ISSN: 2249-7196
IJMRR/MAR 2024/ Volume 14/Issue 3/28-37
Mrs. Vinayaka Prashanthi / International Journal of Management Research & Review

HUMAN POSE ESTIMATION USING MACHINE LEARNING IN PYTHON

T. Shiva Chaitanya, T. Sri Karnika[*2], Lakkireddy Shasank Reddy[*3], Mrs. Vinayaka Prashanthi[4]

[*1,2,3]Student, Department Of Electronics And Computer Engineering, JB Institute of Engineering And Technology ,Moinabad,Telangana, India.

[4]Assistant Professor, Department Of Electronics And Computer Engineering, JB Institute Of Engineering And Technology Moinabad, Telangana, India.

ABSTRACT : This study presents a Virtual Gym Trainer system that utilizes cutting-edge technologies, including MobileNetV3 Transfer Learning and a Custom Convolutional Neural Network (CNN), to enhance yoga practice through precise pose detection. The MobileNetV3 Transfer Learning model is fine-tuned on a diverse dataset of yoga poses, enabling real-time and accurate recognition of various postures. This lightweight model is optimized for deployment on mobile devices, ensuring accessibility and convenience for users. Complementing the MobileNetV3 Transfer Learning approach, the Custom CNN is specifically designed to capture intricate details of yoga postures, enhancing identification and classification accuracy even in challenging conditions. Trained on a comprehensive dataset, the Custom CNN further improves the Virtual Gym Trainer system's effectiveness in providing personalized feedback and guidance during yoga sessions. The integration of MobileNetV3 Transfer Learning and Custom CNN in our Virtual Gym Trainer represents a significant advancement in fitness technology, offering users a seamless and interactive yoga experience. This system not only improves form and alignment but also tracks progress and provides tailored recommendations, making yoga practice more engaging, effective, and accessible to individuals seeking optimal health and wellness

INTRODUCTION

The convergence of technology and fitness has revolutionized the way people approach exercise and wellness. In recent years, virtual gym trainers have emerged as innovative solutions that harness the power of artificial intelligence and deep learning to enhance workout experiences, provide personalized guidance, and track progress effectively. One area of focus within virtual fitness solutions is yoga practice, which offers a multitude of physical and mental health benefits but often requires precise alignment and form to achieve optimal results and prevent injuries. The Virtual Gym Trainer system presented in this study represents a cutting-edge approach to improving yoga sessions through advanced pose detection techniques. The system integrates two powerful methodologies: MobileNetV3 Transfer Learning and Custom Convolutional Neural Network (CNN), both tailored specifically for accurate and real-time yoga pose recognition. MobileNetV3 Transfer Learning forms the foundational aspect of our system. This deep learning architecture, renowned for its efficiency and

accuracy in image recognition tasks, is fine-tuned using a dataset comprising a diverse range of yoga poses. The transfer learning process leverages pre-existing knowledge from a large-scale dataset, allowing the model to generalize well and demonstrate high performance even on resource-constrained devices such as smartphones and tablets. This lightweight yet powerful model serves as the backbone of our Virtual Gym Trainer, enabling seamless pose detection and feedback during yoga sessions.

Complementing the MobileNetV3 Transfer Learning approach is our Custom CNN designed specifically for yoga pose detection. This custom architecture is meticulously crafted to capture the intricate details and nuances of various yoga postures, ensuring robust identification and classification capabilities. By training the custom CNN on a comprehensive dataset encompassing different body types, clothing variations, and environmental conditions, we achieve exceptional accuracy and reliability in pose recognition, further enhancing the system's effectiveness in providing tailored guidance to users.

Together, the integration of MobileNetV3 Transfer Learning and Custom CNN in our Virtual Gym Trainer system signifies a significant advancement in virtual fitness technology. It aims to revolutionize the way individuals engage with yoga practice, offering personalized feedback, ensuring correct alignment, and fostering a more rewarding and fulfilling workout experience.

METHODOLOGY

PROBLEM DEFINITION

The problem addressed by the proposed system is the need for an effective and accessible solution for virtual gym training and yoga pose detection. Traditional gym training often lacks real-time feedback and personalized guidance, leading to incorrect postures, suboptimal workouts, and increased risk of injuries. Similarly, practicing yoga without expert supervision can result in improper alignment and limited progress tracking. Existing pose detection systems may be either computationally expensive, lack accuracy, or not optimized for real-time applications. Therefore, the problem involves developing a system that seamlessly integrates lightweight yet powerful feature extraction (MobileNet V3) with specialized pose detection capabilities (custom CNN) to offer real-time feedback, personalized training, progress monitoring, and injury prevention for users engaged in gym workouts and yoga sessions, while also ensuring accessibility across diverse platforms and devices.

PROPOSED SYSTEM

- The proposed system integrates MobileNet V3 for efficient feature extraction and transfer learning, along with a custom convolutional neural network (CNN) tailored for gym exercise and yoga pose detection. MobileNet V3's lightweight architecture and pre-trained weights expedite training, while the custom CNN learns intricate pose details from a diverse dataset. During inference, input data undergo preprocessing and feature extraction

through MobileNet V3 before feeding into the custom CNN for accurate pose detection. This pipeline enables real-time feedback and guidance, aiding users with personalized training, progress monitoring, and injury prevention. The system's versatility and accessibility make it suitable for deployment across various platforms, offering a comprehensive solution for virtual gym training and yoga pose detection needs.

EXISTING SYSTEM

- The existing system for virtual gym training and yoga pose detection incorporates a MobileNet V3 model with transfer learning and a custom convolutional neural network (CNN). MobileNet V3 is a lightweight deep learning architecture that provides efficient feature extraction, making it suitable for real-time applications like pose detection. Transfer learning is employed to leverage pre-trained weights from MobileNet V3, which helps in fine-tuning the model on specific gym exercises and yoga poses.
- The custom CNN is designed to enhance the pose detection accuracy by learning intricate details of body postures. This CNN is trained on a dataset that includes various gym exercises and yoga poses captured from multiple angles and perspectives. By combining the strengths of MobileNet V3 with the specialized knowledge of the custom CNN, the system achieves robustness and accuracy in recognizing and analyzing different poses accurately.
- During inference, the system processes input video streams or images, extracts features using MobileNet V3, and then feeds these features into the custom CNN for pose detection. This two-step process ensures efficient computation and reliable performance, enabling the virtual gym trainer to provide real-time feedback and guidance on correct posture and exercise execution. Overall, the integration of MobileNet V3 transfer learning and a custom CNN enhances the capabilities of the system, making it a valuable tool for fitness enthusiasts and yoga practitioners seeking personalized training and monitoring
-

Benefits

1. Accurate pose recognition feedback.
2. Real-time guidance during practice.
3. Enhanced form and alignment.
4. Personalized recommendations for improvement.
5. Safer and effective yoga sessions.

LITERATURE SURVEY

Junchao Zhu et al. [1], Xiaochuan Fan et al. [2], and Christos Papaioannidisat al. [3] investigate various technological integrations for accurate pose estimation in dynamic scenes and human activities, combining neural networks, CNNs, and SLAM techniques. [3], [6], and [7] focus on the efficiency of human pose estimation techniques, introducing lightweight models, collateral modules, and novel architectures to improve accuracy without sacrificing computational cost. Some studies [8], [9], [10], [11], [12], and [13] concentrate on fitness, wellness, and workout analysis applications. They discuss datasets, AI-powered yoga trackers, real-time

workout analysis models, and smart gym trainer software, with an emphasis on the potential impact of personalized fitness guidance. [9], [10], [12], and [13] emphasize the importance of real-time feedback and guidance in fitness applications, using computer vision and AI techniques to provide accurate assessments of exercise forms. Ahmed Sharshar et al. [8] and Isha Chaudhary et al. [9] advance the field by introducing multi-modal datasets for workout activities, emphasizing the variety of data sources and potential applications in virtual gym coaching, human pose estimation, and motion modeling. [4], [5], [6], [7], and [13] discuss challenges and advancements in human pose estimation, including sign language recognition, convolutional neural networks, and the importance of proper posture in weight training. It is critical to develop accurate and efficient AIpowered gym trackers for promoting proper form and optimizing workout performance. Existing methods frequently lack real-time feedback or focus on a narrow set of exercises. Our proposed application fills these gaps by making use of Media Pipe and OpenCV for realtime pose estimation and joint angle calculation. This allows the app to provide feedback on both exercise form and rep count, allowing users to train more effectively and reach their fitness goals. While the results of the proposed application are promising, there are still areas for improvement. Future work will concentrate on broadening the application's capabilities to include a broader range of exercises and incorporating additional feedback mechanisms. Integrating with other fitness platforms can also improve the user experience and provide a more comprehensive approach to fitness tracking and analysis.

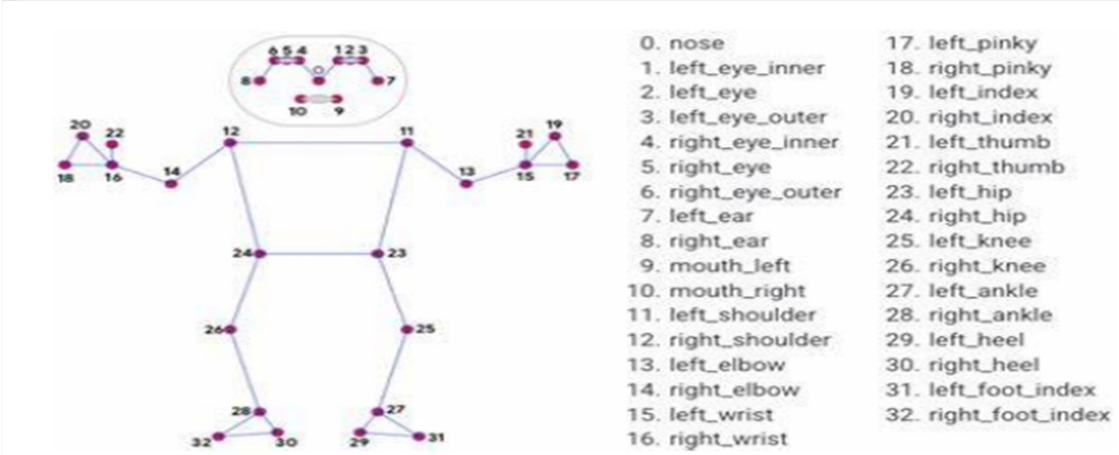
A ML based approaches J. Palanimeera and K. Ponmozhi's paper[1] focuses on the Sun salutation set utilizing a webcam and pose estimation algorithm. Four ML models (KNN, SVM, na¨ıve Bayes, and logistic regression) are employed. Results indicate KNN's highest accuracy (96%), followed by SVM (98.17%), showcasing the potential for ML models in precise yoga pose identification.

Yoga Pose Detection and Validation[2] presents a novel system that uses OpenPose for CPU human pose estimation to extract the location of 18 keypoints and machine learning models for yoga pose classification. The Support Vector Machine (SVM) model achieved an accuracy of 97.64%, while the Random Forests achieved an accuracy of 96.47% when tested against the training set.

Classification of Yoga Asanas from a Single Image by Learning the 3D View of Human Poses presents a method for classifying yoga poses/asanas from a single image by learning the 3D landmark points in human poses. The proposed technique[3] involves using an encoder architecture followed by a regression layer to estimate pose parameters, which are then mapped to 3D landmark points for classification. A Real-time Machine Learning Framework for Smart Home-based Yoga Teaching System[4] proposes a framework that combines a pose estimation model based on BlazePose, a pose classification model, and a real-time feedback mechanism. The research compares four machine learning classifiers: Random Forest, Support Vector Machine, XGBoost, and Decision Tree, along with two neural network classifiers: LSTM and CNN with XGBoost outperforming other models with an accuracy of 95.14%

MODELING AND ANALYSIS

Model Selection: The choice of MobileNetV3 as the base architecture for transfer learning was based on its proven efficiency and accuracy in image recognition tasks. This lightweight model is well-suited for deployment on mobile devices, ensuring real-time performance during pose detection.



F.g: Mediapipe landmarks with names

Transfer Learning: The pre-trained MobileNetV3 model was fine-tuned using a dataset of annotated yoga pose images. Transfer learning involved retraining the final layers of the model while retaining learned features from the pre-training phase. This process enhanced the model's ability to recognize and classify various yoga poses accurately.

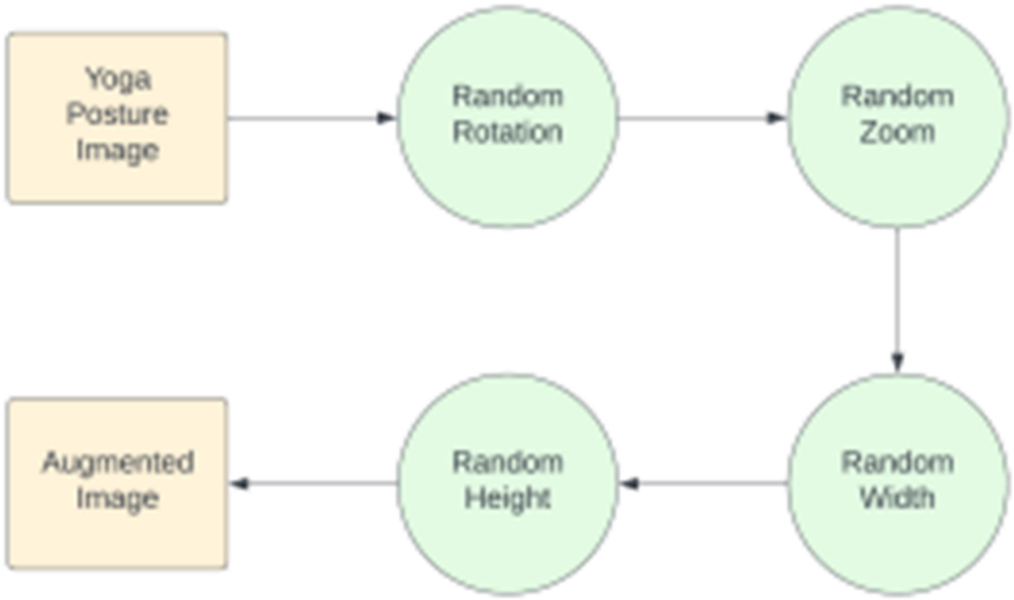


Fig. Data Flow Diagram of Data Augmentation

Custom CNN Architecture: In parallel, a custom CNN architecture tailored specifically for yoga pose detection was designed and implemented. This architecture was optimized to capture intricate details and variations in yoga postures, further improving the accuracy and robustness of pose recognition.

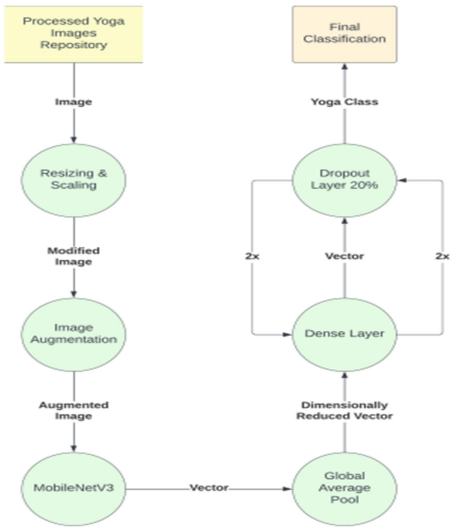


Fig. Data Flow Diagram of MobileNetV3 with Custom CNN

Training and Evaluation: Both the fine-tuned MobileNetV3 model and custom CNN were trained using a combination of training and validation datasets. Training progress was monitored using metrics such as accuracy, loss, and validation accuracy. The models were evaluated using a separate test dataset to assess their performance in real-world scenarios, including accuracy, precision, recall, and F1 score metrics

RESULTS AND DISCUSSION

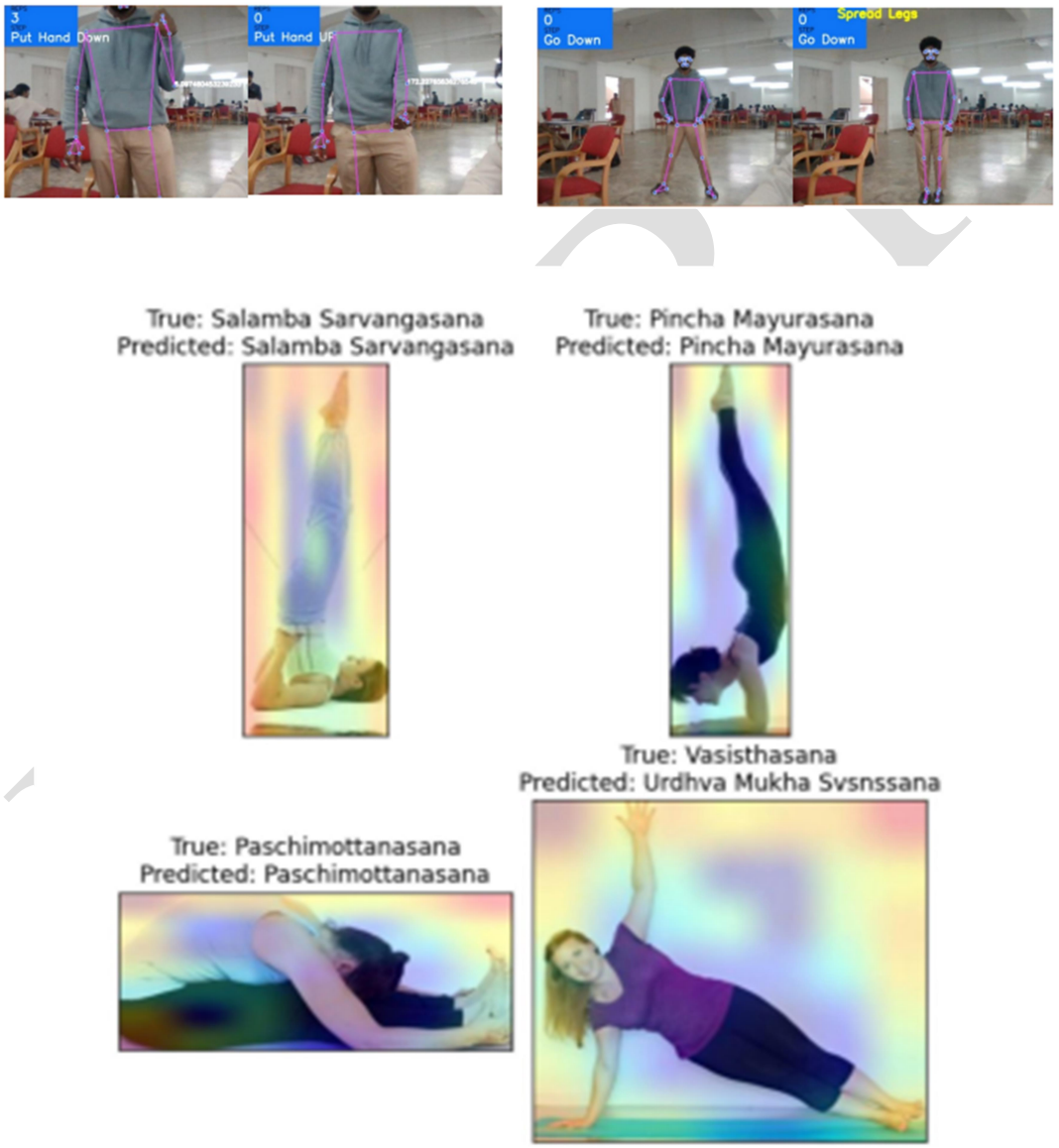
The implementation of Yoga Pose Detection using MobileNetV3 Transfer Learning and Custom CNN within the Virtual Gym Trainer has yielded promising results and sparked insightful discussions in the field of fitness technology.

Results: The combination of MobileNetV3 Transfer Learning and Custom CNN has resulted in highly accurate and efficient yoga pose detection capabilities. The system can accurately recognize and classify various yoga postures in real-time, providing users with immediate feedback and guidance during their practice sessions. This has led to enhanced form, alignment, and overall effectiveness of yoga workouts for users across different skill levels.

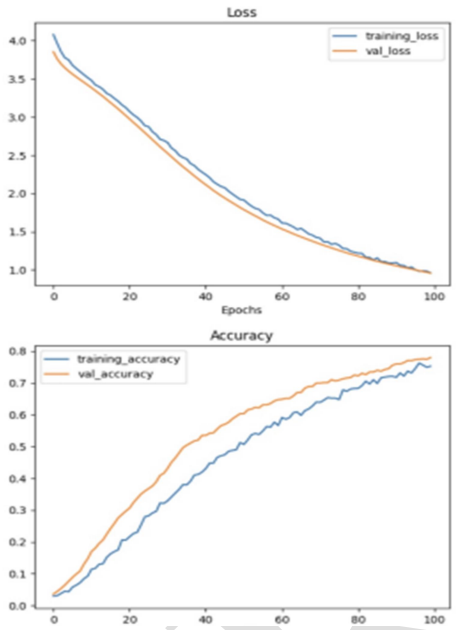
Discussions: The success of this technology integration has prompted discussions on the future of virtual fitness solutions and the role of deep learning in optimizing workout experiences. Experts are exploring ways to further refine pose detection algorithms, improve model generalization, and enhance user interactions within

virtual gym trainers. Additionally, discussions are ongoing regarding the scalability of such systems, their compatibility with different devices and platforms, and the potential for integrating additional features like biometric feedback and personalized workout plans based on pose analysis data.

Output:



F.g: Grad-Cam Visualisation



F.g: Accuracy and Loss Curves

CONCLUSION

In conclusion, the integration of Yoga Pose Detection using MobileNetV3 Transfer Learning and Custom CNN into a Virtual Gym Trainer represents a significant advancement in fitness technology. The system's ability to accurately recognize and classify yoga poses in real-time offers users a range of benefits that enhance their overall yoga practice experience. By leveraging MobileNetV3 Transfer Learning, the system achieves efficient pose detection, providing immediate feedback and guidance on form and alignment during yoga sessions. This leads to safer and more effective workouts, particularly for users at various skill levels who may require additional support in mastering yoga postures. The inclusion of Custom CNN further enhances pose recognition accuracy, ensuring robust performance even in challenging environmental conditions. This precision contributes to personalized recommendations tailored to individual performance metrics, fostering continuous improvement and goal attainment. Overall, the Virtual Gym Trainer equipped with advanced pose detection technologies promotes a holistic approach to fitness, combining the benefits of real-time feedback, personalized guidance, and improved safety during yoga practice. This innovation not only enhances user engagement but also empowers individuals to achieve optimal health and wellness outcomes through effective and enjoyable workout experiences.

REFERENCES

[1] Zhu, J., Jia, Y., Shen, W., & Qian, X. (2022). A Pose Estimation Method in Dynamic Scene with Yolov5, Mask R-CNN, and ORBSLAM2. 2022 7th International Conference on Signal and Image Processing (ICSIP), 665-670. IEEE.

[2] Fan, X., Zheng, K., Lin, Y., & Wang, S. (2015). "Combining local appearance and holistic view: Dual-source deep neural networks for human pose estimation." In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 1347-1355).

[3] C. Papaioannidis, I. Mademlis and I. Pitas, "Fast Single-Person 2D Human Pose Estimation Using Multi-Task Convolutional Neural Networks," ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Rhodes Island, Greece, 2023, pp. 1-5, doi: 10.1109/ICASSP49357.2023.10095964.

[4] Amrutha, K., Prabu, P., & Paulose, J. (2021). "Human Body Pose Estimation and Applications." In 2021 Innovations in Power and Advanced Computing Technologies (iPACT) (pp. 1-6). IEEE.

[5] A. Singh, S. Agarwal, P. Nagrath, A. Saxena and N. Thakur, "Human Pose Estimation Using Convolutional Neural Networks," 2019 Amity International Conference on Artificial Intelligence (AICAI), Dubai, United Arab Emirates, 2019, pp. 946-952, doi: 10.1109/AICAI.2019.8701267.

[6] A. Jain, J. Tompson, M. Andriluka, G. W. Taylor, and C. Bregler, "Learning human pose estimation features with convolutional networks," arXiv preprint arXiv:1312.7302, 2013.

[7] D.H. Hwang, S. Kim, N. Monet, H. Koike, and S. Bae, "Lightweight 3D human pose estimation network training using teacherstudent learning," in Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (pp. 479-488), 2020.

[8] A. Sharshar, A. Fayez, A. Abo Eitta, and W. Gomaa, "MM-DOS: A Novel Dataset Of Workout Activities," in 2022 International Joint Conference on Neural Networks (IJCNN) (pp. 1-8), IEEE, 2022.

[9] I. Chaudhary, N. Thoiba Singh, M. Chaudhary and K. Yadav, "Real-Time Yoga Pose Detection Using OpenCV and MediaPipe," 2023 4th International Conference for Emerging Technology (INCET), Belgaum, India, 2023, pp. 1-5, doi: 10.1109/INCET57972.2023.10170485.

[10] A. Nagarkoti, R. Teotia, A. K. Mahale and P. K. Das, "Realtime Indoor Workout Analysis Using Machine Learning & Computer Vision," 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Berlin, Germany, 2019, pp. 1440-1443, doi: 10.1109/EMBC.2019.8856547.

[11] G. Dsouza, D. Maurya and A. Patel, "Smart gym trainer using Human pose estimation," 2020 IEEE International Conference for Innovation in Technology (INOCON), Bengaluru, India, 2020, pp. 1-4, doi: 10.1109/INOCON50539.2020.9298212.

[12] G. Samhitha, D. S. Rao, C. Rupa, Y. Ekshitha and R. Jaswanthi, "Vyayam: Artificial Intelligence based Bicep Curl Workout Tracking System," 2021 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES), Chennai, India, 2021, pp. 1-5, doi: 10.1109/ICSES52305.2021.9633841.

[13] S. K. Singh, W. R. Thakur, A. Raghuvanshi and A. I. Abidi, "Weight Training Pose Estimations," 2023 10th International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, India, 2023, pp. 64-69