Motion Analysis for Physical Therapy Using Deep Learning

# Gaurang Kalappa Kodimanianda

Roll Number: 20

Registration Number: 220962067

Phone: +91 9663760423

Email: [gaurangkalappa4@gmail.com](mailto:gaurangkalappa4@gmail.com)

# Sanman Pattnaik

Roll Number: 27

Registration Number: 220962109

Phone: +91 7735955112

Email: [sanmanp24@gmail.com](mailto:sanmanp24@gmail.com)

# Ammapuram Sriharsha

Roll Number: 9

Registration Number: 220962019

Phone: +91 9063363624

Email: [asriharsha250305@gmail.com](mailto:asriharsha250305@gmail.com)

***Abstract*—Skin diseases are a widespread health concern, with early detection and classification critical for effective treatment and management. This project aims to develop a computer vision approach to classifying four skin diseases namely basal cell carcinoma, benign keratosis-like lesions, melanocytic nevi, and melanoma—using a machine learning (ML) pipeline that incorporates image preprocessing, data augmentation, feature extraction, and ensemble classification models. The methodology emphasizes preprocessing techniques, including adaptive thresh- olding, grayscale conversion, and sharpening, are employed to enhance lesion features.Augmentation techniques are applied to diversify the dataset and improve model generalization.Feature extraction leverages local binary patterns (LBP), LAB color histograms, and Gray-Level Co-occurrence Matrix (GLCM) texture descriptors. A voting classifier combining Random Forest and Gradient Boosting classifiers is utilized for classification and optimized through parameter tuning. In the conclusion, a com- parative analysis with a deep learning (DL) approach is provided, discussing the respective advantages and limitations, and suggest future directions for optimizing skin disease classification.**

***Index Terms*—pose classification, repetition counting, physio- therapy, deep learning, mediapipe, PyTorch, real-time monitor- ing, human activity recognition, computer vision, rehabilitation exercises**

1. Introduction

Physical therapy is a cornerstone of recovery and rehabil- itation for people with musculoskeletal injuries, neurological disorders, or post-surgical conditions. It empowers patients to regain mobility, improve strength, and restore normal function. A critical element of physical therapy is the accurate execution of prescribed exercises, as improper movements can delay recovery, worsen existing conditions, or even lead to new injuries. However, without the continuous supervision of a physiotherapist, ensuring that patients perform their exercises correctly can be a significant challenge. This gap in supervi- sion often results in reduced adherence to therapy regimens, limiting their effectiveness.

This project aims to bridge that gap by developing a deep learning-based motion analysis system designed to assist patients in performing their exercises accurately and safely. By leveraging cutting-edge advancements in artificial intelligence, particularly in the domains of computer vision and time-series analysis, the system will monitor human motion, identify deviations from correct exercise patterns, and provide real- time feedback to patients. This can significantly enhance the

quality of rehabilitation, minimize risks, and improve recovery outcomes.

Recent innovations in deep learning have unlocked re- markable capabilities for understanding and analyzing human motion. Techniques like pose estimation and activity recogni- tion, which were once considered highly complex, are now accessible through advanced models and algorithms. Pose estimation, for example, allows for precise tracking of human joints and body movements, while activity recognition enables systems to classify and assess complex actions over time. By integrating these techniques, the proposed system will not only detect improper movements but also guide patients toward correcting them in real time.

To achieve this, the project will leverage publicly available datasets such as Human Activity Recognition (HAR) and PoseTrack. These datasets provide extensive, high-quality data on human motion, which can be used to train and validate the deep learning models. The combination of HAR’s focus on activity classification and PoseTrack’s emphasis on detailed pose tracking will enable the system to deliver comprehensive and accurate motion analysis.

Ultimately, this project is a step toward democratizing access to quality physical therapy. By providing patients with an intelligent, automated tool that supports their rehabilitation journey, we can reduce the dependency on constant physio- therapist supervision while maintaining high standards of care. This innovative solution has the potential to transform the field of physical therapy, making recovery more accessible, efficient, and effective for patients everywhere.

1. Literature Review

Recent advancements in deep learning and computer vision have significantly contributed to the field of motion analysis for physical therapy. Various studies have explored the appli- cation of deep learning models for pose estimation, movement assessment, and real-time feedback systems to assist in reha- bilitation exercises.

Mun˜oz-Salinas et al. introduced the UCO Physical Reha- bilitation dataset, addressing the scarcity of publicly available datasets for rehabilitation exercises. Their dataset contains 2,160 videos of patients performing prescribed exercises, and

they evaluated state-of-the-art pose estimation methods in dif- ferent positions. The results demonstrated that most methods performed well for upright positions but failed when the subject was lying down, emphasizing the need for specialized pose estimation techniques [1].

Vakanski et al. proposed a deep learning-based framework for evaluating rehabilitation exercises. The model includes three components: movement performance metrics, scoring functions for movement quality, and deep neural networks for supervised learning-based assessments. Their framework demonstrated effective rehabilitation performance assessment through a novel deep spatio-temporal neural network that organizes data into temporal pyramids and processes joint displacements [2].

Kulkarni et al. developed a physiotherapy assessment sys- tem that utilizes deep learning to enhance movement evalu- ation. Their approach uses OpenCV to extract video frames and MediaPipe to identify key body points, connecting them to compute joint angles. The system provides real-time au- dio feedback and tracks patient progress through visualized improvement reports, offering an efficient and cost-effective solution [3].

Yang et al. conducted a systematic review of machine learning-based computer vision techniques for physiotherapy movement assessment using depth cameras. They analyzed 18 studies from 2020 to 2024, identifying key implementation scenarios and technological challenges, such as limited real- world validation and dataset diversity [4].

Wang et al. explored how deep neural networks recognize and interpret motion for rehabilitation applications. Their research introduced a vector network (VecNet) combined with an LSTM model to analyze relative positional changes. Their approach supports motion recognition and prediction, enabling more accurate assessment of patient progress and rehabilitation outcomes [5].

Ben-Younes et al. conducted a review on deep learning- based pose estimation for movement feedback. They analyzed 45 research papers, highlighting the dominance of convolu- tional neural networks (CNNs) in pose estimation and the need for improved feedback accuracy. Their findings emphasized the importance of refining movement assessment algorithms for enhanced rehabilitation [6].

Mennella et al. developed a cost-effective system for remote rehabilitation assessment, focusing on motion classification and compensatory pattern recognition. Their system achieved high accuracy in recognizing rehabilitation exercises and has potential applications in unsupervised rehabilitation settings [7].

Lu et al. explored pose detection techniques in gait analysis, leveraging deep learning for improved movement tracking. Their study emphasized the importance of deep learning models in monitoring walking patterns, particularly for reha- bilitation applications [8].

He et al. introduced the Residual Neural Network (ResNet), which addressed the vanishing gradient issue in deep networks by incorporating residual connections. This model has since in-

fluenced numerous deep learning applications, including pose estimation and motion tracking for rehabilitation purposes [9]. Redmon et al. developed YOLO (You Only Look Once),

a real-time object detection system that processes entire im- ages in a single pass, significantly enhancing efficiency. This method has been adapted for real-time pose estimation and motion analysis in rehabilitation settings [10].

Wang et al. proposed a depth-aware pose estimation tech- nique for exoskeleton gait analysis, improving motion track- ing accuracy in human-machine integration. Their approach enables precise assessment of movement patterns in rehabili- tation exercises [11].

In summary, the reviewed studies highlight the critical role of deep learning in motion analysis for physical therapy. Advances in pose estimation, real-time feedback mechanisms, and deep neural networks continue to enhance rehabilitation outcomes by improving accuracy, efficiency, and accessibility.

1. Research Gaps and Objectives
2. *Research Gaps*
   * Lack of affordable and accessible systems for remote physical therapy.
   * Limited integration of sensor-based data with video-based pose estimation for a comprehensive motion analysis.
   * Insufficient capability in existing models to detect minor deviations in posture or movement accurately.Variations in lighting, clothing, and patient diversity can signifi- cantly impact the accuracy of pose estimation.
   * Difficulty in deploying real-time solutions on consumer- grade devices due to high computational requirements.
   * Lack of diverse and comprehensive datasets that ade- quately capture the full range of human motion variations encountered in rehabilitation scenarios.
   * Despite progress, the application of these technologies in clinical practice requires validation in real-world settings. Developing user-friendly and accessible interfaces for physiotherapists and patients is essential.
3. *Objectives*
   * Develop a deep learning-based system that integrates video-based pose estimation with sensor-based activity recognition for precise motion analysis.
   * Train models using publicly available datasets, such as Human Activity Recognition (HAR) and PoseTrack, to accurately detect improper postures.
   * Design a real-time feedback mechanism to guide patients during physical therapy exercises.
   * Create a lightweight system optimized for affordability and accessibility, especially in remote or underserved areas.
4. Methodology

The proposed system aims to monitor and classify ex- ercise postures in real-time, providing valuable feedback to users for enhancing their exercise form. The methodology

follows an end-to-end pipeline consisting of Data Acquisition and Annotation, Data Preprocessing, Model Architecture and Training, and Real-Time Inference and Feedback. The detailed methodology is outlined as follows:

## Data Acquisition and Annotation

* 1. Pose Detection with MediaPipe
     + Exercise Selection and Labeling: The user selects a specific exercise (e.g., Shoulder Abduction, Elbow Flexion, or Standing Knee Raise) from the interface. The user is then prompted to perform the exercise while the system records the pose data.
     + Data Logging: During each exercise session, pose landmarks are continuously recorded and labeled in real-time. The data is saved in CSV format, where each row corresponds to a frame and contains the 132 pose features (4 features per landmark × 33 landmarks). Additionally, metadata is saved, including the exercise type (e.g., Shoulder Abduction), the frame’s timestamp, and whether the performance is labeled as ”correct” or ”incorrect” based on the user’s posture. The annotation process is user-guided with visual cues, such as real- time feedback for correction.
  2. Key Exercise Types
     + Shoulder Abduction: Movement of the arm away from the body.
     + Elbow Flexion: Bending and straightening of the elbow joint.
     + Standing Knee Raise: Raising one leg at a time while maintaining balance.

## Data Preprocessing

* 1. Data Preprocessing
     + Normalization: Each frame’s landmarks are normalized based on the distance between the left and right shoul- der to ensure consistency across various body sizes and camera positions. This technique helps mitigate spatial variability across different individuals performing the same exercises.
     + Standardization: After normalization, all features (land- mark coordinates) are standardized to zero mean and unit variance using the StandardScaler from scikit- learn. This ensures that the model is not biased by differences in feature scales and improves training performance.
  2. Data Augmentation

To increase the robustness of the model and prevent overfitting, data augmentation techniques were applied. These included:

* + - Horizontal flipping: To simulate different orientations of the body.
    - Random rotations and translations: These transforma- tions help simulate slight changes in posture and view.
    - Scaling: Slight random zooming of the body to handle different distances from the camera.
  1. Data Splitting

The dataset is split into training and validation sets, with a ratio of 80:20. To ensure balanced class distribution, stratified sampling was used. This approach helps prevent bias in the training process by ensuring that both ”correct” and ”incorrect” labels are well-represented in each set.

## Model Architecture and Training

* 1. Model Design

The system uses a fully connected feedforward neural network (FNN) to classify each frame into one of two categories: correct or incorrect posture. The architecture consists of:

* + - Input Layer: 132 neurons corresponding to the 33 landmarks × 4 features (x, y, z, visibility).
    - Hidden Layers:
      * First hidden layer: 128 neurons with ReLU activa- tion function.
      * Second hidden layer: 64 neurons with ReLU acti- vation function.
      * Dropout layers (0.3 and 0.2) are added after each hidden layer to mitigate overfitting.
    - Output Layer: A single neuron with a sigmoid activa- tion function to output a probability value between 0 and 1, where values closer to 1 correspond to correct posture and values closer to 0 correspond to incorrect posture.
  1. Model Training

The training procedure follows these steps:

* + - Loss Function: Binary cross-entropy is used as the loss function since the classification task is binary (correct vs. incorrect posture).
    - Optimizer: Adam optimizer was chosen for efficient gradient-based optimization.
    - Epochs and Early Stopping: The model was trained for a maximum of 50 epochs, with early stopping criteria to halt training if validation accuracy does not improve for 10 consecutive epochs, preventing overfitting.
    - Evaluation Metrics: Accuracy, precision and F1-score are computed to evaluate model performance during training and validation.

## Real-Time Inference and Feedback System

* 1. Pose Detection and Classification

For real-time inference, a continuous video stream from the webcam is processed. In each frame, the pose land- marks are extracted using MediaPipe, and the feature vector is constructed by extracting the 132 values cor- responding to the x, y, z, and visibility coordinates of each landmark.

* + - The feature vector is passed through the trained model to predict whether the posture is correct or incorrect.
  1. Repetition Counting

The repetition counting system is built on the principle of detecting significant state transitions in the exercise motion. For example:

* + - Elbow Flexion: Repetitions are counted by detect- ing transitions between the ”elbow bent” and ”elbow straightened” positions.
    - Shoulder Abduction: Repetitions are detected based on the arm movement crossing a threshold angle.
    - Standing Knee Raise: The system tracks the knee’s vertical motion and counts the number of times the knee reaches a certain height.

Each valid repetition is logged and displayed in real time on the user interface.

* 1. User Interface and Feedback

A user-friendly interface was developed using Tkinter, where users can:

* + - Select the exercise they wish to perform.
    - View real-time feedback, such as the correctness of the posture and the number of repetitions completed.
    - Track performance over time with visual markers (such as a green tick or red cross for correctness).
    - Save session data, including timestamps and feedback, for further analysis.
  1. The system also provides visual feedback by drawing the user’s pose skeleton on the screen. Posture analysis results are overlaid on this skeleton, showing whether the exercise is being performed correctly.

1. Results

## Dataset Collection and Preprocessing

A custom dataset was constructed using webcam-based data acquisition, capturing three rehabilitation exercises: shoulder abduction, elbow flexion, and standing knee raise. For each exercise, data were recorded for both correct and incorrect posture classes. The dataset comprises 33 pose landmarks per frame, with each landmark represented by x,y,z and visibility attributes, resulting in 132 features per sample. Data normal- ization was performed by centering all landmarks with respect to the midpoint between the left and right shoulders, enhancing invariance to camera position and subject orientation.

The dataset was partitioned into training and validation sets using an 80:20 stratified split to maintain class balance. No separate test set was used; model evaluation was conducted on the validation set.

## Dataset Collection and Preprocessing

A feedforward neural network was developed for posture classification, consisting of three fully connected layers with ReLU activation and dropout regularization. The network architecture is as follows: an input layer of 132 units, two hidden layers of 128 and 64 units respectively, and an output layer with a sigmoid activation for binary classification. The model was trained using the Adam optimizer and binary cross- entropy loss. Early stopping was applied based on validation accuracy with a patience of 10 epochs to prevent overfitting.

## Training Performance

Figures 1, 2, and 3 present the training and validation performance metrics—loss, accuracy, F1 score, and preci- sion—across epochs for each exercise: standing knee raise, elbow flexion, and shoulder abduction, respectively.

* 1. Standing Knee Raise

As shown in Figure 1, both training and validation loss decreased rapidly within the first few epochs and stabi- lized at low values, indicating effective optimization. The accuracy curves (top right) demonstrate that the model achieved high and stable classification accuracy (*>* 99%) on both training and validation sets after approximately five epochs. The F1 score and precision (bottom row) also improved sharply in the initial epochs, plateauing above 0.98, which reflects strong balance between precision and recall and reliable detection of correct and incorrect postures. The close alignment of training and validation curves across all metrics suggests minimal overfitting.

* 1. Elbow Flexion

Figure 2 illustrates a similar trend for the elbow flex- ion exercise. The loss curves (top left) show a steep decline during the first five epochs, with both training and validation loss converging to near zero. Accuracy (top right) consistently exceeds 99% after epoch five, and both F1 score and precision (bottom row) remain above

0.98 throughout training. Occasional minor fluctuations in precision do not impact the overall high performance, and the close match between training and validation metrics indicates robust generalization.

* 1. Shoulder Abduction

In Figure 3, the shoulder abduction model also demon- strates effective learning. Training and validation loss decrease steadily and remain closely aligned, while accu- racy (top right) improves rapidly and stabilizes near 99%. The F1 score and precision (bottom row) both plateau above 0.98, confirming that the model maintains a strong balance between sensitivity and specificity. The consistent overlap between training and validation curves across all metrics further confirms the absence of overfitting.

Across all three exercises, the model converges quickly, achieving high accuracy, F1 score, and precision on both training and validation sets. The close correspondence between training and validation curves for all metrics demonstrates strong generalization and reliable performance for posture classification in physical therapy exercises.

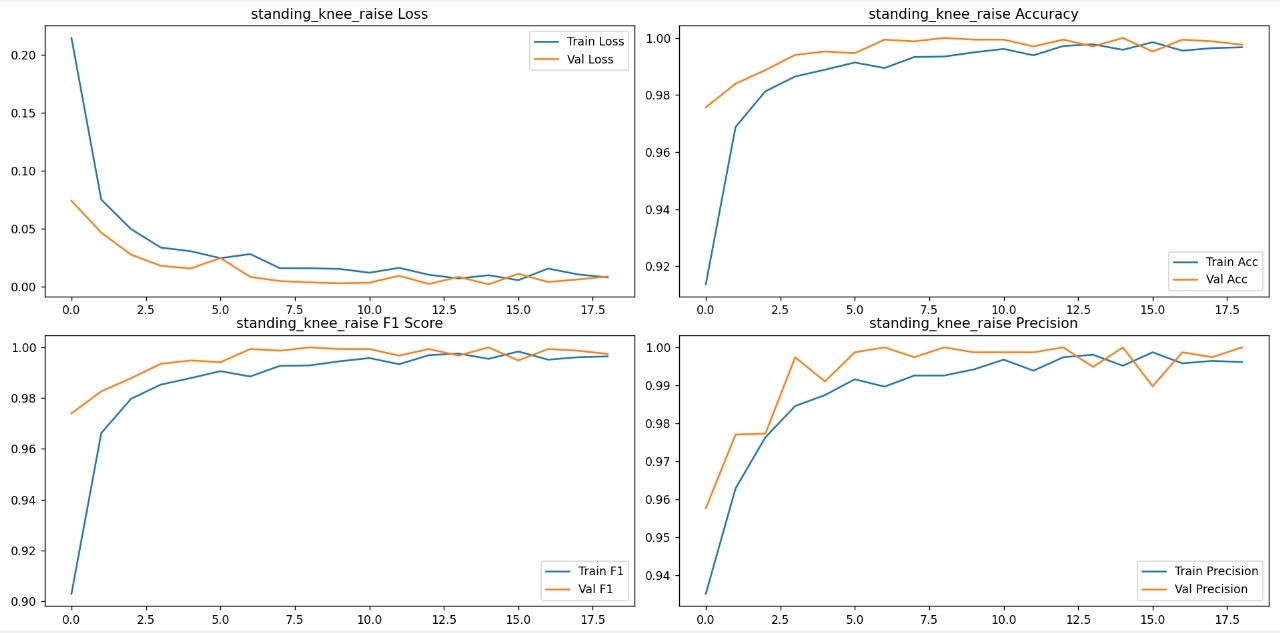


Fig. 1. Standing knee raise—loss, accuracy, F1 score, and precision curves

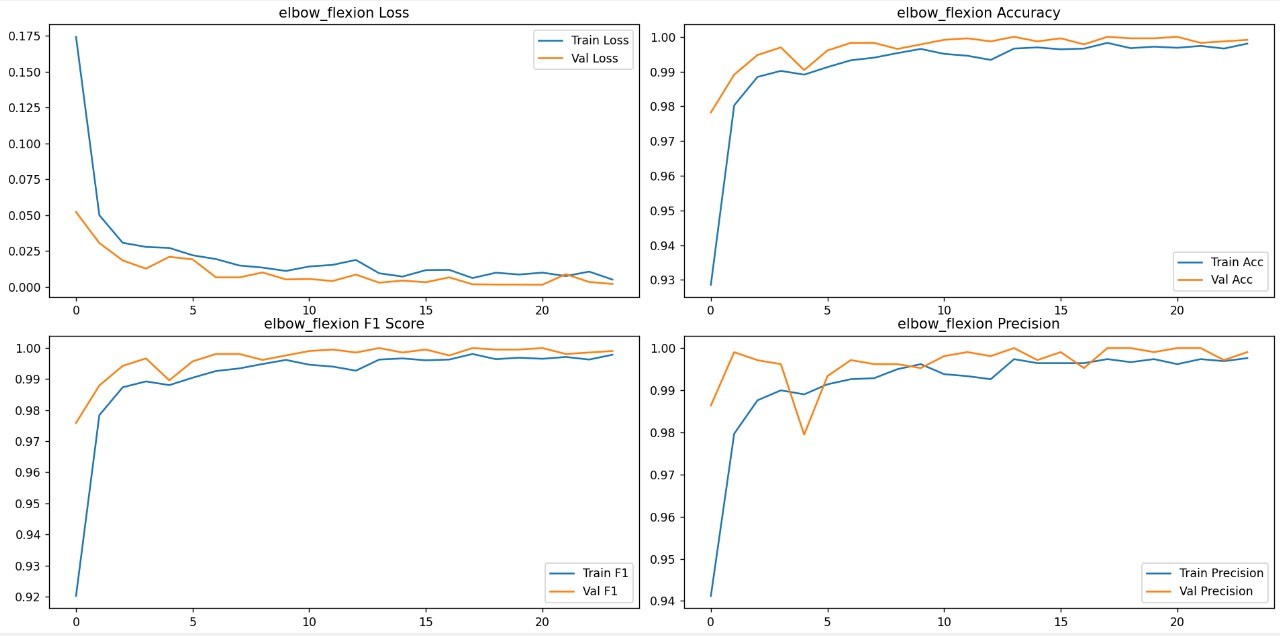


Fig. 2. Elbow flexion—loss, accuracy, F1 score, and precision curves

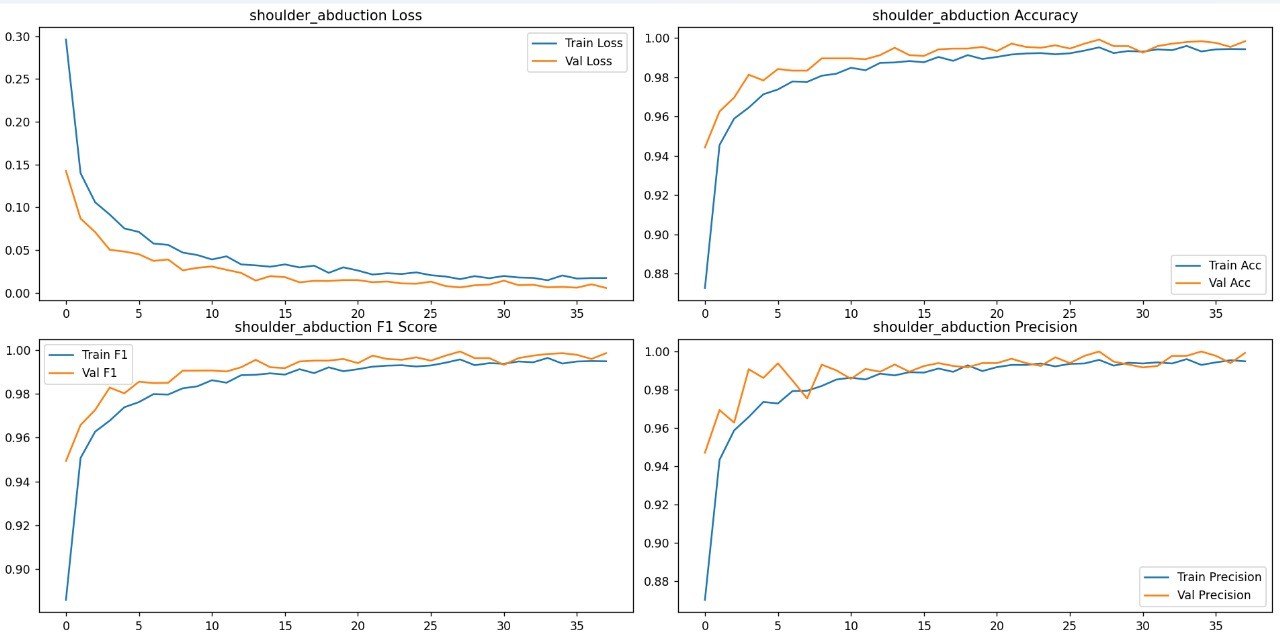


Fig. 3. Shoulder abduction-loss, accuracy, F1 score, and precision curves

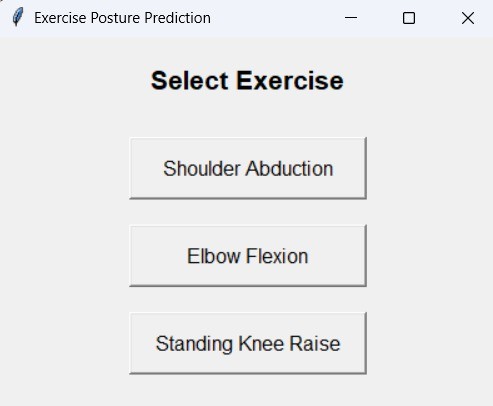


Fig. 4. Graphics User Interface

## GUI and Predictions

Fig. 4 shows the GUI made using tkinter and fig. 5, 6 and 7 show the predictions for elbow flexion, shoulder abduction

and standing knee raise. The prediction for elbow flexion(fig 5) also shows the angle at which the elbow is placed in degree celcius. The repetition count of each exercise performed is stored in a text file with its respective date, time and exercise name to keep track of everyday progress(fig. 8).

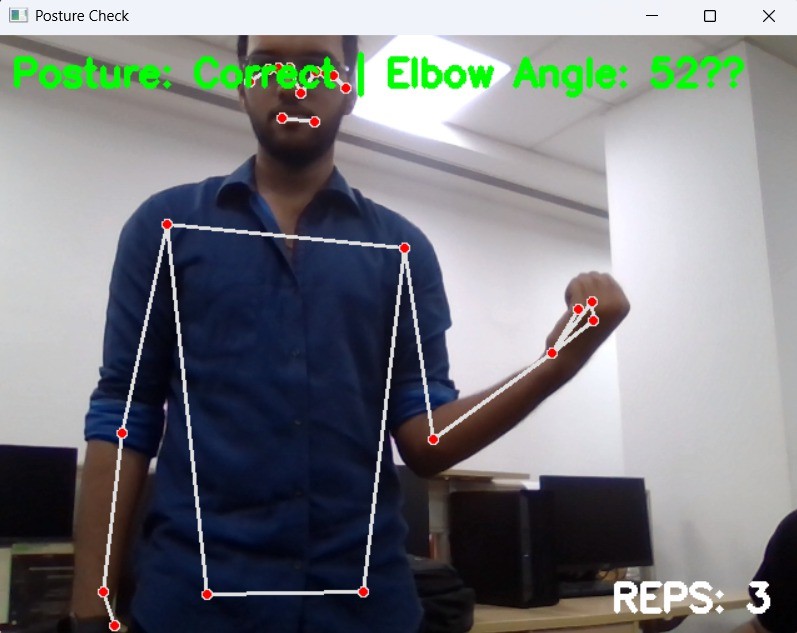


Fig. 5. Elbow Flexion Prediction

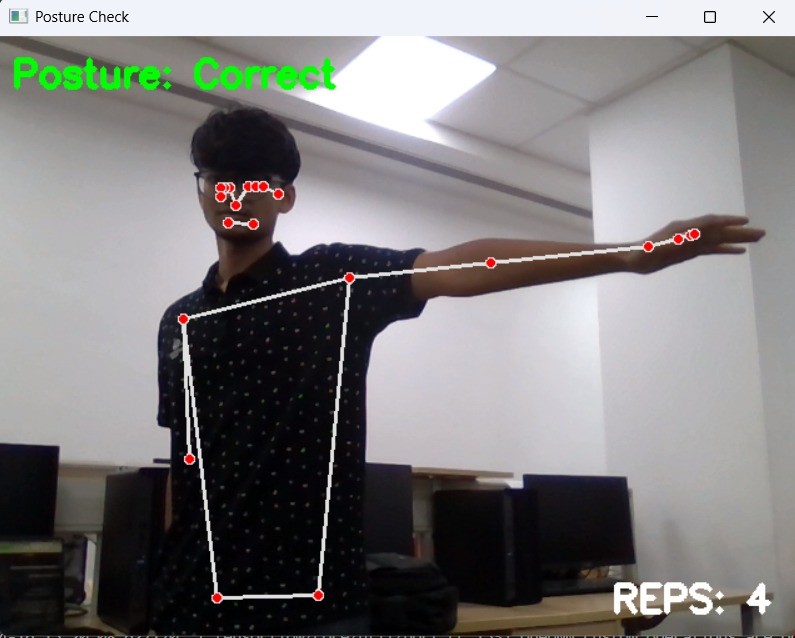


Fig. 6. Shoulder Abduction Prediction

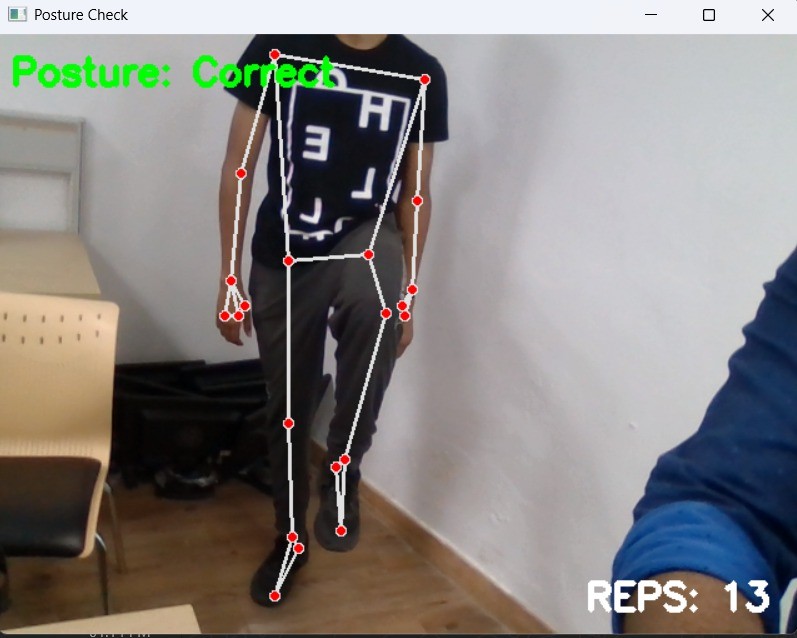


Fig. 7. Standing Knee Raise Prediction

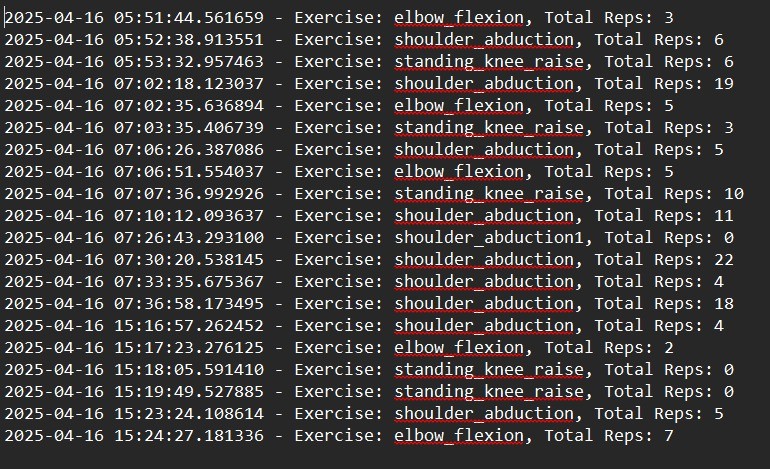


Fig. 8. Storage of Repetition Counts for each exercise

1. Conclusion

This report successfully demonstrates the implementation of a real-time pose classification and repetition counting system for physiotherapy exercises using a combination of Mediapipe, PyTorch, and computer vision techniques. By capturing and preprocessing body landmarks and training a deep learning model for posture recognition, the system achieves reliable classification performance across various exercise postures. The addition of a repetition counter, which operates based on class transitions, further improves its practicality in physical therapy settings. The GUI-based data collection and predic- tion modules streamline the interaction between users and the system, ensuring both usability and accuracy. While the current implementation supports a limited set of exercises, the framework is extensible to a broader range of movements and rehabilitation routines. Future enhancements may include sup- port for multi-angle pose estimation, automated labeling, and integration with cloud platforms for remote health monitoring. Overall, the system demonstrates strong potential for assisting physiotherapists and patients through real-time, automated exercise tracking.

1. K. He, X. Zhang, S. Ren, and J. Sun, ”Deep Residual Learning for Image Recognition,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 770–778.
2. J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, ”You Only Look Once: Unified, Real-Time Object Detection,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 779–788.
3. Y. Wang, Z. Pei, C. Wang, and Z. Tang, ”Depth-Aware Pose Estimation Using Deep Learning for Exoskeleton Gait Analysis,” Scientific Reports, vol. 13, no. 1, Dec. 2023.

References

1. R. Mun˜oz-Salinas et al., ”UCO Physical Rehabilitation: New Dataset and Study of Human Pose Estimation Methods for Physical Rehabilitation,” Sensors, Oct. 2023.
2. A. Vakanski, J. M. Ferguson, and S. Lee, ”A Deep Learning Framework for Assessing Physical Rehabilitation Exercises,” IEEE Trans. Neural Syst. Rehabil. Eng., Jan. 2020.
3. S. B. Kulkarni, A. Kuppelur, A. Shetty, S. Bidarakatti, and T. Sangre- sakoppa, ”Analysis of Physiotherapy Practices using Deep Learning,” Int. J. Res. Appl. Sci. Eng. Technol., Apr. 2024.
4. Y. Yang et al., ”Machine Learning-Based Computer Vision for Depth Camera-Based Physiotherapy Movement Assessment: A Systematic Review,” Mar. 2025.
5. Wang et al., ”How Deep Neural Networks Understand Motion? Toward Understanding Motion Recognition by Neural Networks,” Intell. Com- put., vol. 2, Feb. 2023.
6. H. Ben-Younes, R. Cadene, M. Cord, and N. Thome, ”Deep Learning- Based Human Body Pose Estimation in Providing Feedback for Physical Movement: A Review,” Heliyon, vol. 10, no. 17, p. e36589, Sep. 2024.
7. K. M. Bell et al., ”Verification of a Portable Motion Tracking System for Remote Management of Physical Rehabilitation of the Knee,” Sensors, vol. 19, no. 5, p. 1021, 2019.
8. J. Lu, J. Yang, D. Batra, and D. Parikh, ”Pose Detection Using Deep Learning for Gait Assessment,” IEEE Trans. Med. Imaging, vol. 42, no. 2, pp. 145-153, Feb. 2023.