HEALTH AND WELLNESS MONITORING OF COMPUTER USERS USING MACHINE LEARNING

[1]Terone Lorence, [2]Pamuditha Senadeera, [3]Himantha Amarathunga, [4]Wishva Perera, [5]Dr. Anuradha Karunasena, [6]Lokesha Weerasinghe

[1,2,3,4]Student/Sri Lanka Institute of Information Technology (SLLIT) Faculty of Computing Malabe Sri Lanka
[5]Asistant Professor/ Sri Lanka Institute of Information Technology (SLLIT) Faculty of Computing Malabe Sri Lanka
[6]Lecturer / Sri Lanka Institute of Information Technology (SLLIT) Faculty of Computing Malabe Sri Lanka

ABSTRACT

People whose occupation is to work all day sitting at a desk has a high possibility of having long-term serious health-related consequences due to the sedentary lifestyle they spend. This lack of movement, bad posture, exposure to monitor screens all day long, and lack of water intake will give an individual a range of medical issues all over the body, some of which can be life-threatening. This research has been carried out in order to identify the behavioral patterns that will lead to medical complications in the future and inform the user on what actions must be taken to mitigate the possibility of having such complications in the future. This is done by actively monitoring different physical activities and aspects of the user such as whether the user is sitting for too long without physical movement, whether the user has bad posture when sitting in front of the computer, and whether the user is hydrating themselves in between regular intervals. This monitoring is done by utilizing image processing. The system will also monitor whether the user is stressed for long periods of time, this is done by examining the user's audio with voice analysis. The system demonstrates high success rates with high recognition and user-friendly interfaces along with noninvasive notification, the system is both effective and productive when implemented in the real world.

Keywords*: Health, Artificial Intelligence, Image Processing, Voice Analysis, Posture Detection, CVS, Sedentary Behavior, Hydration Monitoring*

---

1. INTRODUCTION

In the last few decades, Information Technology (IT) has taken over every business sector and has become an essential tool in every business no matter what the scale is. IT has become a part of almost every aspect of the day-to-day life of people and an asset that helps businesses reach further and accomplish greater goals that otherwise would not be possible [1]. With the increased use of IT, computer usage has increased and the usage of computers to generate valuable information from the raw data of companies with much higher precision and accuracy has enabled business organizations to make better decisions [2]. And these decisions are backed by features that come along with the use of computers such as reliability, accuracy, performance, and availability ensuring that the information generated by computers is precise and correct [3].

Although IT has many benefits, extensive computer usage has many adverse effects as well. With the increase in the use of IT, it is inevitable that people will be victims of these adverse effects. Adverse effects caused by high computer usage could range from addictions to physical health effects and mental health effects [4].

Computer Vision Syndrome (CVS) is a medical condition that is a result of the excessive computer. The condition only causes ocular discomforts such as eyestrain, blurred vision, dry eyes, and double vision but it also causes physical discomforts such as stiff shoulders, neck pain, headaches, and dizziness [5]. It was observed that the blinking rate of an individual while using the computer for a prolonged period of time was 7 blinks and the normal rate is 22 blinks under relaxed conditions [6].

Apart from CVS, prolonged computer uses with bad postures such as a hunched back and sitting for a long period of time without taking any breaks could lead to serious physical injuries in the long run. Complications such as slip disks are very common among computer users. It has been identified through research work and medical studies that the main reason for slip disks is bad posture [7]. In addition to physical discomfort caused by excessive computer usage as above, being seated while working for a long period of time may be stressful for employees as well. This would lead to lower concentration, and it also drastically drops the productivity and performance of an individual.

In order to stay healthy during computer usage, one must be used to developing good habits and precautions. Precautions such as maintaining a good posture, taking breaks frequently, and looking away from the screen for at least a few minutes are helpful to reduce the physical and mental discomforts of excessive use of computers [8]. Even though these activities seem to consume time, relaxed and happy individuals have shown to be more productive and accurate in their work than individuals who are stressed. Hence this application is developed to help users keep in touch with their physical and mental health by tracking their movements and analyzing their behavior. When using this application, the users do not need to stay vigilant about their stress levels, blinking, posture, physical activity, or hydration pattern. The application will track the user and inform the user when they need to do a certain action. And with time, the user will automatically adopt good health habits because of this application.

In the context of this research, some of the critical health issues that arise in office workers who use computers for prolonged hours are identified and these health issues will be mitigated with the use of technology and by making the user aware of it [2]. Machine learning algorithms have been used to identify the users' behavioral patterns in order to correctly identify the patterns that negatively affect the user and advice the user to correct them accordingly which would ensure that the user is healthy and the productiveness increases.

2. LITERATURE REVIEW

Computer use by employees has been increasing rapidly in the past decade. However, the main issue with the use of computers is that it contributes to the increase in sedentary behavior and disrupts the physical activity of an individual [1]. Not only does the prolonged use of computers affects the physical health of an individual but it also contributes to the mental health of an individual.

In the research work done in [4], a relationship has been identified between the prolonged use of computers and the mental health-related symptoms among university students and office employees associated with work that can only be done with the use of computers. A qualitative study has been conducted with samples of 16 women and 16 men both of the ages of 21 to 28. According to this research, it is shown that users who use computers for a long period of time were stressed but they were not aware of it. It is shown that they did not want to even answer phone calls and just wanted to be left alone. Many companies do not consider the mental and physical well-being of individuals.

Stress causes several mental diseases. Stress causes sadness and suicide. These diseases develop slowly. Long-term stress causes them. Thus, early stress detection and treatment are crucial. Stress and its auditory effects have been studied extensively.

According to the research work done in [5] a real-time stress detection model could be developed using voice analysis. A set of features in the voice could be extracted from a given voice audio sample. Features such as intensity and maximum amplitude can be taken into consideration. Features such as spectral and time domains can be used for emotion detection in speech. In [6] Various types of algorithms have been used for the detection of the effective states from the voice features that are extracted from the audio signals. Some algorithms that have been used are Gaussian Mixture Model (GMM) and Hidden Markov Model (HMM). This analysis is done without trying to understand the content or the meaning of the actual speech.

The usage of computers for a prolonged period of time causes eye strain. Prolonged stress on the eyes can lead to Computer Vision Syndrome (CVS). Existing research has used eye-tracking methods to detect the possibility of having CVS. However, such research is focused on fewer aspects of CVS. Blink detection using variance map computation and eye corner analysis was suggested [7]. Real-time findings on 320 x 240 pictures are 95% positive and accurate, however, head motions impair variance map consumption and performance. The following program uses the eye corners, eyelids, and iris to detect eye blinks. Normal flow estimates motion. An affine model reconstructs head motion apart from ocular movements. The device can track iris and eyelid movements with over 90% accuracy, but not in real-time [6]. A deterministic finite state machine was then incorporated into a study in order to estimate the blink characteristics and examine normal flow. According to the study [7], a system that identifies blinks in real time by correlating 320 x 240 photos from a webcam with an open-eye template is 95% accurate in spotting blinks. The low computing complexity of template-based techniques is an advantage, but they do not differentiate between the two states of the eyes, open and closed. Any transition between them is not clearly defined. The degree of eye closure is identified in this approach [6] using a boosted classifier. A Hidden Markow Model can capture the movement of eye states. Their technique identifies more than 96% of eye blinks in real time on 320 x 240 webcam video. The approach has difficulties when non-standard eye movements, including partial blinks, occur since examples of natural eye motions are utilized to train the model. Two appearance-based trackers were suggested in the research project [6] for tracking the iris, eyelids, and blink detection. The approach apparently worked in real-time and consistently delivered tracking results using input video of low resolution and a simple and direct appearance model.

Another factor that causes serious physical effects is dehydration. The dehydration of office workers can be insidious. Studies have found that even mild dehydration such as 1% dehydration can decrease the productivity of the employees by 12%. If the employee is 3-4% dehydrated their productivity will drop by 25% [1]. Hence researchers have taken interest in finding solutions for the detection and measuring of water intake by means of computer vision and machine learning. One of the ways this is done is with the utilization of 3D joint position and Local Occupancy Pattern (LOP). These features have the capability to characterize the motions of humans and the interactions that take place between humans and objects [2]. With the use of these two techniques, the researchers have been able to identify a series of actions taken by the test subjects including the detection of water drinking [2]. The next research was carried out to recognize the actions using the Bed Aligned Map (BAM) dataset for an elderly care application. The data is captured using a depth camera that is placed above the bed and the data is stored in the BAM descriptor [3]. A Bag of Words (BoW) encodes activity sequences to identify them. K-means clustering learned a 1000-word codebook. Then they use Locality-constrained Linear Coding (LLC) or Vector Quantization (VQ) with sum polling to create a BoW representation [3]. Finally, the researchers normalized the features to unit variance and zero means and applied the liner multi-class SVM for action recognition [3]. The following paper uses max-min features and key poses with differential evolution random forests classifier to recognize human activities. This application contributes in three ways. [4]. They are, the researcher was able to provide a simple and effective approach in extracting extremal skeleton information that is based on the actions that are variable in size and have been determined by key poses, this requires very few training examples and has a training time that is very fast making it real-time application suitable, Also it is said that the Random Forest (RF) classifier which has no thresholds to tune and where the nodes that split best in the decision tree are found [4].

Just like dehydration, sedentary behavior will cause ill effects on the body in the long run. Most office workers settle down to a sedentary lifestyle with time. This carries a major health issue since studies have shown that sedentary behavior can cause different health complications such as a doubled risk of cardiovascular diseases, diabetes, colon cancer, and overall increased risk of mortality [5]. The researchers have developed a mobile application that interacts with a set of sensors that acquire the data and send it to the application. This data is then used by the application to perform intelligent algorithms and then the application will suggest whether the user needs to change their posture [6]. The researchers' "posture correcting chair" must be used with another mobile app. If the user needs to adjust position, this chair will notify the phone [7]. In the final study, the researchers explored the feasibility of using SenseCam image data to determine and estimate the sedentary behavior of adults. The research has demonstrated that the use of SenseCam is possible, but the data collection and procession is time-consuming and also it yields a low level of usable data [8].

Finally, posture is important while using a computer. Most computer users ignore seating positions while working. Researchers have employed force sensors to track users' working postures. They utilized a chair with force sensors. [8].

Basically, the chair acts as an IoT device in this system. And the mounted force sensors are supposed to take force value as a parameter. The parameters that were taken will be fed to a server through the network connection which is connected to the IoT device [8]. The system then processes the data and depending on the input data the system determines if the posture is good or bad. After they will notify the user of the generated output. But this process is not cost-efficient, and they are not following those steps in real-time. Instead of using sensors, some researchers have used multiple cameras to capture the user's posture. They have located multiple cameras around the user capturing the user from different angles. Captured data will be fed to the system and generate the output base on images. This method is more expensive than the other methods.

3. METHODOLOGY

In this system, the main issues that are addressed is the health issues that office employees that use computers face while engaging in the use of computers for prolonged periods of time. This context may be applied in an office environment or a working-from-home environment. Many individuals are stressed due to the prolonged use of computers which leads to many health effects both mental and physical. This stress may be due to overworking or pressure from the employer to work beyond the natural capacity of that individual. Another issue these individuals face is the development of Computer Vision Syndrome (CVS) which is caused by long-term straining of the eye without blinking or looking away from the screen for a long period of time. Getting used to an incorrect posture is something that is most common among individuals using the computer for long hours. Apart from these one of the major issues with working with computers is getting used to sedentary behavior. Due to this sedentary behavior, the hydration patterns of employees using computers are highly affected and most people tend to forget to drink fluids to replenish their systems. All these health will be addressed to mitigate the risks of developing conditions that could be harmful in the long run.

3.1. Stress Detection using Voice Analysis

In this module, the voice of the user is captured using the computer microphone. The audio file is saved in the computer of the user. This audio will be recorded every two minutes for a time period of one minute. Hence, every two minutes the recorded audio file is overridden Once the recording is saved in the device, the audio file is used against a trained MLP Classifier Model for the prediction of stress. This model was trained with two datasets combined into one. The datasets that were used is WESAD (Wearable Stress and Affect Detection) and RAVDESS (Ryerson Audio-Visual Database of Emotional Speech and Song). Once the recorded audio is analyzed for indications of stress, the model will give a prediction of the emotional state of the user whether stressed or not stressed. This emotional state of the user will be recorded in the database against the timestamp thereby allowing the database to have a timestamp and the emotional state of the user. Once this is done the system will then get the emotional state of the user for the past thirty minutes and if the user is found to be stressed 75% of the time, then the system will return to the user a notification notifying them of their stressed behavior. Along with this, some tips for managing stress are also given to the user.

3.2. Detection of factors that lead to Computer Vision Syndrome (CVS)

In this component, the feed of the user is captured by the video camera. After the video feed is captured, it is analyzed to detect the factors that lead to CVS and to identify whether the user is following the 20-20-20 rule. In this component cvzone library is used to access the facial landmark module. The input video is separated into individual frames to apply the facial landmark module. From this facial landmark module, the specific points of the eyes are extracted to read the eye blinks. After identifying the blinks, the blinks are calculated, and the blink rate is identified. The blink rate is a major factor that can predict whether the user is showing signs of CVS since a low blink rate shows that the user is in danger of getting CVS. After calculating and gathering the data, it is then saved in the database. In the front end, the application will fetch the data from the database every 10 seconds and will update the statistical values on the client side. If the rate of eye blinks is lower than the expected value, the system will send a non-interruptive notification to the user suggesting taking a break and relaxing the eyes.

Fig. 1 Highlighted specific facial landmarks

3.3. Detection of Posture using Image Processing

In this module, the feed of the user is captured by the video camera. After capturing the video feed, it is analyzed to detect the posture of the user to identify whether they are sitting in a good posture or a bad posture. This is identified by initially extracting the user's facial data from the pre-trained data set. The trained dataset is extracted as a .pkl file. These facial data points contain three axes which means every point has a 3D view. The extracted data points are then stored in a CSV file and that will be fed into a decision three classification model. This tree structure enables the prediction of data based on the data of the video feed. Hence when the user sits in a good position the component will detect it as a good sitting posture and it will log it as such. Similarly, when the user is sitting in a bad posture, it will detect it as a bad posture, and it will log it accordingly. This data log is saved in the database to be used to display statistical data in the front end. In the front end, the application will fetch the data from the database every 10 seconds and will update the statistical values on the client side. And if the user is sitting with a bad posture, the system will send a non-interruptive notification to the user suggesting correcting the user's posture.

Fig. 2. Posture detection landmarks

3.4. Detection of sedentary behavior

The body movement of the user is captured by the video camera in this module. After capturing the video feed, it is analyzed to detect the body movement of the user. There are multiple pre-defined body movements that are in the classifier. These are natural body movements people make when they have been sitting for a long period of time such as different types of stretches. When the user does one of these body movements, the system identifies this movement and logs it in the database. For body movement detection, Mediapipe was used to identify different body components and landmarks since it provides a very accurate representation of body movements. After calculating different body movements, the system calculates the average number of movements to get the average amount. Apart from that, the system also calculates the time between two movements. This allows the application to recognize the time that has passed after a movement, and this also allows the system to understand whether the user is sedentary for more that the desired time. After every movement, the timer is reset to zero so it can start the time calculation from the beginning. In the front end, the application will fetch the data from the database every 20 seconds and will update the values on the client side. And if the average activity rate is less than the expected rate, the system will send a non-interruptive notification to the user suggesting engaging in some physical activities.

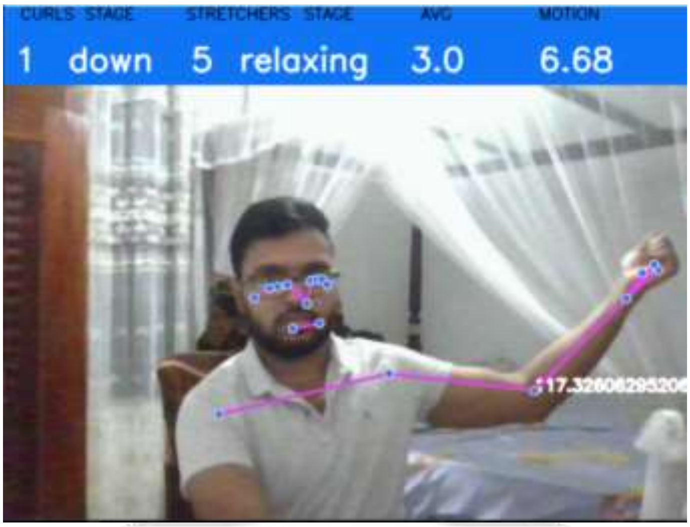

Fig. 3 Detection of sedentary behavior Implementation

3.5. Detection of Hydration pattern

In this component, the user's movements are captured with the video camera. After capturing the feed, it is analyzed to identify whether the user is hydrating themselves. This component uses a combination of several machine-learning libraries to ease the process of identification. The system detects the face, hands, and bottle to identify and register a hydration movement. This component utilizes Mediapipe and cvzone to detect different components and extract landmarks of the user to identify the hydration behavior. Here, the number of times the user rehydrated is calculated and the time between two rehydration detections is calculated and stored in the database. This time is used to identify whether the user is hydrating themselves regularly or whether they are dehydrated. In the front end, the application will fetch the data from the database every 20 seconds and will update the values on the client side. And if the average hydration rate is less than the expected rate, the system will send a non-interruptive notification to the user suggesting rehydrating themselves.

Fig. 4 Detection of hydration pattern Implementation

4. RESULTS AND DISCUSSION

4.1. Stress Detection using Voice Analysis

Stress detection was carried out with the data that was captured and analyzed from the computer microphone. The component uses a MLP Classifier Model to analyze the voice of the user. A MLP classifier is used because it is the most reliable, easiest and most accurate way to find variations in the voice. When the recorded audio is run against the classifier it will detect whether the user's voice is stressed or relaxed based on the characteristics of the user's voice. The component will record the audio of the user for 30 minutes and analyze it to determine whether the user's voice exhibit signs of stress. This is highly effective in a place where the user is speaking often because the system will have plenty of data to analyze and give a highly accurate output. When the system registers the characteristics of a stressed voice, it will send a notification to the user.

4.2. Detection of factors that lead to Computer Vision Syndrome (CVS)

Detection of factors that lead to CVS is identified with the calculation of eye blinks since the eye blink denotes whether the user's eyes are lubricated or whether they are dry. When analyzing the eyes of the user, the analysis was carried out with the video data from the video camera. With the use of cvzone face detection module, the component was able to successfully identify the eyes and detect when the user blinks their eyes. The cvzone library was due to the fact that it has an extensive mapping of the face which includes 468 facial points. Hence it is possible to track the subtle movements of the eyes since there are multiple landmarks to track from. Due to the accuracy of the face detection model, it was possible to get a highly accurate reading from the video camera. Hence it was possible to get a good calculation and a highly accurate blink rate when used. The accuracy could be increased when used with a high-definition camera with good light effects.

4.3. Detection of Posture using Image Processing

The posture detection analysis was carried out with the input video data from the video camera. The component was able to successfully identify the different postures the user was sitting. This identification was done with a combination of Mediapipe and extracted data points that were pre-trained. The trained dataset was able to perform with near-perfect accuracy which in turn detected the user postures very accurately with little to no delay. In this component, Mediapipe was used because it has the capability to easily track the body posture and a trained dataset was used to provide the predefined bad and good postures, so the application is able to refer to them when determining the posture. The application would output a value of 0 if the posture of the user is bad and if the posture is a good posture that was pre-recorded in the database it would output a positive value between 1-10.

4.4. Detection of sedentary behavior

The sedentary behavior was analyzed with the feed from the video camera. The application was able to accurately identify the different movements made by the user and to measure those movements accurately. When running this application with different cameras it was evident that depending on the quality of the camera a better result with higher accuracy could be obtained. However, with a general 720p web camera a highly accurate reading could be obtained. Since Mediapipe is a highly tested and evaluated library, its detection capabilities certainly helped the improvement of accuracy. Mediapipe was used due to the fact that it is capable of detecting body landmarks and calculating different body movements. There was little to no delay when identifying movements and logging them which made this a highly accurate component.

4.5. Detection of hydration pattern

The analysis of the hydration behavior was done with the feed from the video camera. The component is successfully identified when the user is hydrating themselves. The application was able to detect and map the body parts and the objects that were required for hydration. This component too had an accuracy increase when using a higher resolution camera, however, the results were still good with a high level of accuracy when used with a general-purpose web camera. Since this component used both Mediapipe and cvzone it helped the increase in accuracy and detection since they are highly used libraries. Both Mediapipe and cvzone help in identifying and determining specific actions accurately, hence they are the easiest and most reliable libraries to be used for determining body movements and actions.

5. CONCLUSION

After thorough research, it was evident that even though there are systems to monitor the individual aspects of the system there are no systems in existence that provide the total monitoring package provided by Whealth or one that is similar. During the research process, the importance of employee health was identified, and it was directly entangled with the performance and the success of the company. Hence it could be concluded that a system like Whealth is able to improve the health of the employees and in turn, it will have a positive effect on the success and performance of the company. The aspects that are been monitored by this system are the most crucial and sensitive points of health failure in a work environment. Hence active monitoring and correction will help minimize the health risks and will improve the overall health of the employees. It must be mentioned that there are employee monitoring systems in existence, and these came to be quite popular after the pandemic. However, these systems are designed for the benefit of the company rather than the benefit of the employees. Unlike these systems, Whealth is designed for the benefit of the employees. When the employees benefit the company benefit as a result of this. Hence it is always better to improve the benefits of the employee since it will definitely improve the company.

In identifying the posture of the users, the system makes use of the MediaPipe library. This library extracts some of the pre-defined facial points and shoulder points and calculates the summary of the data to determine whether the posture is good or bad. These predefined data were used to train the model with respect to the decision tree classification method. When used in a real-world scenario the system showcased a 96% accuracy in identifying the posture of the users. When the system is running in the production environment, the system will capture the user in real-time and will feed those captured data to the model. Then the model will process the data to determine whether the posture of the user is good or bad. If the user is sitting in a bad posture, the system will notify the user to correct their posture.

6. REFERENCES

[1] F. MJ, W. RL and &. O. N, Computer use and physical inactivity in young adults: Public health perils and potentials of new information technologies, Ann Behav Med 22, 269–275, 2000.

[2] S. Thomée, L. Dellve, A. Härenstam and M. Hagberg, Perceived connections between information and communication technology use and mental symptoms among young adults - a qualitative study, vol. 10, BMC Public Health, 2010, p. 66.

[3] R. D. Arushi and A. N. Teoh, "Real-time Stress Detection Model and Voice Analysis: An Integrated VR-based Game for Training Public Speaking Skills," Vols. pp. 1-4, doi: 10.1109/CoG52621.2021.9618989., 2021.

[4] A. Dementyev and C. Holz, "DualBlink: A Wearable Device to Continuously Detect, Track, and Actuate Blinking For Alleviating Dry Eyes and Computer Vision Syndrome," ACM Journals, vol. 1, no. 1, pp. 1-19, 30 March 2017.

[5] M. T, B. P and Z. F, "Blink Detection for real-time eye tracking," Network and Computer Applications, vol. 25, pp. 129-143, April 2002.

[6] M. P. R. H. Slavomir Matuska, "A Smart System for Sitting Posture Detection Based on Force Sensors and Mobile Application," Mobile Information Systems, vol. 2020, p. 13, 2020.

[7] Premium Waters, Inc, "How Dehydration Impacts Workplace Productivity," 02 08 2017. [Online]. Available: https://premiumwaters.com/blog/dehydration-impacts-workplace-productivity/#:~:text=Just%20one%20percent%20dehydration%20has,percent%20decline%20in%20worker%20productivity.. [Accessed 15 06 2022].

[8] J. Wang, Z. Liu, Y. Wu and J. Yuan, "Learning Actionlet Ensemble for 3D Human Action Recognition," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 36, no. 5, pp. 914-927, 2014.