Vision-Based Posture Detection for Rehabilitation Program

Sudhir Gaikwad(✉), Shripad Bhatlawande, Atharva Dusane, Dyuti Bobby, Krushna Durole, and Swati Shilaskar

Department of Electronics & Telecommunication, Vishwakarma Institute of Technology, Pune, India

sg22jn@gmail.com, {atharva.dusane20,dyuti.bobby20, krushna.durole20,swati.shilaskar}@vit.edu

Abstract. Individuals with disabilities frequently struggle to do simple tasks. Recurrent workouts have been demonstrated to aid affected patients in rehabilitation. Physical rehabilitation therapy that can be self-managed provides a convenient solution for people with motor disabilities who may find it challenging to attend regular in-person therapy sessions. Analyzing body postures is instrumental in assisted living and health monitoring at home. Tracking body postures is a profound issue in computer vision. Monitoring the upper-limb posture of the body is the primary goal, while considering the complication of human pose, despite having no publicly available dataset. In this text, a self-data procurement system followed by real-time body posture recognition is implemented using LSTM. The posture classification accuracy is 93.75 percent. If current frames are incorrect, immediate results will be displayed. As a result, the user can instantly improve their posture if they complete their exercises inaccurately, by viewing the correctness of their performance in real-time.

Keywords: Posture Detection · LSTM · Machine Learning · Media Pipe · OpenCV · Rehabilitation · Random Forest

1 Introduction

Motor disabilities often restrict individuals' ability to perform various basic tasks and induce a limited range of motion. Recurrent exercises as part of rehabilitation have been found to allow people with motor disabilities to overcome these issues. However, the bracket of people performing the exercises as recommended is limited to a mere 31% [1]. This results in varied negative outcomes for the affected.

Continuous activity monitoring by medical practitioners/physiotherapists could be used to fight this disregard. However, this process is labor and time - intensive, subject to oversight, and, in some situations, prohibitively expensive [2]. Therefore, home-based physical rehabilitation therapy (HBPT) which can be self-managed is the need of the hour [3, 4].

B. Raj et al. (Eds.): ICETE 2023, AER 223, pp. 473–483, 2023.
https://doi.org/10.2991/978-94-6463-252-1_50

In this work, an exercise posture monitoring system is proposed for individuals during trunk rotations, leaning forward, standing, and shoulder hike activities. Human motion modeling using computer vision (CV) is performed. The accuracy of each posture being done is monitored, with a message displayed for inaccurate posture. This prevents the patient from performing incorrect exercises, thus reaping maximum benefit from the ones performed.

The following paper is divided into multiple sections. The second section presents an overview of related studies in this topic. Section 3 describes the techniques utilized in the study. Section 4 covers the findings of the study and discusses their consequences. Finally, Sect. 5 summarizes the findings and discusses prospective avenues for additional experimentation.

2 Related Work

Posture detection for the ease of rehabilitation to prevent the discomfort of the spinal cord using various technologies and methods is becoming increasingly popular. The NITE skeleton tracking library along with a consumer depth camera are used in [5] for tracking real-time patient data. Upper limb rehabilitation processes are monitored by the proposed system. As the monitoring is done in real-time, the patient can be prompted into the correct pose immediately. An alternative skeletal tracking tool in use is Kinect SDK [6]. 20 points positions are estimated for each frame. Microsoft's Skeleton API has also been employed. The Kinect II sensor can also be used to assess sleep posture [7]. The single depth sensor was employed to acquire selective depth signals under two conditions: with and without a thin blanket, and to curate a dataset. Features were then extracted using fast Fourier transformation and used to prepare a support vector machine (SVM). Skeletal feature extraction can also be performed using the open-source OpenPose [8] library.

Posture detection has also been previously performed on various tasks including body posture recognition on bed [9]. BCG signals were captured in 4 basic postures with average accuracy greater than 97%. The kappa values additionally show that gender and physique have no impact on posture recognition. Distracted Driver Dataset by American University in Cairo was used in [10] to implement posture recognition to classify driver distraction into categories including "talking to passenger" and "reaching behind" with accuracy of 92.7%. CNN, SVM, k-means clustering, and ANN have been utilized by Hasib et al. [11] on the MS-COCO and custom datasets for both human posture recognition as well as fall detection.

The sensors and cameras engaged in posture detection are critical components [12] investigated a method for training neural networks on a Virtex platform. The Weizmann human outline-based database, which contains ten different actions, was used to train the model. Orengo et al. [13] suggested a new approach for sensor profiling that allows users to anticipate their performance in various applications, resulting in a reduction in mean error.

Different models have been using rehabilitation-centered posture detection. Unzila et al. [14] tested a trained network to achieve 0% misclassification of remedial postures. KNN and random gradient descent algorithms have also been used in posture classification during rehabilitation of stroke patients [15]. Exergames have been been used to increase patient enthusiasm in [16] for telerehabilitation. Home based telerehabilitation of stroke patients has also been suggested by Zheng et al. [17] in their review of various position sensing technologies.

Real-time posture monitoring using joint angles generated from the OpenPose library was used to identify high-risk occupational postures [18]. Recent developments can also classify multi-person positions [19]. An algorithm for fall detection centered around body pose estimation aimed at the elderly population is proposed in [20]. The authors review existing fall detection methods and the challenges they face. Neural network for human posture estimation in physical training [21] is proposed covering existing studies on human pose estimation using deep learning techniques such as CNNs with 2D/3D joint positions and RGB/depth images, as well as lightweight networks like MobileNet, ShuffleNet, and SqueezeNet. Authors of the paper [22] review techniques for estimation of pose and using machine learning algorithms. It covers various detection techniques and the challenges they face, as well as different algorithms, including CNNs, RNNs, and DBNs.

Most papers reviewed present limitations as wearable/hardware devices that can hamper movement during exercise. Thus, this paper presents a computer vision-based posture detection structure. Further, the proposed system provides real time classification of postures from each category into correct and incorrect posture aiding in immediate feedback.

3 Methodology

The outline of the research work can be clearly understood with the help of the diagram (see Fig. 1). Key point detection using Media Pipe on dataset collected followed by classification using LSTM and Random Forest were the major techniques undertaken.

3.1 Dataset Collection

The dataset used for the research is custom created by recording videos of individuals exercising. For each action category, videos of 5 people are recorded and merged into a single video from which a total of 1600 frames are extracted. Videos were collected for two classes - correct and incorrect postures. Standing, leaning forward, trunk rotations, and shoulder hikes are activities that are recorded.

The frames were extracted for different postures from the combined dataset and the background was removed by first fetching the pose landmarks from each frame, different masks were created, and the background excluding the pose image was colored gray using MediaPipe and CV2 application.

The image with background and image without background is demonstrated (see Fig. 2).

3.2 Landmarks Marking

33 body posture landmarks and a background segmentation mask are marked using MediaPipe, an open-source library. The landmarks of the body are connected for feature extraction (see Fig. 3).

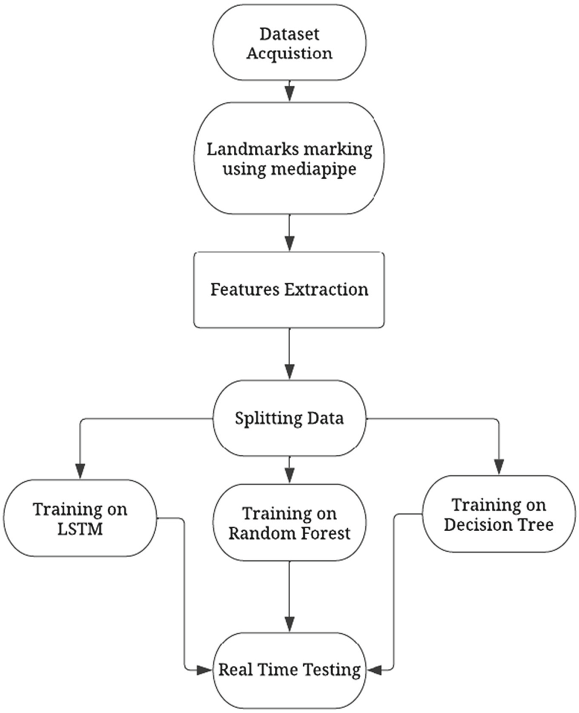

Fig. 1. Outline of work

Fig. 2. Removal of background

3.3 Feature Extraction

The x, y and z coordinates of all landmarks are marked in 3-Dimension with respect to image's height and width. Coordinates x and y are normalized within a range of [0.0, 0.1] corresponding to image's height and width. Coordinate z corresponds to the depth

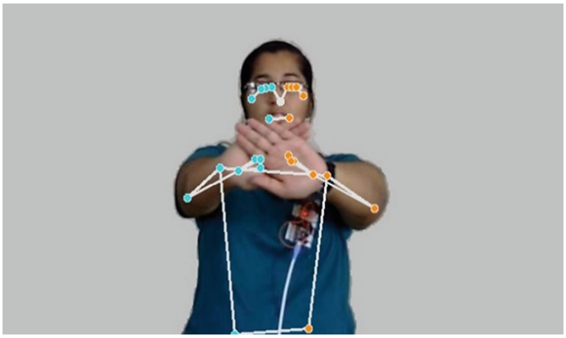

Fig. 3. Landmarks Marking

(a) (b) (c)

Fig. 4. Feature extraction with background **(a)** Original image captured with background; **(b)** Landmarks are detected on the captured image; **(c)** 3-D plot of the landmarks detected.

of landmark where origin was kept at mid of hips. The magnitude of x-coordinate is same as that of z-coordinate.

The list of detected landmarks converted into their original scale and a 3-D plot are illustrated (see Figs. 4 and 5).

3.4 Classification

Depending upon the exercise performed by the user, postures will be classified into the correct and incorrect postures. For the classification two models are trained and tested to obtain good accuracy. If the exercise which the user is performing belongs to standing, lean forward, trunk rotation or shoulder hike then that body posture will be classified as correct one else incorrect one. The percentage of correct and incorrect will be displayed to the user-on-user interface so that user can correct the posture while doing exercises. The models which trained are Decision Trees, LSTM, and Random Forest.

Decision Trees. DT is a Classification Tool that Utilizes a Tree Structure Comprising of Various Decisions and Their Corresponding Outcomes.

Long Short-Term Memory. The Long-Term Dependencies Are Learnt by the LSTM Models Which is a Special Type of RNN.

(a) (b) (c)

Fig. 5. Feature extraction without background **(a)** Original image captured without background; **(b)** Landmarks are detected on the captured image; **(c)** 3-D plot of the landmarks detected.

Random Forest. RF is a Classifier that Computes an Average of Decision Trees Trained on Different Batches of the Dataset.

Algorithm 1 shown underneath deals with the acquisition of data from different subjects.

```
Algorithm 1: Data Acquisition
Start
Setup Holistic model and Drawing utilities.
Input: a path to a folder of First posture.
For correct and incorrect videos in folder Do
     Access the correct/ incorrect posture's video.
     For 40 small videos in correct/incorrect video Do
          Draw landmarks, Extract keypoints into a list.
     End For
End For
Repeat for all posture's folder.
Output: List
End of Algorithm
```

Algorithm 2 describes the labelling of the data acquired and storing of the data from the NumPy files into a temporary variable.

```
Algorithm 2: Label and Data

Start
Initialize 2 empty list → Data= [] and Label= [].
Input: a path to a folder of First posture
For correct/incorrect folder Do
     For 40 folders in correct/incorrect Do
          Initialize an empty list → Temp = [].
          For 40 NumPy files Do
               Load NumPy file and append it into Temp.
          End For
          Append the Temp list into Data list.
          Append the label as 0 and 1.
     End For
End For
Repeat for all posture's folder.
End of Algorithm
```

Algorithm 3 is used for building an LSTM model for correct and incorrect classification of body posture. Here, the data is divided into a ratio of 80:20 and then passed on to the sequential LSTM model for training.

```
Algorithm 3: LSTM Model Building
Start
Data [] ← NumPy array and store.
Labels [] ← Categorical values and store.
Split the data in the ratio 80:20.
Create a Sequential model's object, Add Layers.
Dense layer with 2 neurons.
Compile the model with optimizer, loss, and metrics.
Fit the data.
End of Algorithm
```

The system user flow for the interface is demonstrated (see Fig. 6).

4 Results and Discussions

Initially, the image dataset created was resized to 100x100 frames in grayscale. Denoising was performed on these frames using the Gaussian Blur filter. As part of enhancement, edges were detected with the Prewitt operator. Features were then described using the BRISK and FREAK constructors. K-means clustering was performed on the features and PCA was used for Diminishing the dimensions. Classifiers DT and RF were trained and tested on this data achieving results (see Table 1). All values are in percentage.

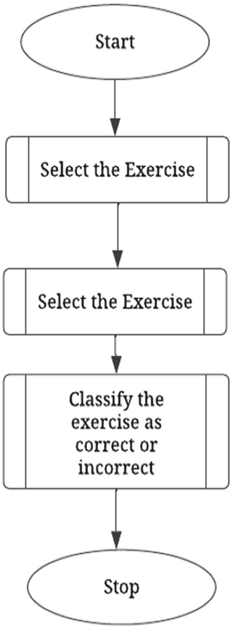

Fig. 6. User flow of the system

Table 1. Decision tree, LSTM, and Random Forest results

	Accuracy	Precision	Recall	F1
Decision Trees	85.50	85.32	85.41	85.50
LSTM	93.75	100	87.50	93.33
Random Forest	94.25	93.65	94.25	94.67

However, this approach does not yield accurate results in real time deployment which was the primary goal of this work.

As a result, RNN (Recurrent Neural Network) is used, which uses sequential or time series data. RNN is trained with train data and tested in real time. The idea behind an RNN is to store a particular layer's output and feed it back to the input to predict the outcome of that layer.

In this, the data was trained on the RNN but drawback of RNN is not being able to connect or use its previous data to preprocess with current frames. In this classification, the model should relate the current frames with previous frames because body posture classification depends on long term frames. Hence here RNN fails. Long-term dependencies can be learned using LSTMs, a special kind of RNN. After training the data on LSTM model we calculated the testing accuracy to be 87.5% (see Table 1).

The final system is user driven, where the individual can select actions from "trunk rotation", lean "forward", "shoulder hike", and "standing".

The scenario where the user performs the exercise as optimally required is shown in figure, the percentage of exercise being done correctly increases, with the green color on the "correct" label increasing. The display on the frame is "correct" if the individual continues performing the action perfectly. If any of the frames contains an incorrect posture for any of the exercises, "incorrect" is immediately displayed (see Fig. 7). The blue color on the "incorrect" label starts increasing. Thus, the user can view the accuracy of their performed exercises in real time and correct their posture immediately if incorrectly performed.

A comparison of the results for RF, DT and LSTM models is demonstrated (see Fig. 8).

(a) (b)

Fig. 7. Real time deployment **(a)** Correct Posture; **(b)** Incorrect Posture

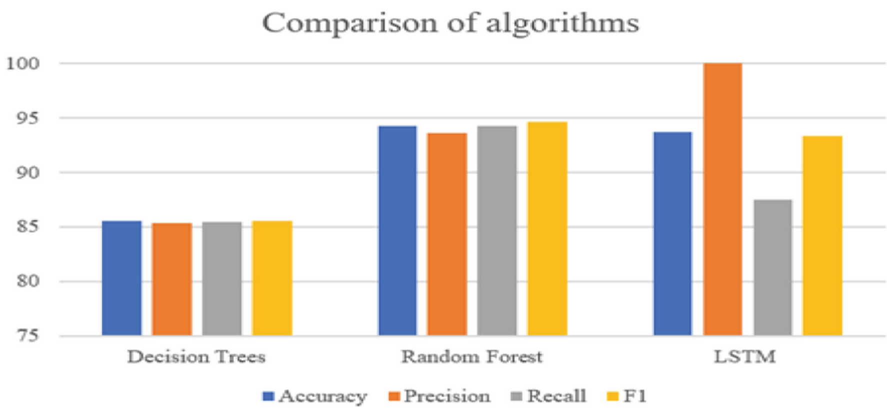

Fig. 8. Results comparison

5 Conclusion and Future Scope

The main emphasis of the study is on the unmanned estimation of posture during rehabilitation therapy for four different exercises viz., trunk rotation, lean forward, shoulder hike, and standing. The developed system employs a real-time body landmark tracking library for a prototype capable of recognizing correct and incorrect body postures in real-time, with good accuracy. The system would prompt the user if incorrect posture were recognized in real time for effective rehabilitation. Future work will focus on combining all different postures important for rehabilitation programs.

Acknowledgement. We are immensely grateful to La Fondation Dassault Systèmes for providing technical support to conduct this research work.

References

1. Chang, Y.-J., Chen, S.-F., & Huang, J.-D. A Kinect-based system for physical rehabilitation: A pilot study for young adults with motor disabilities. Research in developmental disabilities 32(6), 2566–2570 (2011).
2. Debnath, B., O'Brien, M., Yamaguchi, M., & Behera, A. A review of computer vision-based approaches for physical rehabilitation and assessment. Multimedia Systems, 1–31 (2021).
3. Essery, R., Geraghty, A. W. A., Kirby, S., & Yardley, L. Predictors of adherence to home-based physical therapies: a systematic review. Disability and Rehabilitation, 39(6), 519–534 (2017).
4. Bassett, S. F. The assessment of patient adherence to physiotherapy rehabilitation. New Zealand Journal of Physiotherapy, 31(2), 60–66 (2003).
5. Xu, Y., Chen, J., Yang, Q., & Guo, Q. Human posture recognition and fall detection using Kinect V2 camera. In: 2019 Chinese Control Conference (CCC), pp. 8488–8493. IEEE (2019).
6. Maryam, S. R. D., & Payandeh, S. A novel human posture estimation using single depth image from Kinect v2 sensor. In: 2018 Annual IEEE International Systems Conference (SysCon), pp. 1–7. IEEE (2018).
7. Le, T. L., & Nguyen, M. Q. Human posture recognition using human skeleton provided by Kinect. In: 2013 international conference on computing, management and telecommunications (ComManTel), pp. 340–345. IEEE (2013).
8. Ghazal, S., & Khan, U. S. Human posture classification using skeleton information. In: 2018 international conference on computing, mathematics and engineering technologies (iCoMET), pp. 1–4. IEEE (2018).
9. Liu, M., & Ye, S. A novel body posture recognition system on bed. In 2018 IEEE 3rd International Conference on Signal and Image Processing (ICSIP), pp. 38–42. IEEE (2018).
10. Mase, J. M., Chapman, P., Figueredo, G. P., & Torres Torres, M. A hybrid deep learning approach for driver distraction detection. In: 2020 International Conference on Information and Communication Technology Convergence (ICTC), pp. 1–6. IEEE (2020).
11. Hasib, R., Khan, K. N., Yu, M., & Khan, M. S. Vision-based human posture classification and fall detection using convolutional neural network. In: 2021 International Conference on Artificial Intelligence (ICAI), pp. 74–79. IEEE (2021).
12. Desai, S. J., Shoaib, M., & Raychowdhury, A. An ultra-low power, "always-on" camera front-end for posture detection in body worn cameras using restricted boltzman machines. IEEE Transactions on Multi-Scale Computing Systems, 1(4), 187–194 (2015).

13. Orengo, G., Lagati, A., & Saggio, G. Modeling wearable bend sensor behavior for human motion capture. IEEE Sensors Journal, 14(7), 2307–2316 (2014).
14. Jawed, U., Mazhar, A., Altaf, F., Rehman, A., Shams, S., & Asghar, A. Rehabilitation posture correction using neural network. In: 2019 4th International Conference on Emerging Trends in Engineering, Sciences and Technology (ICEEST), pp. 1–5. IEEE (2019).
15. Yu, X., Xiao, B., Tian, Y., Wu, Z., Liu, Q., Wang, J., Sun, M., & Liu, X. A control and posture recognition strategy for upper-limb rehabilitation of stroke patients. Wireless Communications and Mobile Computing, 2021.
16. Rosique, F., Losilla, F., & Navarro, P. J. Applying vision-based pose estimation in a telerehabilitation application. Applied Sciences, 11(19), 9132 (2021).
17. Zheng, H., Black, N. D., & Harris, N. D. Position-sensing technologies for movement analysis in stroke rehabilitation. Medical and Biological Engineering and Computing, 43(4), 413–420 (2005).
18. Lin, P.-C., Chen, Y.-J., Chen, W.-S., & Lee, Y.-J. Automatic real-time occupational posture evaluation and select corresponding ergonomic assessments. Scientific Reports, 12(1), 1–9 (2022).
19. Iqbal, U., & Gall, J. Multi-person pose estimation with local joint-to-person associations. In: European conference on computer vision, pp. 627–642. Springer, Cham (2016).
20. Sun, G., & Wang, Z. Fall detection algorithm for the elderly based on human posture estimation. In: 2020 Asia-Pacific Conference on Image Processing, Electronics and Computers (IPEC), pp. 172–176. IEEE (2020).
21. Sun, Y., Li, G., & Zhao, T. Research on Lightweight Network of Human Posture Estimation for Physical Training. In: 2022 4th International Conference on Robotics and Computer Vision (ICRCV), pp. 62–67. IEEE (2022).
22. Gupta, P. A Review on Posture Detection and Correction using Machine learning and Deep learning. In: 2022 Second International Conference on Next Generation Intelligent Systems (ICNGIS), pp. 1–5. IEEE (2022).