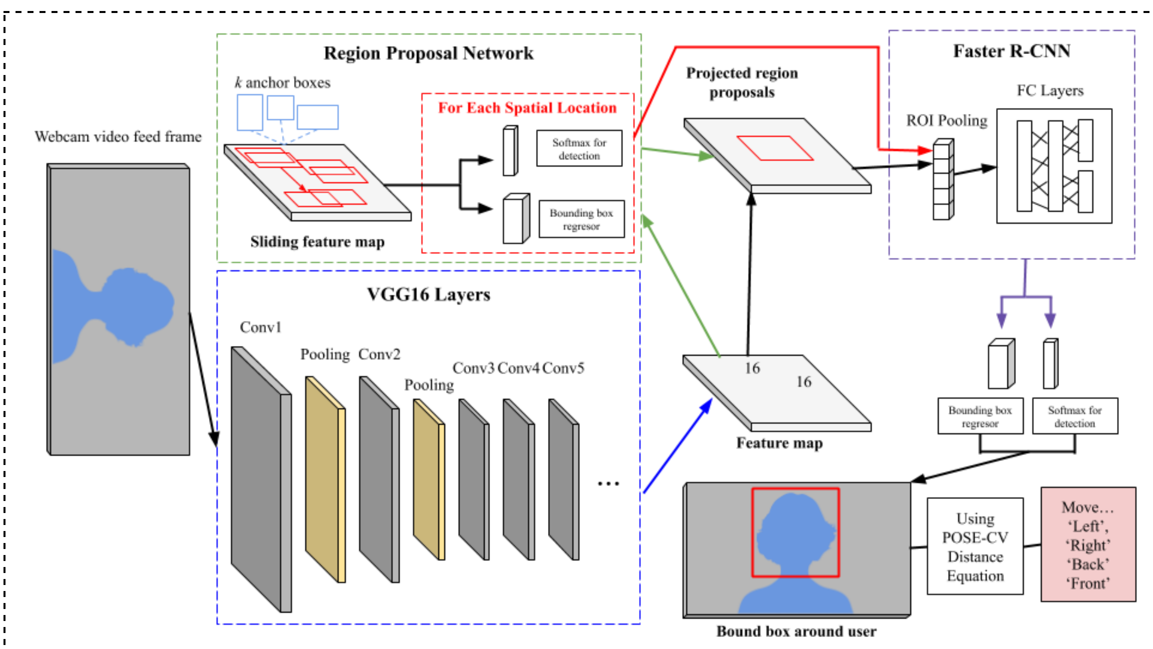

Figure 5: POSE-CV Architecture

As shown in Fig. 5 at the highlighted red box, there is a final output for the WebPOSE program. After detecting what the posture currently is with the POSE-CV Distance Equation, a ‘text-to-speech’ output is played to inform the user to move a certain direction. Alongside that, the user’s screen brightness goes down, to further inform that some sort of posture change is necessary. These visual and auditory indicators make the program effective in any scenario.

2.3 User Experimental Setup

19 users were contacted and consented to the testing of WebPOSE’s effectiveness. All 19 subjects filled out consent forms, allowing the results of their usage of WebPOSE to be used in this paper anonymously. Over the course of one work week, users were instructed to run the application on their laptops and stay seated on a chair, placing their device on a flat surface (e.g. desk, table, etc.). After their work/school day was completed, they were to turn off the application. Their average posture result would automatically be recorded in a spreadsheet after the application was stopped. For each of the five days of use, a different part of the application would be turned on. Day 1 was a baseline day, where just the posture was recorded with no posture correction notifications. On Day 2, just the audio alert feature was turned on, where the ‘text-to-speech’ output would play. On Day 3, just the visual alert feature was turned on, where the screen would dim considerably whenever the user maintained poor posture (no position change suggestion was displayed). Days 4 and 5 were when all of the application’s features were turned on (visual and auditory). An example user interacting with a GUI is shown in Fig. 6.

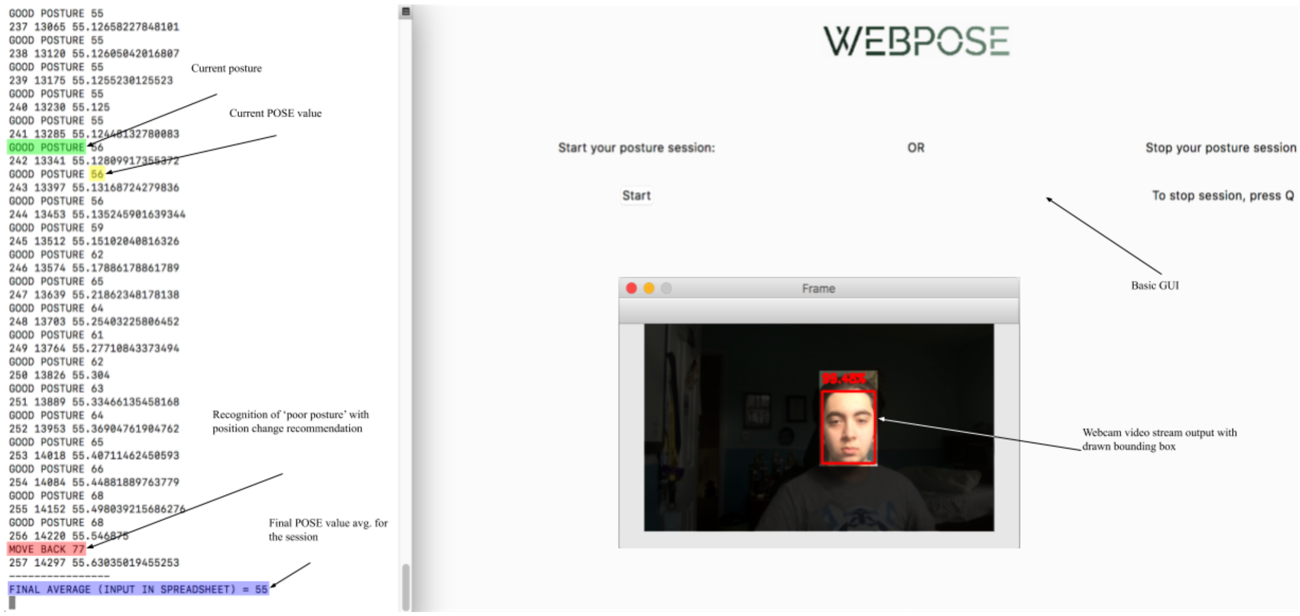

Figure 6: Annotated trial run of WebPOSE (Image is of Shivam Syal)

3 Results

As mentioned previously, the testing done for each individual was over a 5 day period. After the user had used their device with WebPOSE on for their work time, or >2 hours, they were to end the program and the average POSE value for that day was automatically inputted into the spreadsheet using the Google Sheets API. We can see in Fig. 7, below, Day 1 (baseline) was the day with the highest POSE values. They all exceeded the 50-70 threshold, indicating poor posture amongst all participants. However, there is a continuous decline for all participants till Day 5.

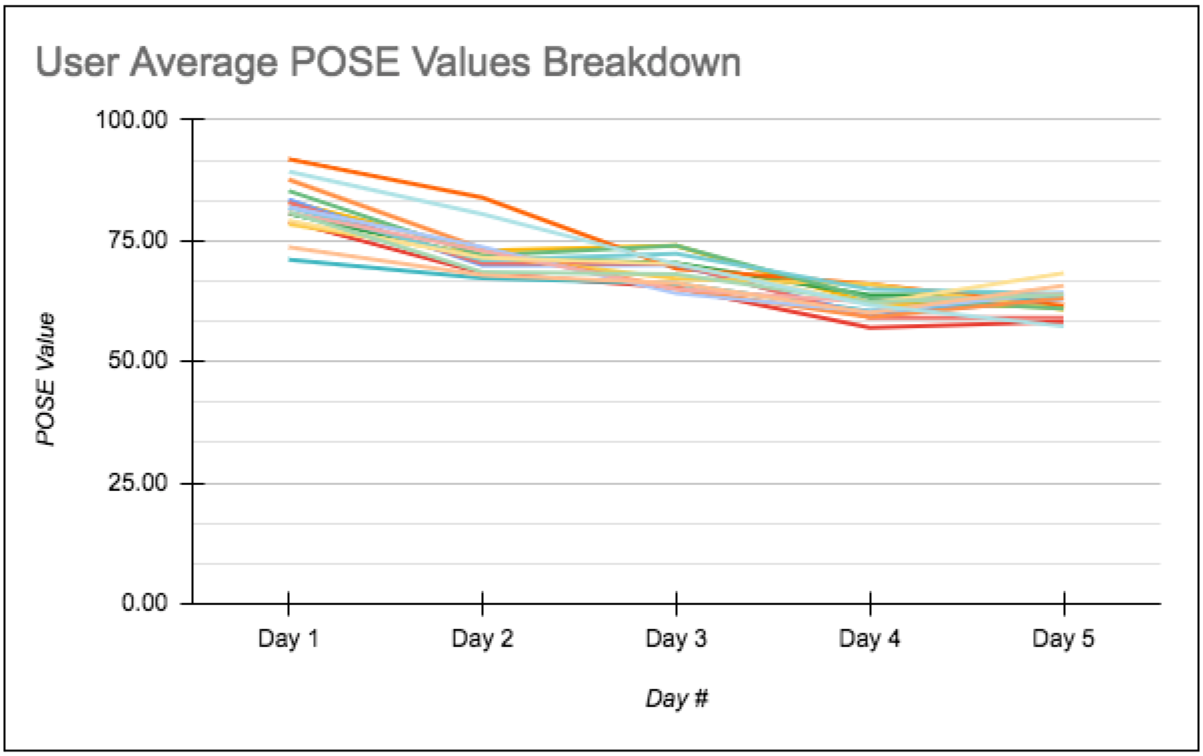

Figure 7: Average values by user, by day. Each line indicates a unique user.

We see in Fig. 8, below, the cumulative averages have a linear negative correlation, with a coefficient of determination of $R^2 = -0.901$. However, there is somewhat of a rise in the average of Day 5.

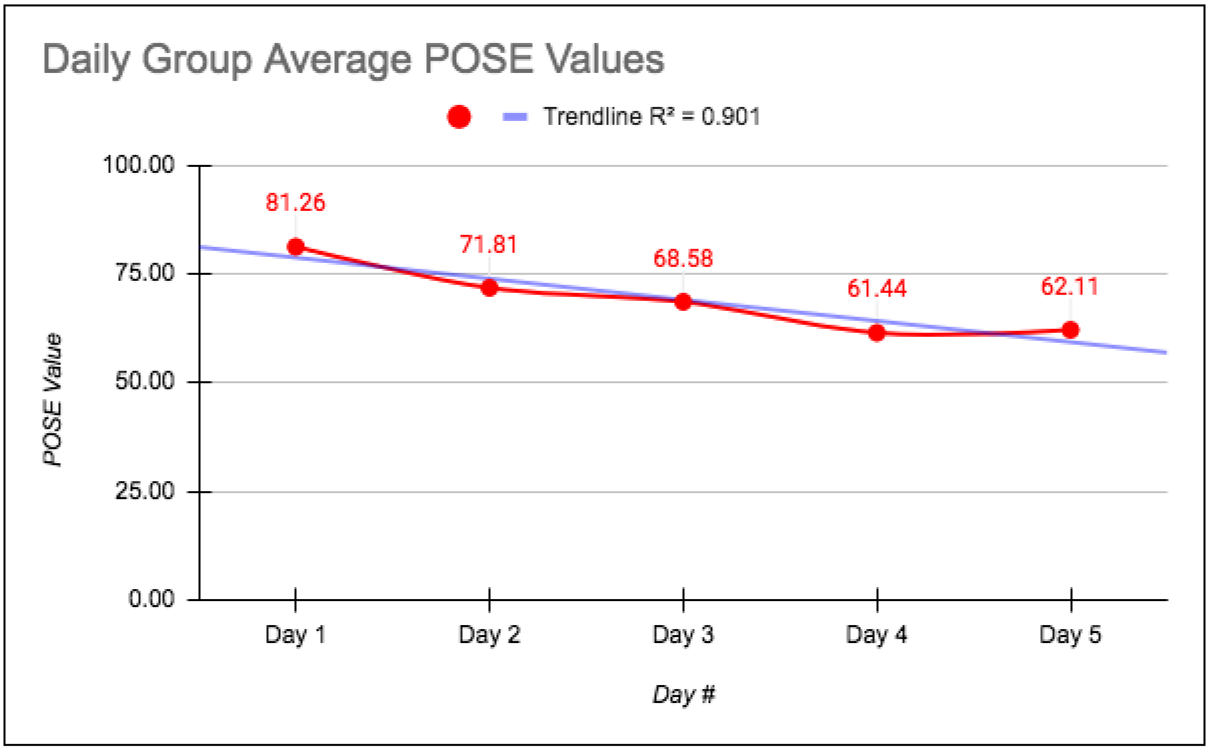

Figure 8: Cumulative user average of each day observed.

4 Discussion and Conclusions

4.1 Discussion

We see in Fig. 8, the average POSE values of all users have a linear negative correlation, proving that the application is moving users' postures closer to the 135 degree optimal posture mark [11]. However, what is interesting is the fact that the visual alert (dimming of the screen - Day 3) was more effective overall than the auditory alert (text-to-speech output - Day 2). Nevertheless, the combined alerts (dimming of screen and text-to-speech output - Days 4 and 5) proved to be most effective. Regardless of the slight increase in cumulative average from Day 4 to Day 5 (61.44 to 62.11), full implementation of the WebPOSE features has proven to correct users' position from ~93 degree posture to a ~132 degree posture.

4.2 Conclusions

Besides exclusively posture correction, WebPOSE has a direct effect on osteoarthritis prevention. The primary causes of osteoarthritis are being overweight and injury. Good posture, however, can reduce overall weight and generally prevent development of osteoarthritis [14][15]. This makes WebPOSE a prospective osteoarthritis prevention application due to its proven effectiveness in bettering poor posture.

What has been presented in this paper is a novel, cost-free, efficient posture correction method which uses minimal CPU intensive and accurate Faster R-CNNs to detect a user's face in the webcam video feed and make a position change prediction with the newly developed POSE-CV algorithm, all while exploiting existing computer vision libraries/technologies. Usage of Faster R-CNNs enables a unified, deep-learning-based object detection system to run at near real-time frame rates [8].

The uses of WebPOSE are not only endless, including student engagement detection, long-term disease prevention, worker productivity quantification, and more; they are highly beneficial to our technology dependent society in a time of need.

5 Related Work

The objective of my project, correcting user posture for potential osteoarthritis prevention using his/her webcam, is similar to that of Taieb-Maimon et al., (2011) [10]. The paper outlined an automatic frequent-feedback system that displayed on the computer screen a photo of the worker's current sitting posture together with the correct posture photo taken earlier during office training. This interval iterating system aims to make the user more aware of his/her posture, thus reducing musculoskeletal risk. Results showed that sustained improvement was attained with the photo-training method [10]. Although the Taieb-Maimon et al. system proved effective, it contained many flaws. The system required numerous data captures before usage and did not provide true feedback on current posture. If the calibration image taken during office training represented bad posture, all feedback would be maligned and useless. Alongside that, the users needed an external webcam set up on either side of their body in order to take the interval based pictures which were displayed on the user's device's screen. These flaws represent the system as inefficient and not able to provide a true, adaptive feedback system. However, WebPOSE fixes all the aforementioned drawbacks with the implementation of machine learning and using only the user's webcam. WebPOSE can truly help millions, and potentially billions, of desktop users inexpensively, efficiently and effectively.

Literature Cited

[1] Cole, A. K., Mcgrath, M. L., Harrington, S. E., Padua, D. A., Rucinski, T. J., & Prentice, W. E. (2013). Scapular Bracing and Alteration of Posture and Muscle Activity in Overhead Athletes With Poor Posture. *Journal of Athletic Training, 48*(1), 12-24. doi:10.4085/1062-6050-48.1.13.

[2] B. (2020). Mobile Vs. Desktop Internet Usage (Latest 2020 Data). Retrieved November 4, 2020 from http://www.broadbandsearch.net/blog/mobile-desktop-internet-usage-statistics.

[3] Obembe, A. O., Johnson, O. E., Tanimowo, T. O., Onigbinde, A. T., & Emechete, A. A. (2013). Musculoskeletal pain among undergraduate laptop users in a Nigerian University. *Journal of Back and Musculoskeletal Rehabilitation, 26*(4), 389-395. doi:10.3233/bmr-130397.

[4] Clement, J. (2020, September 10). Topic: Coronavirus: Impact on online usage in the U.S. Retrieved November 4, 2020, from http://www.statista.com/topics/6241/coronavirus-impact-on-online-usage-in-the-us/.

[5] Daneshmandi, H., Choobineh, A., Ghaem, H., & Karimi, M. (2017). Adverse Effects of Prolonged Sitting Behavior on the General Health of Office Workers. *Journal of Lifestyle Medicine, 7*(2), 69-75. doi:10.15280/jlm.2017.7.2.69.

[6] Gauchard, G. C., Vançon, G., Meyer, P., Mainard, D., & Perrin, P. P. (2010). On the role of knee joint in balance control and postural strategies: Effects of total knee replacement in elderly subjects with knee osteoarthritis. *Gait & Posture, 32*(2), 155-160. doi:10.1016/j.gaitpost.2010.04.002.

[7] Wan, S., & Goudos, S. (2020). Faster R-CNN for multi-class fruit detection using a robotic vision system. *Computer Networks, 168*, 107036. doi:10.1016/j.comnet.2019.107036.

[8] Ren, S., He, K., Girshick, R., & Sun, J. (2017). Faster R-CNN: Towards Real-Time Object

Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 39*(6), 1137-1149. doi:10.1109/tpami.2016.2577031.

[9] Lecun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., & Jackel, L. D. (1989). Backpropagation Applied to Handwritten Zip Code Recognition. *Neural Computation, 1*(4), 541-551. doi:10.1162/neco.1989.1.4.541.

[10] Taieb-Maimon, M., Cwikel, J., Shapira, B., & Orenstein, I. (2012). The effectiveness of a training method using self-modeling webcam photos for reducing musculoskeletal risk among office workers using computers. *Applied Ergonomics, 43*(2), 376-385. doi:10.1016/j.apergo.2011.05.015.

[11] Torio, T., M.D., Pope, M., Ph.D., Takahashi, K., M.D., & Smith, F. W., M.D. (2006, November 27). Aching Back? Sitting Up Straight Could Be the Culprit. Retrieved November 14, 2020, from https://press.rsna.org/timssnet/media/pressreleases/pr_target.cfm?ID=294.

[12] Posturite. (Unknown). *Art of Sitting - Good Slouch* [Image]. Posturite. https://www.posturite.co.uk/media/resources/art-of-sitting-good-slouch.png.

[13] Zhang, Z., Song, Y., & Qi, H. (2017). Age Progression/Regression by Conditional Adversarial Autoencoder. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). doi:10.1109/cvpr.2017.463.

[14] BayCare. (n.d.). How to Prevent Osteoarthritis. Retrieved November 10, 2020, from https://baycare.org/health-library/how-to-prevent-osteoarthritis

[15] Gulf Coast Bariatrics Staff. (2012, October 2). Sit Up Straight, Lose Weight. Retrieved November 11, 2020, from https://www.gulfcoastbariatrics.com/weight-loss-surgery-blog/sit-up-straight-lose-weight