IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

Gym-Buddy
(using Machine Learning)

Kritika Pandey, Snehal Kumbhar, Rhea Malpekar, Poonam Vengurlekar
Computer Science and Technology, Usha Mittal Institute of Technology,
SNDT Womens's University,
Juhu-Tara Road, Sir Vitthaldas Vidyavihar,
Santacruz(W), Mumbai, Maharashtra 40049.

Abstract: Over the past few years, there has been a notable increase in interest within the fitness industry, as more people are actively participating in a variety of exercises and sports. The provision of precise and immediate feedback on exercise form is paramount for optimizing performance, reducing the risk of injuries, and attaining fitness objectives. The "Gym-Buddy" initiative represents a machine learning endeavor aimed at transforming the fitness landscape by delivering comprehensive pose detection capabilities for exercise and sports enthusiasts. Gym-Buddy utilizes cutting-edge computer vision and deep learning methodologies to analyze and evaluate users' body positions during different exercise regimens and sports engagements. The key objectives of this project encompass:

Pose Detection: Gym-Buddy employs advanced algorithms for pose estimation to identify and monitor crucial body joints and movements in real-time. This functionality enables the system to furnish accurate feedback regarding exercise form, posture, and movement precision.

Exercise Guidance: The Gym-Buddy system is crafted to provide interactive exercise guidance, assisting users in aligning their bodies correctly, refining their movements, and maintaining proper form throughout their workout sessions. Users can receive visual prompts and auditory feedback via their preferred device, making it a versatile companion for home workouts, gym sessions, and outdoor activities.

Gym-Buddy is Accessible via a user-friendly web application, Gym-Buddy is designed to be seamlessly integrated into the routines of individuals at all fitness levels. The machine learning models within the system continuously enhance based on user input, ensuring personalized and precise feedback.

The overarching goal of this project is to empower individuals to achieve their fitness and sports performance aspirations by furnishing them with an intelligent and supportive training ally. Gym-Buddy not only enhances workout experiences but also advocates for healthier lifestyles and injury prevention, thereby contributing significantly to the overall well-being of its users.

Key-words:Machine Learning,Computer Vision,Deep Learning,Pose estimation,Mediapipe blazepose,Convolution neural network.

INTRODUCTION

Human pose estimation plays a pivotal role in localizing key body points to accurately identify individuals' postures within an image. This process serves as a fundamental step for various computer vision tasks, including human action recognition, tracking, human-computer interaction, gaming, sign language interpretation, and video surveillance. While fitness exercises offer numerous health benefits, they can also pose risks if performed incorrectly, leading to inefficacy and potential injuries. Exercise errors often stem from improper form or posture

adopted by the user.

In this study, we introduce Pose Trainer, an innovative application designed to detect users' exercise poses and offer personalized, comprehensive recommendations for improving their form. Utilizing state-of-the-art pose estimation techniques, Pose Trainer precisely identifies the user's pose and evaluates the vector geometry throughout the exercise, providing actionable feedback to enhance performance and safety.

LITERATURE REVIEW

Sr No.	Title	Methodology/Technology	Observation
1	Alexander Toshev Christian Szegedy Google DeepPose: Human Pose Estimation via Deep Neural Networks . **2014**	Linear Regression Alex Net	In this approach, pose estimation is formulated as a CNN-based regression problem towards body joints. The model consisted of an Alex Net backend (7 layers). The model is trained using L2, loss for regression. This approach reasons about pose in a holistic fashion, i.e., even if certain joints are hidden, they can be estimated if the pose is reasoned about holistically. The paper argues that CNNs naturally provide this sort of reasoning and demonstrate strong results.
2	Joao Carreira,Pulkit Agrawal ,Katerina Fragkiadaki, Jitendra Malik.Human Pose Estimation with Iterative Error Feedback . **2018**	deep google.net Iterative Error Prediction ConvNet	This is an interesting paper that follows a method that predicts the estimates and then corrects them iteratively. Instead of directly predicting the outputs in one go, they use a self-correcting model that progressively changes an initial solution by feeding back error predictions, and this process is called Iterative Error Feedback (IEF). It employs a standard ConvNet architecture pretrained on ImageNet: the very deep google net.

3	Jonathan Tompson, Ross Goroshin, Arjun Jain, Yann LeCun, Christoph Bregler New York University Efficient Object Localization Using Convolutional Networks. **2015**	Heat Model Multi-resolution CNN Architecture	This approach generates heatmaps by running an image through multiple resolution banks in parallel to simultaneously capture features at a variety of scales. The output is a discrete heatmap instead of continuous regression. A heatmap predicts the probability of the joint occurring at each pixel. A multi-resolution CNN architecture (coarse heatmap model) is used to implement a sliding window detector to produce a course heatmap output.
4	Mihai Fieraru ,Anna Khoreva Leonid ,Pishchulin Bernt Schiele. 'Learning to Refine Human Pose Estimation', Max Planck Institute for Informatics,Saarland Informatics Campus,Saarbrücken, Germany. **2018**	MPII Human Pose Pose Track Benchmarks Google API System	Learning to Refine Human Pose Estimation have given insights on refining the human pose estimation where It takes the help of RGB image and body pose estimate as input to render out the suitable output and Exploiting the dependencies between the image and the predicted body pose makes it easier for the model to identify the errors in the initial estimate and how to refine them along with methods to achieve state-of-the-art results on the challenging MPII Human Pose and Pose Track benchmarks.
5	Simen Thys,Wiebe Van Ranst,Toon Goedem´e. 'Fooling automated surveillance cameras:adversarial patches to attack person detection'.EAVISE, Technology Campus De Nayer, KU Leuven, Belgium **2019**	YOLOv2 object detector CNN RNN	Object detection of this paper they target the popular YOLOv2 object detector, An example of real-world adversarial attack. Sharif et al. demonstrate the use of printed eyeglasses that can be used to fool facial recognition systems for making the patches more robust. Thus, vulnerability of certain neural networks has been exposed.

NEED OF THE STUDY

Despite the abundance of working models showcased in research papers and news articles, there is a noticeable absence of an application in the market that addresses the need for precise exercise posture detection and feedback. We possess a functional model ready for commercialization, which leverages cutting-edge technologies such as object detection and motion detection. Our aim is to democratize these advanced technologies and make them accessible to the general public, enabling individuals to improve their gym stance, form, and overall health effortlessly.

With the proposed application, many users will be able to train effectively without the constant need for human assistance. The app's ability to provide real-time feedback on exercise posture is revolutionary in the health and fitness industry, as it can pinpoint mistakes and offer corrective guidance without human intervention.

Our course project, Pose Trainer, is designed to assist individuals in maintaining correct exercise postures by utilizing a combination of state-of-the-art pose estimation techniques and machine learning algorithms. The primary objective of Pose Trainer is to reduce the risk of injuries and enhance the efficacy of workouts using only a computer and a webcam. The initial phase of Pose Trainer focuses on human pose estimation, a challenging yet highly relevant area within computer vision.

3.1 ARCHITECTURE

Figure 1: **Architecture of Gym-Buddy**

3.2 ALGORITHM

Mediapipe BlazePose algorithm

The Mediapipe BlazePose algorithm, provided by Mediapipe, is a robust human body detection method capable of calculating the positions of 33 key points on the human body, as depicted in the accompanying Figure. This algorithm is specifically designed for detecting human body postures and can accurately determine the coordinates of each joint in the body. These coordinates range from 0 to 32, with the exception of coordinate 0, which represents the ”nose”; all other coordinates are symmetrical.

Fitness movements primarily involve the articulation of body joints, making Mediapipe well suited for detecting joint movements during fitness activities. The algorithm comes pre-trained for human body detection, eliminating the need for additional data collection to train the model. Mediapipe excels at identifying fitness movements due to its comprehensive joint detection capabilities.

Presently, there exists a method that employs Mediapipe to recognize fitness movements. This approach initially utilizes Mediapipe to identify the key points of the entire body and acquire their coordinate positions. Subsequently, each coordinate is utilized to determine the current fitness category using a K-nearest-neighbor (K-NN) classifier. This method simplifies the process of counting body nodes and categorizing fitness activities. By considering a wide range of angles, user variations, and background scenarios, the classification accuracy for different fitness types can be significantly improved

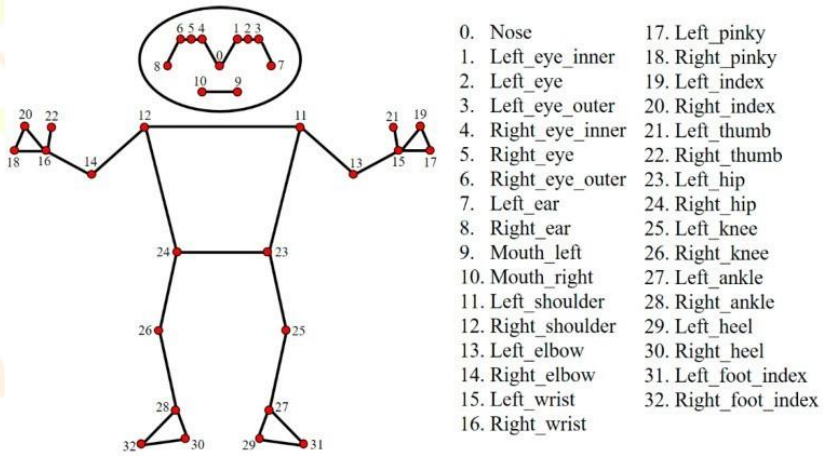

Figure 2: **MediaPipe**

Convolutional neural network

A Convolutional Neural Network (CNN) is a specialized type of neural network designed for processing data with grid-like structures, such as images or videos. Unlike traditional multi-layer perceptrons (MLPs), CNNs use convolutional layers to extract meaningful features from the input data, such as edges, textures, and patterns. These features are then passed through additional layers, including pooling layers for downsampling and fully connected layers for classification or regression tasks. CNNs have revolutionized tasks like image recognition, object detection, and pose estimation by automatically learning hierarchical representations of the input data, making them highly effective for a wide range of computer vision applications.

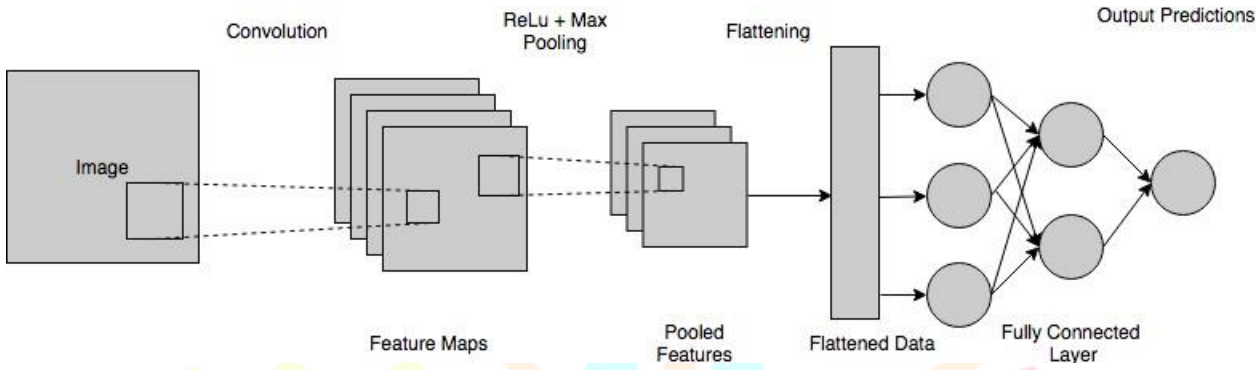

Figure 3: **Convolution Neural Network**

RESEARCH METHODOLOGY

The research methodology for Gym Buddy involves utilizing machine learning algorithms like Convolutional Neural Networks (CNNs) to analyze video frames captured by a camera and identify key points of a person's body, such as joints and limbs, using pre-trained models in Mediapipe. The OpenCV library is then employed to track the movement of these body parts over time and provide feedback on the person's form and technique.

A. Data Collection

Data collection plays a crucial role in training AI models. Popular libraries like Mediapipe and OpenCV are utilized for data collection and processing in computer vision tasks. Connecting a camera to the computer and configuring it to capture the required images is the initial step. OpenCV is used to capture images or video frames, which are then processed using Mediapipe to extract features like hand gestures and body pose.

B.Pre-processing

Data processing is essential for preparing data for machine learning models. Mediapipe and OpenCV aid in cleaning the data by removing invalid or redundant samples and filling in missing data. The pre-processing step involves normalizing the data, augmenting it, and extracting features to enhance its quality for training.

C. Training

Training the AI model using Mediapipe and OpenCV involves iterative steps where the model's parameters are adjusted to improve its accuracy and reliability. After pre-processing the data, it is split into training, validation, and testing sets. A CNN can be used for image classification tasks, and the training set is utilized to train the model by feeding pre-processed data and adjusting its parameters to enhance performance. The validation set helps fine-tune the model's parameters for improved accuracy.

D. Validation

Validation is crucial for assessing the model's accuracy and reliability on new data. By using separate validation data from the training set, the model's ability to generalize well and make accurate predictions or classifications on unseen data can be evaluated. This validation process ensures the model's effectiveness in real-world scenarios

beyond the training data.

4.2 FEATURES

We have researched several projects and research papers and have come across a lot of ideas on how to move forward with it. We have as of yet decided on what our app will look like, the functions that it will have. As in, it will have an introduction or 'About' page on startup. It will have a section in the navigation bar that will guide users to the formation check, where the user can check their gym formation and will be guided on how to pose properly to avoid injury and have maximum benefits. It will have a gym scheduler and a registration page too.

FEATURE 1 – **REAL-TIME MONITORING** Real-time monitoring of Gym buddy application can involve several different aspects: User activity monitoring, Performance monitoring, Data Monitoring, Model Monitoring, Security Monitoring. Overall, real-time monitoring of application can help ensure its effectiveness, reliability, and security. .

FEATURE 2 – **PROGRESS TRACKING** The application allows users to track key performance metrics such as weight, body fat percentage, muscle mass, and BMI. The application allows users to set fitness goals based on their personal preferences, fitness level, and overall health objectives. Users can set goals for weight loss, muscle gain, cardiovascular health, or other specific areas of focus.

FEATURE 3 – **ALERT** The application will pops-up a message if the person is doing excess workout than require. It will analyze user data and provide valuable insights and feedback to help improve its performance and user experience

FEATURE 4 – **Diet-Recommendation System** Our Machine learning Diet Recommendation System is Based on Python's LangChain Libraries.LangChain is a framework for developing applications powered by language models. Our System works with the help of OpenAI API ,which is engineered to recommend Breakfast, Dinner and Workouts for the users after taking their age,weight,Allergies,Generic Diseases,height,preference foods.

FEATURE 5 – **BMI Calculator** The BMI calculator is a fundamental tool used in health and fitness assessments to estimate an individual's body composition based on their height and weight. BMI is a simple yet widely accepted metric that provides a rough indicator of whether a person's weight falls within a healthy range relative to their height.

FUTURE SCOPE

In our future endeavors, we plan to integrate our trained convolutional neural network and enhanced pose estimation model into our web application, expanding its capabilities to provide real-time feedback and guidance during workouts across a wider range of exercises and sports forms. To achieve higher accuracy, we aim to increase the training and validation images, necessitating additional hardware resources such as a dedicated graphics card for efficient processing. Our overarching goal is to not only critique form but also motivate users through personalized feedback, ultimately enhancing their fitness journey and experience.

CONCLUSION

In conclusion, our work has resulted in the development of an end-to-end trained convolutional neural network for action recognition, capable of generating silhouettes by subtracting static backgrounds and classifying intermediate workout actions. This model lays the groundwork for a workout application in sports and fitness, leveraging workout-specific action recognition. Additionally, our enhanced pose estimation model showed promising results and has considerable potential for further improvement and expansion. Our top-performing model achieved an impressive testing accuracy score of 79 out of 100 on squatting images from the PennAction dataset. Looking ahead, we aim to incorporate additional exercises and sports forms into our training data to enhance accuracy further. This expansion will necessitate additional hardware resources, such as a dedicated graphics card, to handle increased computational demands. Our ultimate objective is to enhance user feedback by not only providing critique but also offering motivation to help individuals stay motivated and committed to their workouts

REFERENCES

References

[1] Moez Baccouche, Franck Mamalet, Christian Wolf, Christophe Garcia, and Atilla Baskurt *Sequential deep learning for human action recognition* Albert Ali Salah and Bruno Lepri, editors, Human Behavior Understanding, pages 29–39, Berlin, Heidelberg, 2011. Springer Berlin Heidelberg.

[2] Z. Cao, T. Simon, S.-E. Wei, and Y. Sheikh *Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields* . ArXiv e-prints, November 2016.

[3] Yu-Wei Chao, Jimei Yang, Brian Price, Scott Cohen, and Jia Deng. *Forecasting human dynamics from static images* In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017.

[4] Steven Chen and Richard Yang. *Pose trainer: Correcting exercise posture using pose estimation.* ResearchGate, 03 2018.

[5] Yigithanˇ Dedeoglu,ˇ B. Ugurˇ Töreyin, Ugurˇ Güdükbay, and A. Enis Çetin. *Silhouette-based method for object classification and human action recognition in video. In Thomas S. Huang, Nicu Sebe, Michael S. Lew, Vladimir Pavlovi´c, Mathias Kölsch, Aphrodite Galata, and Branislav Kisaˇcanin, editors, Computer Vision in Human-Computer Interaction,* pages 64–77, Berlin, Heidelberg, 2006. Springer Berlin Heidelberg.

[6] Jeff Donahue, Lisa Anne Hendricks, Marcus Rohrbach, Subhashini Venugopalan, Sergio Guadarrama, Kate Saenko, and Trevor Darrell. *Long-term Recurrent Convolutional Networks for Visual Recognition and Description.* arXiv e-prints, page arXiv:1411.4389, November 2014.