Real Time Human Body Posture Analysis Using Deep Learning

Ram krishna[1], Vaibhav Shekhar [2], Aman Raj[3], Prof. Ganesh V. Madhikar[4]

[1, 2, 3]*Student,* [4]*Assistant Professor, Dept. of Electronics and Telecommunication Engineering, Sinhgad College of Engineering, Maharashtra, India*

Abstract: We present a novel approach for accurately estimating the pose of objects in a low-cost and resource-efficient manner, making it suitable for deployment on embedded systems. Our algorithm comprises of two primary stages: object detection and spatial reconstruction. In the first stage, we employ a Convolutional Neural Network (CNN) called PoseNet for object detection. This approach has proven to be effective in detecting and localizing objects in an image. Next, utilizing stereo correspondences, we 3D reconstruct the spatial coordinates of multiple ORB features within the object's bounding box. This enables us to accurately estimate the position of the object in space.

To calculate the final position of the object, we compute a weighted average of the stereo-corresponded key points' spatial coordinates. The weights are proportional to the level of ORB stereo matching, which enables us to obtain a more accurate estimate of the object's position in space. Our algorithm was tested in a calibrated environment, and we compared the results with a deep learning-based method using various datasets. The results show that our approach outperforms existing methods in terms of accuracy, while maintaining a low cost and efficient resource utilization. Our proposed method has several applications, including the quantitative and qualitative analysis of human posture. By analyzing all aspects of a person's posture, we can determine if there are any postural deviations, imbalances, or muscle weaknesses that may be causing pain or discomfort. This information can then be used to develop personalized rehabilitation programs, reducing the risk of injury and enhancing athletic performance.

Furthermore, our approach can be used in various assistive technology applications, such as the control of robotic arms for pick-and-place tasks. The low-cost and resource-efficient nature of our algorithm make it ideal for deployment in embedded systems, enabling us to develop affordable and accessible assistive technology solutions.

In conclusion, our proposed algorithm provides an accurate, low-cost, and resource-efficient solution for pose estimation, with a wide range of potential applications in human posture analysis, assistive technology, and beyond.

I. INTRODUCTION

Human Pose Estimation (HPE) involves identifying and classifying the joints in the human body to capture a set of coordinates for each joint, also known as a key point, that can describe a pose of a person. There are several methods for pose estimation, including OpenPose, Posenet, and DeepPose. This paper discusses the evolution of human pose estimation over the years and concludes that Posenet is the most suitable technique for a real-world Android application. Posenet provides real-time pose estimation for the human body, allowing for real-time evaluation on the client side. This is achieved with TensorFlow, which enables faster and more privacy-respecting model inference on Android. Posenet is an open-sourced technology that can extract the essential 16-17 points natively and draw a skeleton of the human pose with these points. This skeleton can be used to derive angles between points, enabling effective correction of the user's pose.

Human pose estimation and tracking is a computer vision task encompassing the detection, association, and tracking of semantic key points, such as "right shoulders," "left knees," or the "left brake lights of vehicles." Real-time performance of semantic key point tracking in video footage demands significant computational resources, which limits the accuracy of pose estimation. Pose estimation has diverse applications, including interactive installations that respond to human motion, augmented reality, animation, fitness tracking, and more. We aim to foster experimentation and application of pose detection in unique projects through the accessibility of our model. While various alternate pose detection systems have been open-sourced, all necessitate specialized hardware and/or cameras, as well as extensive system setup.

A. Problem Statement
The optimization of human body productivity and enhancement of athletic performance through various techniques can facilitate the development of numerous assistive technologies in the field of robotics.

B. Objective

The primary aim of this project is to promote healthy living and enhance posture through the use of advanced technologies such as augmented reality, virtual reality, and training robots. The proposed project seeks to investigate new assistive technologies that can positively impact our daily lives.

II. LITERATURE SURVEY

This study involved a carefully designed experiment that included 24 female participants divided into three groups: control, exercise, and nutrition. The participants were subjected to a simulated microgravity environment through a head down bed rest (HDBR) for 60 days. The objective of the study was to investigate the effects of microgravity on the regulation of balance function and to evaluate the effectiveness of countermeasures, including specific exercises and a tailored diet. The participants' orthostatic and dynamic balance were assessed nine and two days before the experiment, on the first day of getting up, the following day, and four and ten days after, under two visual conditions: eyes open and eyes closed. The results indicated that the postural balance performances were better with eyes open than with eyes closed, and that the static and dynamic postural performances were impaired on the first day of recovery following HDBR. This impairment lasted up to four days after getting up, after which the participants recovered their initial performances. Notably, the exercise group demonstrated a faster recovery of static postural performances compared to the other groups, while there were no significant differences in the recovery of the dynamic balance performances.[1]

In the field of human-computer interaction, the use of skeleton data for human posture recognition is an important research topic. This paper presents a new algorithm that aims to improve the accuracy of human posture recognition. The proposed algorithm defines a 219-dimensional vector that includes angle and distance features based on the local relationship between joints and their global spatial location. During human posture classification, the rule learning method is used with Bagging and random sub-space methods to create different samples and features for improved classification. The proposed algorithm is evaluated on four human posture datasets, and the results show that it effectively recognizes many kinds of human postures. The rule-based learning method provides higher interpretability compared to traditional machine learning methods and CNNs.[2]

In 2019, researchers utilized Deep Learning Models for Human Pose Estimation and implemented an Interactive Computer Vision System for Home-based Physical Therapy. However, the system lacked a side-view option for users and did not generate a detailed feedback mechanism that evaluates the patient's performance in-depth. The researchers aimed to develop an algorithm that provides more nuanced feedback rather than just an overall performance evaluation.[3]

This literature survey examines the current state of human pose estimation methods, dividing them into two main categories: regression-based and heatmap-based approaches for single-person estimation, and top-down and bottom-up approaches for multi-person estimation. While there have been significant improvements in these methods, there remains a need for further improvement to make them more applicable for real-world use.. [4]

Sr No	Author	Proposed Method	Softwa-re Method	Accu-racy	Year
1	Marion viguier Philippe Dupui Richard Montova	Described the proposed work of the of pose net system.	Python module Library	Depends On Trained model	2009
2	Weili Ding Han liu	Described the idea and implementation of the machine learning and cybernetics	Python module Library Vs code	Depends On Trained model	2020
3	Yiwen Gu	Deep learning method for pose estimation	Python module Library Vs code	Depends On Trained model	2019

## III. SOFTWARE/HARDWARE REQUIREMENTS

A. Algorithm Theory

Human pose estimation is a computer vision technique used to predict the positions of key joints or body parts in images or videos. This involves identifying joints such as wrists, elbows, knees, and ankles as key points. When input images are fed into a pose estimation model, it outputs the coordinates of these detected body parts, along with a confidence score indicating the certainty of the estimate. There are two types of pose estimation techniques: two-dimensional (2D) and three-dimensional (3D). The 2D approach involves extracting (X,Y) coordinates for each key joint of an RGB image, while the 3D approach estimates (X,Y,Z) coordinates. In this article, we will focus on the working of 3D human pose estimation, which aims to detect the (X,Y,Z) coordinates of joints in an image containing a person. By joining these joints, the posture of a person can be inferred.

Pose — The PoseNet model is capable of detecting and returning a pose object, which consists of a comprehensive list of key points along with an associated confidence score. The model also provides a pose confidence score that reflects the overall confidence level of the estimated pose in an image, ranging from 0.0 to 1.0. The key points represent the parts of the person's pose that are estimated, such as the nose, right ear, left knee, right foot, and so on, and they include both a position and a related confidence score. The confidence score can be used to conceal key points that are not predicted with sufficient confidence.

This implementation employs the PoseNet model architecture, which is available in two versions: single-person pose detector and multiple-person pose detector. The former is simpler and faster than the latter and can estimate the joint positions of a single person in an image, without being responsible for identifying the person. The focus of the model is on detecting the positions of the key joints to facilitate movement tracking.

B. Raspberry Pi 4GB

The project uses Raspberry Pi 4GB, Raspberry Pi Camera Module and a small breadboard with an LED, resistor, and push button. This is the hardware configuration to run the pipeline: A Raspberry Pi can be used for a wide range of applications. Making your Raspberry Pi into a retro arcade machine, using it as a web server, or using it as the brain for a robot, security system, IoT device, or dedicated Android device are all popular uses.
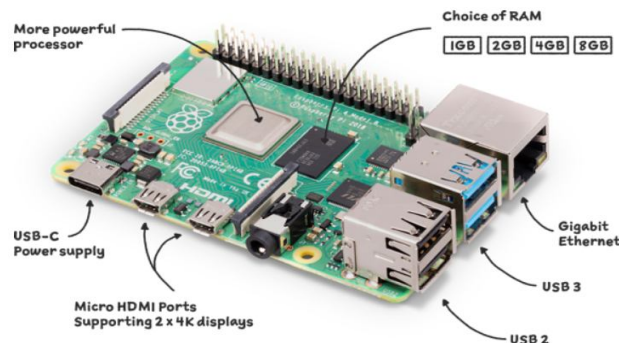

C. Software Implementation

This implementation uses the Pose Net model integrated in TensorFlow Lite, everything is written in Python to be run on the Raspberry Pi 4. The code written for this project develops a pipeline to feed images to the model, process them using TensorFlow pretrained model, decode the model output, and draw key points and limbs on the processed images. Post-processing can be used to convert these images into a video!

D. LCD Display

LCD (Liquid Crystal Display) is a type of flat panel display which uses liquid crystals in its primary form of operation. LEDs have a large and varying set of use cases for consumers and businesses, as they can be commonly found in smartphones, televisions, computer monitors and instrument panels. LCDs were a big leap in terms of the technology they replaced, which include light-emitting diode (LED) and gas-plasma displays. LCDs allowed displays to be much thinner than cathode ray tube (CRT) technology. LCDs consume much less power than LED and gas display displays because they work on the principle of blocking light rather than emitting it. Where an LED emits light, the liquid crystals in an LCD produces an image using a backlight.

IV. RESULT
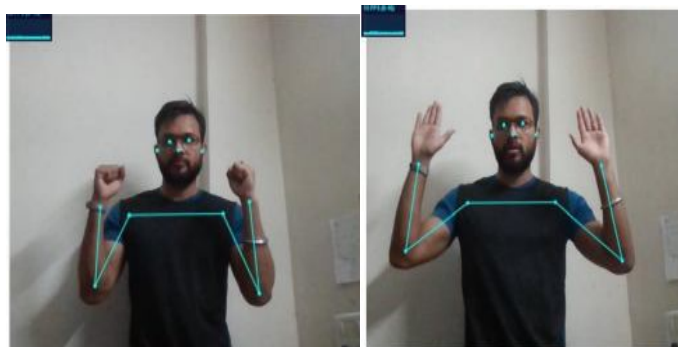
Our research is the first to apply deep convolutional neural networks to end-to-end 6-degree-of-freedom (6-DOF) camera pose localization. We have demonstrated that transfer learning from classifiers trained on relatively small datasets can be used to overcome the need for millions of training images. Our findings show that these networks retain sufficient pose information in their feature vectors, even though they are trained to produce pose-invariant outputs. Our future research aims to investigate the use of Multiview geometry as a source of training data for deep pose regressors and explore probabilistic extensions to this algorithm. Additionally, we acknowledge that there is an upper limit to the physical area that can be localized by a finite neural network, which we leave for future exploration.

V. CONCLUSION

The primary advantage of the approach described is its ability to rapidly and autonomously provide high-probability detections based on conservative estimations of key point parameters in image transformations. A concise list of the most probable pose positions, 3D poses, and sizes is derived, which can be refined in a subsequent step to determine the correct object parameters. Additionally, as the algorithm always obtains the most probable pose for each image position, the parameter space is significantly reduced, allowing more sophisticated techniques to be employed to identify objects in situations where the highest convolution results do not match the object. Therefore, the algorithm, as currently presented, is intended to provide a swift parameter space reduction, which can directly estimate poses or initiate a refined search using alternative methods. Further research is needed to explore the full potential of this approach and to identify its limitations.

VI. FUTURE SCOPE

1) The use of pose estimation has a multitude of applications in various fields. For instance, it can be utilized in robotics to enable robots to imitate the movements of a human instructor, who can demonstrate the desired action through a pose skeleton. This eliminates the need for manual programming of the robots to follow trajectories, as the robot can compute how to move its articulators to perform the same action as the human instructor.

2) In addition, pose estimation can be employed in security and surveillance systems to enhance their capabilities. It can also be leveraged to detect if an individual has fallen or is experiencing a medical emergency.

3) Furthermore, pose estimation technology has been used in interactive gaming applications, such as the Kinect, which utilizes 3D pose estimation using IR sensor data to track the motion of human players. This technology is then utilized to render the actions of virtual characters, providing a highly immersive gaming experience.

4) Applications that can enhance security and surveillance.

REFERENCES

[1] Kepski, M., & Kwolek, B. (2012, September) Human fall detection by mean shift combined with depth connected components. In International Conference on Computer Vision and Graphics (pp. 457-464). Springer, Berlin, Heidelberg.

[2] Shi, G., Zou, Y., Jin, Y., Cui, X., & Li, W. J. (2009, February). Towards HMM based human motion recognition using MEMS inertial sensors. In 2008 IEEE International Conference on Robotics and Biomimetics (pp. 1762-1766). IEEE

[3] Baek, W. S., Kim, D. M., Bashir, F., & Pyun, J. Y. (2013, January). Real life applicable fall detection system based on wireless body area network. In 2013 IEEE 10th Consumer Communications and Networking Conference (CCNC) (pp. 62-67). IEEE

[4] https://github.com/Pose-Group/DCPose

[5] https://paperswithcode.com/task/pose-estimation

[6] https://blog.tensorflow.org/2018/05/real-time-human-pose-estimation-in.html

[7] https://medium.com/analytics-vidhya/pose-estimation-on-the-raspberry-pi-4-83a02164eb8e

BIOGRAPHIES

Ram krishna
Student,
Dept. of Electronics and
Telecommunication Engineering,
Sinhgad College of Engineering,
Maharashtra, India.

Vaibhav Shekhar
Student,
Dept. of Electronics and
Telecommunication Engineering,
Sinhgad College of Engineering,
Maharashtra, India.

Aman Raj
Student,
Dept. of Electronics and
Telecommunication Engineering,
Sinhgad College of Engineering,
Maharashtra, India.

Prof. Ganesh V. Madhikar
Assistant Professor,
Dept. of Electronics and
Telecommunication Engineering,
Sinhgad College of Engineering,
Maharashtra, India

INTERNATIONAL JOURNAL
FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)