International Journal of Research in Computer & Information Technology (IJRCIT),
Vol.10, Issue 2, March 2025
E-ISSN: 2455-3743
Available online at www.ijrcit.co.in

Robotic Mimicry Leveraging Posenet Software for Human-Like Movements

[1]Pooja Chandewar, [2]Yashika Wanjari, [3]Pooja Raut, [4]Dipanjali Warade, [5]Pratiksha Motghare, [6]Dr. Archana Dehankar

Abstract— This research presents a novel algorithm for accurate pose estimation tailored for human posture mimicry in robotics, optimized for low-cost and resource-constrained embedded systems. Integrating Convolutional Neural Network (CNN) PoseNet for object detection with stereo correspondences for spatial reconstruction, the algorithm enables precise estimation of human posture. Leveraging ORB features and weighted averaging of stereo-corresponded key points, it achieves superior accuracy while maintaining resource efficiency. Comparative testing against deep learning-based methods demonstrates its efficacy for robotic applications. The algorithm's affordability and efficiency make it well-suited for implementing human posture mimicry in robotics, with potential applications in personalized rehabilitation and assistive technology. In conclusion, this research presents a versatile and accessible solution for advancing human-robot interaction through posture mimicry.

Keywords— PoseNet, Convolutional Neural Network (CNN), Robotic control, Embedded systems.

I. Introduction

Human Pose Estimation (HPE) is a crucial task in computer vision, involving the identification and classification of key points on the human body to define a pose accurately. Over the years, several methods, including OpenPose, Posenet, and DeepPose, have been developed for this purpose. Among these techniques, Posenet has emerged as particularly suitable for real-world applications, especially in Android environments. Utilizing TensorFlow, Posenet enables real-time pose estimation directly on the client side, offering efficient and privacy-respecting model inference. By extracting essential key points and generating a skeletal structure of the human pose, Posenet facilitates the derivation of angles between points, enabling effective pose correction.

Manuscript Received 10 Feb 2025; Revised 28 Feb 2025 and published on 30 March 2025

Pooja Chandewar, Yashika Wanjari, Pooja Raut, Dipanjali Warade, Pratiksha Motghare, Dr. Archana Dehankar *Department of Computer Technology, Priyadarshini College of Engineering, Nagpur, Maharashtra, India.* ***Mail Id:*** chandewarpooja19@gmail.com, yashikawanjari@gmail.com, poojarautm2022@gmail.com, dipanjali.warade2003@gmail.com, motgharepratiksha6@gmail.com, archana_dehankar@rediffmail.com

Meanwhile, in robotics, humanoid robots have revolutionized human-robot interaction by meticulously emulating human actions and behaviors. Powered by advanced AI, machine learning algorithms, and sophisticated mechanical engineering, humanoid robots can engage in interactions with humans that closely resemble human communication and behavior. These remarkable machines find applications across diverse industries, from healthcare and customer service to manufacturing and entertainment, offering capabilities previously exclusive to humans.

Despite their extraordinary potential, the rise of humanoid robots has raised ethical and societal concerns regarding human employment, human-robot relationships, privacy, and security. As technology progresses, humanoid robots are poised to assume increasingly pivotal roles, forging deeper collaborations between humans and machines and reshaping various aspects of human existence. With ongoing advancements in AI and robotics, the future holds promise for further integration, fundamentally altering the dynamics of human-robot interaction and the fabric of society.

II. Problem Statement

Human pose estimation (HPE) is a critical task in various domains, including robotics, healthcare, and entertainment. Traditional pose estimation methods often require high computational resources, limiting their applicability, particularly in low-cost and resource-constrained embedded systems. While existing pose estimation techniques like Posenet offer real-time performance, they may not fully optimize resource utilization for deployment on embedded systems. Thus, there is a need for a pose estimation algorithm tailored specifically for low-cost embedded systems, balancing accuracy with efficiency. The objectives of this research are:

III. Objecties

The primary objective of this research is to develop a novel pose estimation algorithm optimized for deployment on low-cost and resource-constrained embedded systems. The algorithm aims to achieve the following objectives:

- Enhance accuracy: Develop techniques to improve the accuracy of pose estimation while considering the computational limitations of embedded systems.

- Optimize resource utilization: Design algorithms that efficiently utilize system resources, ensuring real-time performance on low-cost embedded hardware.Enable real-world applications: Create a pose estimation solution that is practical and accessible for deployment in robotics, assistive technology, and other real-world scenarios.
- Comparative evaluation: Conduct thorough comparative evaluations with existing pose estimation methods to demonstrate the effectiveness and efficiency of the proposed algorithm.
- Facilitate human-robot interaction: Enable humanoid robots to accurately mimic human posture and gestures, fostering seamless and natural interactions between humans and robots.

IV. Related Work

Firstly, a study focusing on the effects of microgravity on balance regulation utilized a carefully designed experiment involving 24 female participants divided into control, exercise, and nutrition groups. Through a simulated microgravity environment via head down bed rest (HDBR) for 60 days, the study aimed to investigate balance function regulation and evaluate countermeasures such as specific exercises and tailored diets. Results indicated impaired static and dynamic postural performances following HDBR, with the exercise group showing faster recovery in static postural performances.[1]

In the realm of human-computer interaction, a new algorithm is proposed to enhance the accuracy of human posture recognition by defining a 219-dimensional vector incorporating angle and distance features based on joint relationships and global spatial locations. Leveraging rule learning methods, Bagging, and random sub-space techniques, the algorithm demonstrates effective recognition of various human postures across multiple datasets, offering higher interpretability than traditional machine learning methods and CNNs.[2]

Additionally, researchers in 2019 utilized Deep Learning Models for Human Pose Estimation to develop an Interactive Computer Vision System for Home-based Physical Therapy. However, the system lacked a side-view option for users and detailed feedback mechanisms. The researchers aimed to address this limitation by developing an algorithm providing nuanced feedback on the patient's performance, enhancing the overall therapeutic experience.[3]

Furthermore, the literature survey categorizes human pose estimation methods into regression-based and heatmap-based approaches for single-person estimation, as well as top-down and bottom-up approaches for multi-person estimation. Despite significant advancements, there remains a need for further improvement to enhance the applicability of these methods for real-world scenarios.[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

V. Software/Hardware Requirements

A. Algorithm Theory:

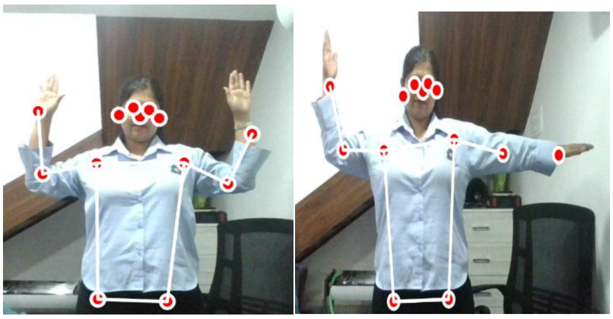
Human pose estimation, a pivotal computer vision technique, predicts the positions of key joints or body parts in images or videos. This entails identifying joints like wrists, elbows, knees, and ankles as key points. When input images are processed by a pose estimation model, it outputs the coordinates of these detected body parts, alongside a confidence score indicating the certainty of the estimate. There exist two types of pose estimation techniques: two-dimensional (2D) and three-dimensional (3D). The 2D approach extracts (X,Y) coordinates for each key joint from an RGB image, while the 3D approach estimates (X,Y,Z) coordinates. This article focuses on 3D human pose estimation, which aims to detect the (X,Y,Z) coordinates of joints in an image containing a person, enabling inference of their posture.

PoseNet Model: The PoseNet model, integral to this implementation, detects and returns a pose object comprising a comprehensive list of key points along with associated confidence scores. Additionally, it provides a pose confidence score reflecting the overall confidence level of the estimated pose in an image, ranging from 0.0 to 1.0. These key points represent the estimated parts of the person's pose, such as the nose, right ear, left knee, and right foot, each accompanied by a position and a related confidence score. Utilizing this confidence score, less confidently predicted key points can be concealed.

LCD Display: LCD (Liquid Crystal Display) technology revolutionized displays by employing liquid crystals for operation. Widely utilized in smartphones, televisions, computer monitors, and instrument panels, LCDs replaced older technologies like LEDs and gas-plasma displays. Compared to traditional cathode ray tube (CRT) displays, LCDs offer significantly thinner displays and consume less power. Unlike LEDs, which emit light, LCDs operate by blocking light, utilizing a backlight to produce images through liquid crystals. This results in reduced power consumption and thinner displays, making LCDs a preferred choice for various consumer and business applications.

- Design an intuitive web interface for algorithm selection and data input.
- Develop modular implementations of key analytics algorithms.
- Enable users to upload and test custom datasets.
- Provide real-time visual feedback for data transformations and algorithm execution.
- Ensure compatibility across devices and platforms.
- Integrate a file handling system for persistent data processing.
- Include guided tutorials, sample datasets, and tooltips for ease of learning.
- Evaluate platform performance through usability testing and feedback.
- Support personalization features such as user history and progress tracking.
- Provide a scalable backend to support future extensions like machine learning integration.

VI. Result And Discussion



Our research marks a pioneering endeavor in applying deep convolutional neural networks to achieve end-to-end 6-degree-of-freedom (6-DOF) camera pose localization. We demonstrate the efficacy of transfer learning from classifiers trained on relatively modest datasets, circumventing the necessity for vast training image repositories. Remarkably, our findings indicate that these networks preserve ample pose information within their feature vectors, despite being trained to yield pose-invariant outputs. Moving forward, our research trajectory aims to delve into leveraging Multiview geometry as a training data source for deep pose regressors and exploring probabilistic extensions to this algorithm. Moreover, we recognize the finite capacity of neural networks to localize within a physical area, an aspect we intend to explore further in subsequent investigations.

Conclusion

The salient advantage of the described approach lies in its ability to swiftly and autonomously furnish high-probability detections grounded on conservative estimations of key point parameters in image transformations. By deriving a succinct list of the most probable pose positions, 3D poses, and sizes, the algorithm sets the stage for subsequent refinement to ascertain the correct object parameters. Moreover, by unfailingly obtaining the most probable pose for each image position, the parameter space undergoes significant reduction, facilitating the adoption of more sophisticated techniques to identify objects even when the highest convolution results fail to match the object. Thus, the algorithm, as presently formulated, serves as a potent tool for expeditious parameter space reduction, capable of directly estimating poses or initiating a refined search using alternative methods. Further exploration is warranted to unlock the full potential of this approach and delineate its constraints.

Future Scope

1.Pose estimation harbors manifold applications across diverse fields. In robotics, for instance, it enables robots to mimic human actions, obviating the need for manual trajectory programming by computing articulator movements to replicate human actions.
2.Security and surveillance systems stand to benefit from pose estimation's augmentation, facilitating functions such as fall detection or identification of medical emergencies.
3.Pose estimation technology finds utility in interactive gaming applications like Kinect, leveraging 3D pose estimation to track human player motions, thereby enhancing gaming immersion.
4.There is immense potential for applications enhancing security and surveillance through pose estimation technology.

Conflict Of Interest

The authors declare that they have no conflict of interest.

Funding Support

The author declare that they have no funding support for this study.

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

[5] X. Nie, J. Feng, J. Xing, and S. Yan, "Pose partition networks for Multi-person pose estimation," in ECCV, 2018

[6] V. Belagiannis and A. Zisserman. Recurrent human pose estimation. In Proc. IEEE Int. Automatic Face & Gesture Recognition, 2017.

[7] OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields 30 May 2019.

[8] "Adversarial PoseNet: A Structure- aware Convolutional Network for Human Pose Estimation" Yu Chen Chunhua Shen Xiu Shen Wei Lingqiao Liu Jian Yang Nanjing University of Science and Technology University of Adelaide.

[9] Deep High-Resolution Representation Learning for Human Pose Estimation 25 Feb 2019/

[10] W. Yang, W. Ouyang, H. Li, and X. Wang. End-to end learning of deformable mixture of parts and deep convolutional neural networks for human pose estimation. In CVPR, pages 3073–3082, 2016.