A Preliminary Study on Poor Posture Warning System for Desk Work

Shuhei Era, Kunio Yamamoto and Masaki Oshita

Kyushu Institute of Technology, 680-4 Kawazu, Izuka-shi, Fukuoka, 820-8502, Japan

Abstract. In this study, we propose a system that uses a depth camera and skeletal tracking to detect poor posture during desk work and alerts users. This approach eliminates challenges associated with sensor attachment and installation, typically faced by conventional systems that rely on multiple sensors, by assessing various postural issues using a single sensor. The system focuses on the alignment of core body parts such as the head and torso and evaluates posture based on misalignment features. An Intel RealSense depth camera and NuiTrack were used for skeletal tracking, whereas an LSTM-based posture classification model was implemented to detect common poor postures, such as slouching, sacral sitting, and scoliosis, from the tracking data. As the camera's field of view is limited to the upper body, the system leverages the time series capability of the LSTM model to effectively detect poor posture even without lower-body data. Initial experiments showed that slouching and sacral sitting could be classified with reasonable accuracy. However, the detection of scoliosis remains challenging. Future work will focus on improving the model with more diverse data and developing a real-time warning system that visually displays a user's core body misalignment.

Keywords: Posture Evaluation. Poor Posture. Skeletal Tracking.

1 Introduction

The increase in deskwork [1] has led to an increase in problems associated with poor posture [2], including forward-leaning postures such as slouching and straight neck syndrome. Figure 1 presents other examples of poor posture in which the body parts tilt forward, backward, or to the sides, such as in lateral scoliosis or sacral sitting. Poor posture can lead to lower back pain and muscular dysfunction [3][4]. According to [5], poor posture is fundamentally defined as a misalignment of body parts along the front-back or side-to-side axes, placing strain on the musculoskeletal system. When a person maintains a posture for an extended period, such as during desk work, their body tends to tense up and seek a more comfortable position. Consequently, as shown in Figure 1, the body parts, particularly those in the trunk region, gradually become misaligned, leading to poor posture. Therefore, there is a demand for systems capable of detecting such poor postures.

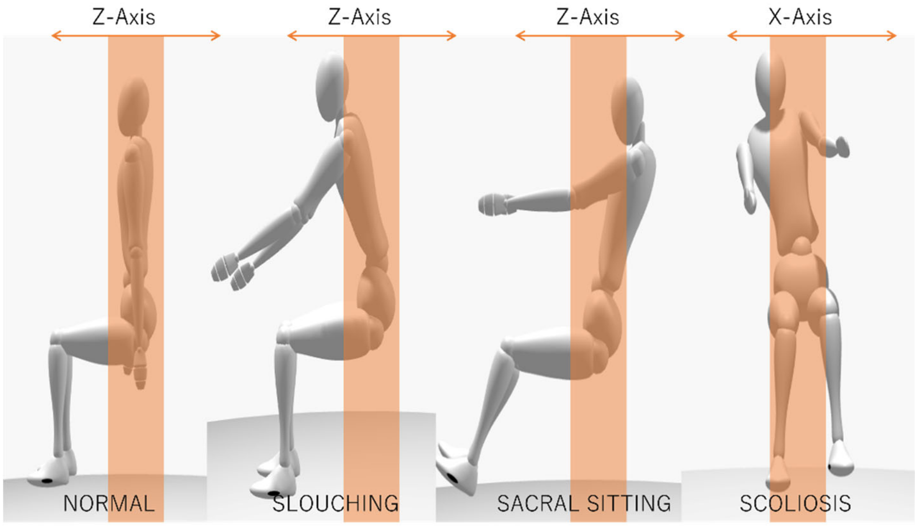

Fig. 1. Poor posture class

For detecting and warning against poor posture, existing systems [6] require attaching multiple accelerometers to the body to estimate the joint angles and evaluate posture. However, these sensors can interfere with work, making them impractical for regular use. Although some approaches [7] involve fewer sensors or avoid attaching them directly to the body, a single accelerometer is typically limited to recording measurements only from a specific part of the body. As a result, these systems are constrained by the estimated number of postures. Moreover, although accelerometers are well suited for determining the orientation of a specific body part, they are not ideal for evaluating spatial relationships between body parts. Further, camera-based methods [8] are limited by the field of view, especially when monitoring seated postures. Additionally, conventional cameras often struggle to capture the spatial relationships between body parts and multiple cameras are required for accurate posture estimation.

In this study, we employed depth camera-based skeletal tracking to capture posture data, offering a more accurate method for detecting body positions than traditional cameras. A depth camera was placed under a PC monitor to track the head, neck, and torso, allowing the evaluation of spatial relationships within the upper body region and supporting the identification of multiple types of poor posture. To evaluate the posture intuitively, we utilized positional differences along certain axes between the head and torso as trunk features. Specifically, as shown in Figure 1, the z-axis represents the forward-backward direction, and the alignment in this direction is evaluated using the head-torso z-axis positional difference v_z, which is a one-dimensional feature. Similarly, the x-axis represents the left-right direction, and the alignment in this direction is assessed using the head-torso x-axis positional difference v_x, which is also a one-dimensional feature. In addition to these two parameters, we employed the head position (three dimensional), torso height (one dimensional), and orientations of the head, neck, and torso (r_x, r_y, r_z), represented by three-dimensional Euler angles for each of the three body parts. In total, 15 parameters were used to evaluate posture. Classification of the four types of poor postures, as shown in Figure 1, was performed using a machine learning model, namely Long Short-Term Memory (LSTM) networks. Our system runs in the background during deskwork. Based on the results of online classification, our system issues alerts when poor posture is detected, providing users with detailed feedback on how their body alignment deviates from the ideal alignment. This feedback promotes effective postural correction. Our goal was to develop a system capable of identifying and addressing multiple classes of poor posture.

Although the prototype classification model used in this study has been developed, a system that issues real-time warnings based on the output of the classification model is yet to be established. The prototype model demonstrated sufficient accuracy for practical use in the three types of postures. However, challenges remain in the classification of scoliosis.

2 Related Work

Mori et al [6]. developed a system in which three accelerometers were directly attached to the body to evaluate posture based on the orientation of the upper body parts. They also proposed a system that provides posture deterioration warnings using sound and confirmed the effectiveness of sound-based posture correction. In our system, we reduce the user burden by using only a single-depth camera and focusing on the relative position of the body parts rather than their orientation.

Watanabe et al [7]. used smartphone sensors to estimate neck angles during smartphone use, specifically evaluating the forward head posture (straight neck). They targeted the neck and assessed poor posture specific to this region. However, considering that multiple types of poor postures can occur during desk work, such as typing or starting at a monitor, our system evaluates three types of poor postures during desk work.

Pinero-Fuentes et al [8]. conducted pose classification using a PC web camera and deep learning. This research evaluates the posture of body parts that are visible on camera, such as shoulders, arms, and neck. Our system also targets sacral sitting.

Tanaka et al [9]. developed a system using helmet-mounted devices to estimate the postures of construction workers. They applied the OWAS method to visual posture assessment, and linked data from accelerometers to posture evaluation. Using these mapped data, they employed LSTM to estimate the postures of civil workers. Although our system differs in providing real-time posture evaluation, we are inspired by the use of sensor data and visual posture assessment with LSTM; however, we focus specifically on desk work rather than physical labor.

Tang et al [10]. used IMUs and LSTM to classify seven postures with high accuracy. This study is similar to ours in that it employed a method for inputting sensing data into LSTM. However, we aim to classify postures using an alternative approach that does not require attaching sensors directly to the body, while also providing warnings. In addition, because we only need to classify the information necessary for users to improve their poor posture, we focus on four specific postures.

Ran et al [11]. classifies posture using a pressure sensor. A warning is given by vibration according to the classification result. We intend to alert users to alignment awareness.

Nadeem et al [12]. summarized recent research on locus recognition. We are focused on creating systems that help users improve their posture more fundamentally by providing posture assessments and alerts that are aware of body alignment.

3 Posture Evaluating and Warnings

3.1 System Overview

In this section, we present a brief overview of the proposed system. Figure 2 illustrates the system flow. In this study, a depth camera was positioned in front of the user for skeleton tracking and obtaining 3D information regarding the position and orientation of the user's joints. However, in this instance, we focus solely on body parts located in the trunk region, such as the head and torso. that can be captured during deskwork. After normalizing the obtained joint positions and orientations, we calculate the positional differences $\mathrm{v_x}$ and v_z between these trunk parts, which we define as trunk features. Both the torso features and posture information were used as inputs to the posture evaluation model. The parameters used for the input include the one-dimensional torso features v_x and v_z, head position $(\mathrm{x_{head}}, \mathrm{y_{head}}, \mathrm{z_{head}})$ (3D), height of the torso $\mathrm{y_{torso}}$ (1D), and orientations $(\mathrm{r_x}, \mathrm{r_y}, \mathrm{r_z})$ of the head, neck, and torso (Euler angles for three parts), totaling 15 parameters. The model evaluates posture based on the input features, and if a poor posture is detected, outputs a posture class such as "slouching" or "sacral sitting." Based on output posture class, the posture-warning UI provides the user with appropriate alerts.

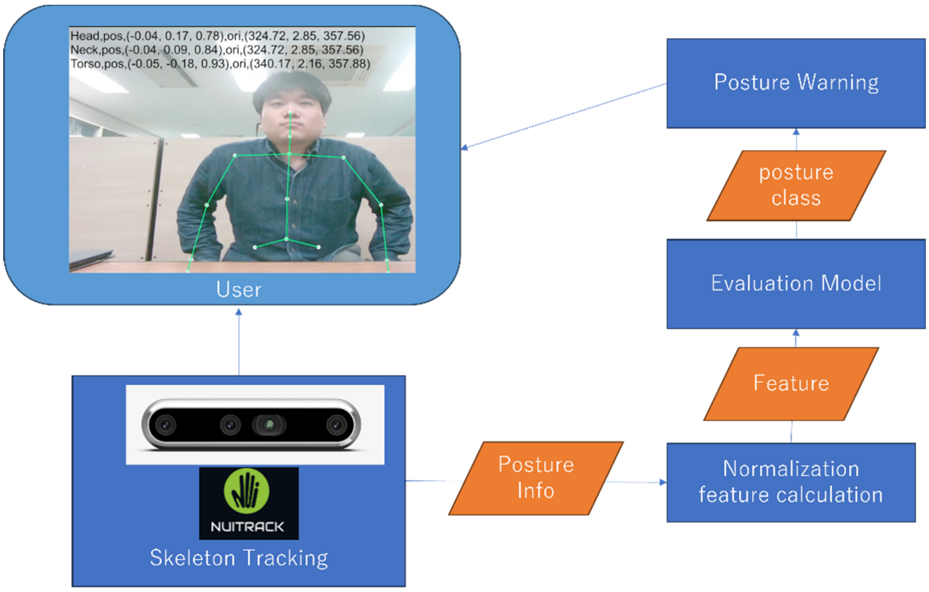

Fig. 2. System Overview

In this study, we utilized the Intel RealSense depth camera. To track the skeleton, we used the NuiTrack library. By capturing the user from the front, as shown in Figure 2, NuiTrack can obtain postural information with sufficient accuracy for postural evaluation. The posture information is obtained in the form of a 3D vector (x, y, z) for the position of each joint point in every frame and a 3D vector in Euler angles (r_x, r_y, r_z) for the orientation. These parameters are expressed in a coordinate system based on the camera, where the z-axis points in the direction facing the user, and the x-axis represents the left-right direction, as shown in Figure 2. In the final system, each posture parameter will be normalized based on the position of the user's head, assuming a normal posture. NuiTrack specifications are limited to shooting from the front. The prototype system draws shoulder and arm skeletons; however, these skeletons are not used.

3.2 Evaluation Based on Trunk Misalignment Features

The four classes of postures, shown in Figure 1, were evaluated with reference to [5], and labeling was performed based on the explanations included in this document. Slouching is a posture in which the back is rounded and the torso and head are slightly tilted downward. Slouching is often associated with forward-leaning postures, such as straight neck syndrome. Lateral scoliosis refers to a posture in which the centers of the head and waist are misaligned, including postures in which the upper body leans to one side or the head is propped up with an elbow. Sacral sitting involves deep leaning against the backrest with the pelvis tilted forward. This posture is common among gamers, and has become a concern in recent years. Correct posture is free of poor posture characteristics, and is defined as one in which the spine is straight and there is no misalignment between the head and waist in either the forward-backward or side-to-side directions.

Understanding the spatial relationships between body parts is crucial for evaluating poor posture. Literature [5] have shown that poor posture occurs when the body parts located in the trunk region, such as the head and abdomen, deviate from the central line. For example, as shown in Figure 1, when maintaining a normal posture, the head and torso are aligned within the trunk region, as depicted in orange. However, in a slouched posture, where the head shifts forward or in the sacral sitting position, where the pelvis moves forward. These body parts deviate from the trunk region, leading to poor posture. Similarly, twisting the waist or raising the elbows can cause a lateral alignment deviation from the trunk. Therefore, the positional differences in the trunk region along the z-axis (anteroposterior direction) and x-axis (lateral direction), expressed as v_x and v_z, respectively, can serve as an effective means for evaluating whether the posture is normal or poor. v_x and v_z are calculated using Equations (1) and (2).

$$v_z = z_{head} - z_{torso} \tag{1}$$

$$v_x = |x_{head} - x_{torso}| \tag{2}$$

Table 1. Input Parameters for Posture Evaluation Model

Parameter	**Description**	**Dim**
v_z	Head-to-torso lateral position difference	1
v_x	Head-to-torso anterior-posterior position difference	1
Head position	3D position of the head (x, y, z)	3
Torso height	Height of the torso (Y-axis)	1
Head orientation	3D Euler angles for head orientation (r_x, r_y, r_z)	3
Neck orientation	3D Euler angles for neck orientation	3
Torso orientation	3D Euler angles for torso orientation	3
Total Dimensions		15

In this study, the parameters v_x and v_z are used as trunk feature quantities and combined with posture information as inputs for evaluating poor posture. Table 1 lists the features used as input for the evaluation model. For posture information, the orientation r_i of three body parts (represented by 3D Euler angles), head position $(x_{head}, y_{head}, z_{head})$ (3D), and torso height y_{torso} are used. These 15 parameters are input into the deep learning model, which performs posture classification. After several iterations of training, features that reduced the accuracy were eliminated, leaving 13 parameters. The model classifies postures into four categories: "slouching," "sacral sitting," "scoliosis," and "normal."

3.3 Evaluation by LSTM

In this study, we considered using a machine learning model, such as LSTM, for posture. classification as the evaluation model. Although evaluating trunk deviations using a simple threshold-based methods allowed for a certain degree of accuracy in assessing postures, such as slouching and lateral scoliosis, the approach has several limitations, the two mains being as follows: First, setting appropriate thresholds for individual users is challenging, and second, some postures are difficult to recognize. In threshold-based models, it is necessary to predefine thresholds based on the body type of the user. For example, when skeleton tracking was performed during desk work with several adult males, some individuals did not appear to slouch unless the trunk regions were misaligned by > 5 cm, whereas the others showed rounded shoulder appearance with only a 3 cm deviation.

Fig. 3. Body tracking from a depth-camera

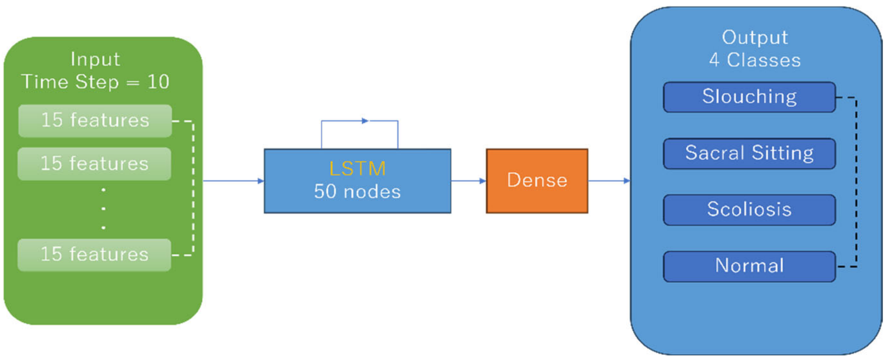

Fig. 4. LSTM networks

Another issue we identified was the accuracy of the sacral sitting classification. Figure 3 shows the sacral sitting position photographed from the front. Given the setup of the depth camera in our method, we avoided using the position and orientation of the hip joints because the waist is not captured directly. In sacral sitting, a certain degree of deviation between the torso and head positions is expected because of leaning back. However, in more severe cases, as illustrated in the figure, the torso and head are nearly vertically aligned, with only the waist protruding forward. Therefore, threshold-based methods used to distinguish the positional differences between the head and torso are insufficient for recognizing sacral sitting. Additionally, increasing the number of thresholds to accommodate all possible postures is impractical, considering individual differences in body types.

To address this issue, we propose using LSTM, which can recognize time-series features, to classify poor posture by leveraging the time-series characteristics of the head gradually being lowered in sacral sitting posture. By inputting both the trunk deviation features and posture information into the LSTM, multiple poor postures can be classified more effectively. We also believe that, with appropriate data preparation and time-series feature recognition, the model can handle classifications regardless of individual differences in body type.

Based on these considerations, we investigated the use of LSTM. Employing LSTM allows the recognition of time-series features associated with sacral sitting, in which the head gradually descends and helps mitigate misclassification of scenarios involving minor body movements. The model predicts poor postures over specified intervals based on posture information and features. Figure 4 shows an overview of the evaluation model. Table 1 lists the 15 features used as inputs for the LSTM. The time step was 10, and the features for 10 frames were input. The LSTM model was configured with 50 nodes and the output layer will perform classification into posture classes.

For constructing the deep learning model, TensorFlow/Keras was used. The training dataset was created by processing data collected from multiple adult males performing deskwork for approximately one hour. The dataset was labeled based on the posture information obtained from depth camera during deskwork and visual assessment of frontal and lateral video footage. The current evaluation model classifies postures into four categories: "slouching," "sacral sitting," "scoliosis," and "normal." During training, the time-step range was set between 10 and 100 to incorporate the time-series features.

3.4 Warning for Poor Posture

After receiving posture class information from the evaluation model, the warning system issues a corresponding alert. In addition to the conventional audio warning method, the system displays a schematic diagram of the identified poor posture.

Moreover, to help users become aware of the cause of their poor posture and correct it, the system displays two lines representing the user's trunk along with an illustration of any body part that deviates from the trunk. This allows users to train themselves to keep their body parts aligned with their trunk while working.

4 Experiment

To validate the posture classification method using trunk deviation features and LSTM, we conducted actual posture classification using the model described in Section 3. The training data comprised approximately 30,000 frames of the work data obtained from a single adult male. Using this model, we performed posture classification on the sequence data containing multiple instances of poor posture, as shown in the top panel of Figure 5 (True Labels). The sequence data to be classified consisted of approximately 4,000 frames and included poor postures, such as slouching, lateral scoliosis, and sacral sitting.

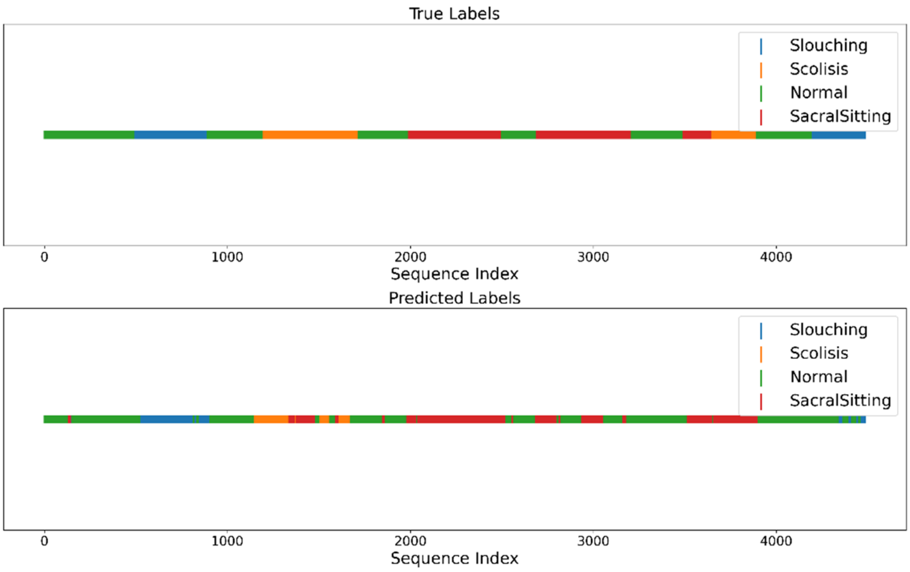

Fig. 5. Comparison: Posture classification with LSTM

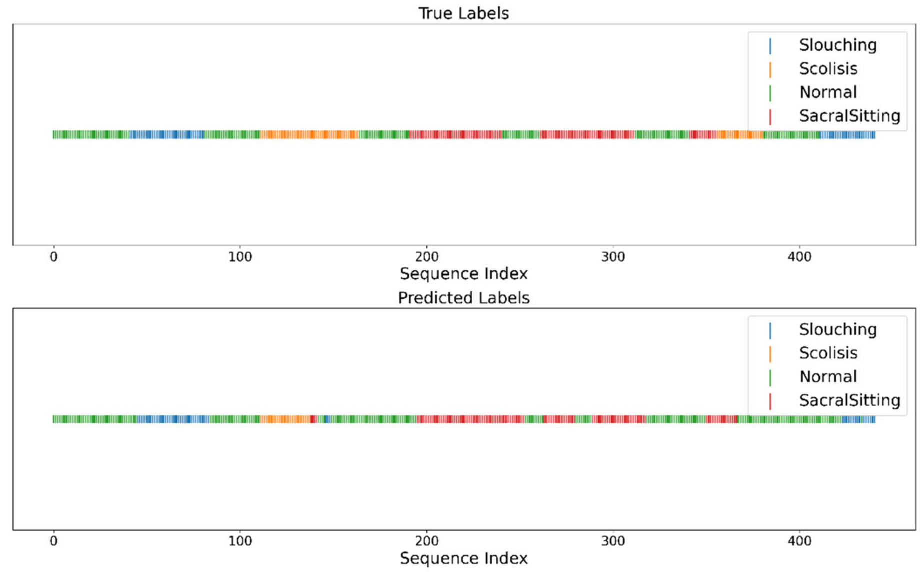

Fig. 6. Comparison: Posture classification with LSTM (frame interval 10 times)

Figures 5 and 6 illustrate the sequence data for repeated poor postures. The top panels show the labels determined by visual inspection, whereas the bottom panels displays the posture classifications using LSTM. Figure 5 shows the classification results with a time step of 10, where the frame intervals remain unchanged. The overall recognition rate was approximately 73% across multiple experiments. Blue and red segments represent the slouching and sacral sitting positions, respectively. Postures within the acceptable range with an upright spine are typically labeled as "Normal". It can be observed that the slouching segments are generally recognized correctly. Please note the yellow label indicating the lateral scoliosis. The evaluation model failed to correctly classify the scoliosis occurring in frame 3700. In addition, some segments of the sacral sitting position, indicated in red, were incorrectly classified, which may have been caused because of the noise introduced by the low time step.

Figure 6 shows the experimental results when the frame interval of the estimated data was adjusted. By increasing the frame interval by a factor of 10, the data were reduced to approximately 400 frames for the posture classification. Because our system aims to evaluate poor postures that persist over long periods, it is unnecessary to focus on postures that occur for minute intervals. Compared to Figure 5, the overall recognition rate improved to 74%, and the occurrence of noise was further reduced. The sacral sitting and terminal phases of slouching were also largely recognized correctly. However, challenges remain regarding accurate diagnosis of lateral scoliosis.

5 Discussion

The proposed model focuses on issuing posture-related warnings, where the primary requirement is not overall accuracy but the ability to continuously predict specific poor postures for extended periods across different classes. The experimental results of our classification model indicate that using the selected features, LSTM can generally predict periods of poor posture, such as slouching or sacral sitting. However, issues regarding the classification of scoliosis were noted. We hypothesize that this is because the scoliosis data comprise less than 10% of the overall dataset.

The features extracted from the camera were primarily limited to the upper body region, excluding the waist because of the frontal viewpoint of the camera. Despite this limitation, the prediction accuracy for sacral sitting is satisfactory. Although the features along the x-axis, given the frontal viewpoint, are expected to be accurate, the evaluation of scoliosis remains problematic.

In this paper, we presented experimental results using data from a single individual. Further validation using data from multiple individuals is required. Although we created a dataset using data from three individuals, the model trained on the combined dataset did not yield favorable results. We believe that this is due to the variability in the dataset, which contains data from individuals who exhibited prolonged periods of slouching and those who mostly maintained the correct posture. This suggests that it is necessary to create a dataset in which the number of samples per label is balanced. Additionally, given the scarcity of scoliosis data, additional scoliosis data may need to be introduced to create a more balanced dataset.

6 Conclusion

We have developed a method to extract trunk feature quantities from posture information captured by a depth camera, as well as an evaluation model capable of classifying four types of poor postures. However, challenges have been identified, such as the inability to capture the waist during desk work owing to the limitations of NuiTrack's viewpoint. However, the use of LSTM has allowed reasonably effective evaluations using measurements from only the positions of the head and torso without relying on the waist. In the future, we plan to explore more effective methods for capturing trunk-related information.

The current evaluation model was somewhat successful in accurately estimating periods of poor posture. Integrating this model into a real-time system could potentially provide effective warnings for postures such as slouching and sacral sitting. However, the evaluation of scoliosis has not been successful and the data used for model training are obtained from a single individual. Therefore, it is necessary to collect data from diverse individuals and retrain the evaluation model.

We are currently developing a system that leverages Unity's Barracuda framework to implement a posture evaluation using LSTM. This development is proceeding in parallel with efforts to improve the accuracy of the model. Simultaneously, we are working on a user interface design to enhance usability and interaction.

Our future plans include gathering more data to refine the model, designing and implementing the warning system, and conducting experiments in which multiple individuals will use an integrated evaluation and warning system. These experiments will assess the accuracy of poor posture recognition, the system's non-intrusiveness during tasks, and the effectiveness of making users aware of trunk misalignment.

References

1. Gould, S.-J., Rudnicka, Anna., Cook, D., Cecchinato, M.-E., Newbold, J.-W., Cox, A.-L.: Remote Work, Work Measurement and The State of Work Research in Human-Centred Computing. Interacting with Computers, Vol.35, Issue.5, pp. 725-734. (2023).
2. Fathima, D., Lobo, J., Angioi, M., Blach, W., Rydzik, L., Ambrozy, T., Malliaropoulos, N.: Sedentary Lifestyle, Heart Rate Variability, and The Influence on Spine Posture in Adults: A Systematic Review Study. Appl. Sci, 14(16), 6985. (2024).
3. Cheung, M., Lai, J., Yip, J., Cheung, J.: Increased Computer Use is Associated with Trunk Asymmetry That Negatively Impacts Health-Related Quality of Life in Early Adolescents. Patient Preference and Adherence, Vol.2021:15, pp. 2289-2302. (2021).
4. Salsali, M., Sheikhhoseini, R., Sayyadi, P.: Association Between Physical Activity and Body Posture: A Systematic Review and Meta-Analysis. BMC Public Health, 23, Article number. 1670. (2023).
5. Takei, H.: Tadasiku Risouteki na Shisei wo Torimodosu: Shisei no Kyokasho (Correcting Ideal Posture: A Textbook on Posture). Natsumesha CO. (2015).
6. Mori, Y., Enokibori, Y., Mase, K.: Support System for Posture Improvement Using Wearable Accelerometers. DICOMO2014, pp. 126-130. (2014).
7. Watanabe, K., Nakamura, S.: A Proposal of A Neck Angle Estimation Method for Posture Correction during Smartphone Use. IPSJ SIG Technical Report Vol.2024-HCI-206, No.25, pp. 1-7. (2024).
8. Pinero-Fuentes, E., Canas-Moreno, S., Rios-Navarro, A., Dominguez-Morales, M. Sevillano, J.-L.: Linares-Barranco, A. A Deep-Learning Based Posture Detection System for Preventing Telework-Related Musculoskeletal Disorders. Sensors, 21, 5236. (2021).
9. Tanaka, R., Yabuki, N., Fukuda, T.: Development of A Workers' Posture Estimation System Based on The OWAS Method Using Helmet-Mounted Terminals. Journal of Japan Society of Civil Engineers, Ser.F3, Vol.77, No.2, I, pp. 114--123. (2021).
10. Tang, H.-Y., Tan, S.-H., Su, T.-Y., Chiang, C.-J., Chen, H.-H.: Upper Body Posture Recognition Using Inertial Sensors and Recurrent Neural Networks. Appl. Sci, 12101, (2021).
11. Ran, X.,Wang, C.,Xiao, Y.,Gao, X.,Zhu, Z., Chen, B.:A Portable Sitting Posture Monitoring System Based on A Pressure Sensor Array and Machine Learning. Sensors and Actuators A: Physical, Vol.331, 112900, ISSN 0924-4247. (2021).
12. Nadeem, M, Elbasi, E, Zreikat, A.-I., Sharsheer, M.: Sitting Posture Recognition Systems: Comprehensive Literature Review and Analysis. Applied Sciences, 14(18), 8557, (2024).