Available online at [*http://www.advancedscientificjournal.com*](http://www.advancedscientificjournal.com)
[*http://www.krishmapublication.com*](http://www.krishmapublication.com)
IJMASRI, Vol. 2, issue 12, pp.787- 791, Dec. -2022
https://doi.org/10.53633/ijmasri

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY ADVANCED SCIENTIFIC RESEARCH AND INNOVATION (IJMASRI)

ISSN: 2582-9130 **IBI IMPACT FACTOR 1.5** **DOI: 10.53633/IJMASRI**

RESEARCH ARTICLE

YOGA POSE ESTIMATION

Ujjwal Rastogi[1] and Mohd Sarim[2]

[1,2]*Department of Information Technology, Maharaja Agrasen Institute of Technology, Rohini, Delhi*

Abstract

Human pose estimation is a deep-rooted problem in computer vision that has exposed many challenges in the past. Analyzing human activities is beneficial in many fields like video- surveillance, biometrics, assisted living, at-home health monitoring etc. With our fast-paced lives these days, people usually prefer exercising at home but feel the need of an instructor to evaluate their exercise form. As these resources are not always available, human pose recognition can be used to build a self-instruction exercise system that allows people to learn and practice exercises correctly by themselves. This project lays the foundation for building such a system by discussing various machine learning and deep learning approaches to accurately classify yoga poses on pre-recorded videos and also in real-time. The project also discusses various pose estimation and key point detection methods in detail and explains different deep learning models used for pose classification.

Keywords: Yoga Pose Estimation, Machine Learning, Artificial Neural Network, Mediapipe

Introduction

The art of practicing yoga helps in controlling an individual's mind, body and soul. It brings together physical and mental disciplines to achieve a peaceful body and mind; it helps manage stress and anxiety and keeps you relaxing. It also helps in increasing flexibility, muscle strength and body tone. It improves respiration, energy and vitality. Practicing yoga might seem like just stretching, but it can do much more for your body from the way you feel, look and move.

Media Pipe

Media Pipe is an open source library by google, which uses deep learning to localize and detect objects and then generates a 33-point embedding which are the landmarks. Media Pipe does detection and tracking for efficiency. images-use detector again and again but for video it’s good to use tracking. It first localizes the object then outputs the landmarks. Media Pipe Pose is a ML solution for high-fidelity body pose tracking, inferring 33 3D landmarks and background segmentation mask on the whole body from RGB video frames utilizing our Blaze Pose research that also powers the ML Kit Pose Detection API. Current state-of-the-art approaches rely primarily on powerful desktop environments for inference, whereas our method achieves real-time performance on most modern mobile phones, desktops/laptops, in python and even on the web.

Human pose estimation from video plays a critical role in various applications such as quantifying physical exercises, sign language recognition, and full-body gesture control. For example, it can form the basis for yoga, dance, and fitness applications. It can also enable the overlay of digital content and information on top of the physical world in augmented reality.

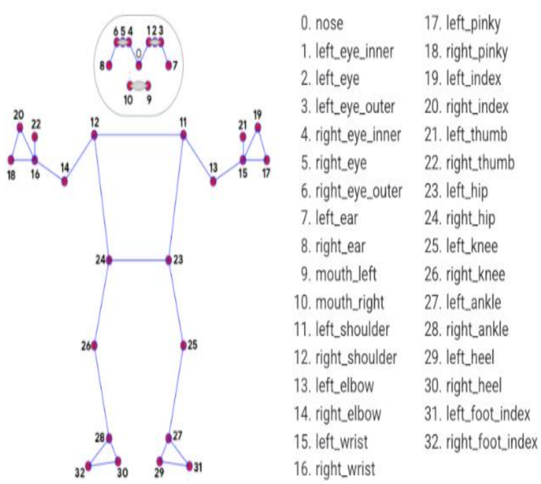


Fig. 1 List of 33 Landmarks detected by media pipe

ARTIFICIAL NEURAL NETWORK

An ANN is based on a collection of connected units or nodes called artificial neurons, which loosely model the neurons in a biological brain. Each connection, like the synapses in a biological brain, can transmit a signal to other neurons. An artificial neuron receives signals then processes them and can signal neurons connected to it. The "signal" at a connection is a real number, and the output of each neuron is computed by some non-linear function of the sum of its inputs. The connections are called edges. Neurons and edges typically have a weight that adjusts as learning proceeds. The weight increases or decreases the strength of the signal at a connection. Neurons may have a threshold such that a signal is sent only if the aggregate signal crosses that threshold. Neural networks learn (or are trained) by processing examples, each of which contains a known "input" and "result," forming probability-weighted associations between the two, which are stored within the data structure of the net itself. The training of a neural network from a given example is usually conducted by determining the difference between the processed output of the network (often a prediction) and a target output. This difference is the error. The network then adjusts its weighted associations according to a learning rule and using this error value. Successive adjustments will cause the neural network to produce output which is increasingly similar to the target output. After a sufficient number of these adjustments the training can be terminated based upon certain criteria. This is known as supervised learning.

Such systems "learn" to perform tasks by considering examples, generally without being programmed with task-specific rules. For example, in image recognition, they might learn to identify images that contain cats by analyzing example images that have been manually labeled as "cat" or "no cat" and using the results to identify cats in other images. They do this without any prior knowledge of cats, for example, that they have fur, tails, whiskers, and cat-like faces. Instead, they automatically generate identifying characteristics from the examples that they process.

- **System Architechture**

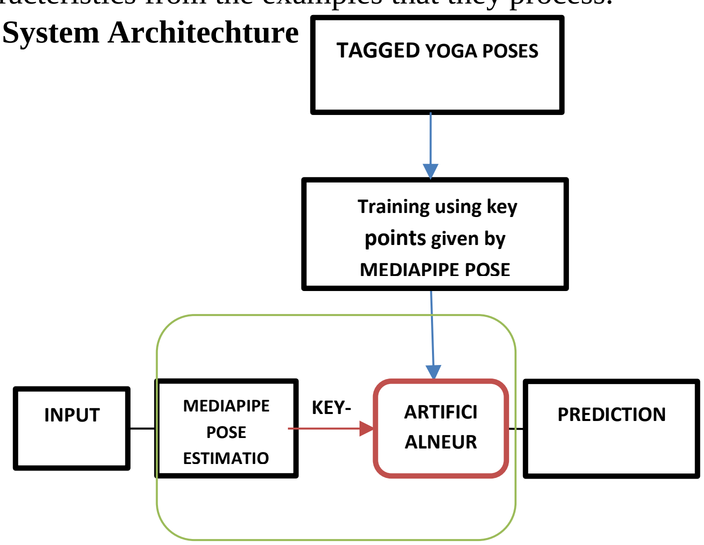


Fig. 2 System Architecture

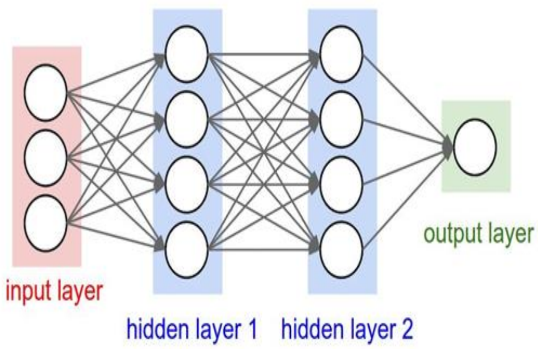


Fig. 3 Artificial Neural Network

Methodology

Data Collection

The dataset used in this project was from Kaggle website. Our dataset contains 107 Yoga poses.

Data Preprocessing

Because the input for our pose classifier is the output landmarks from the mediapipe model, we need to generate our training dataset by running labeled images through mediapipe and then capturing all the landmark data and ground truth labels into a CSV file.

Training And Inference

Model 1 calculates a pose embedding (a.k.a feature vector) from the detected landmark coordinates.

Feature Engineering

Moving the pose center to the origin. Scaling the pose so that the pose size becomes 1.
Flattening these coordinates into a feature vector
Model 2 feeds pose embedding through several Dense layer to predict the pose class.

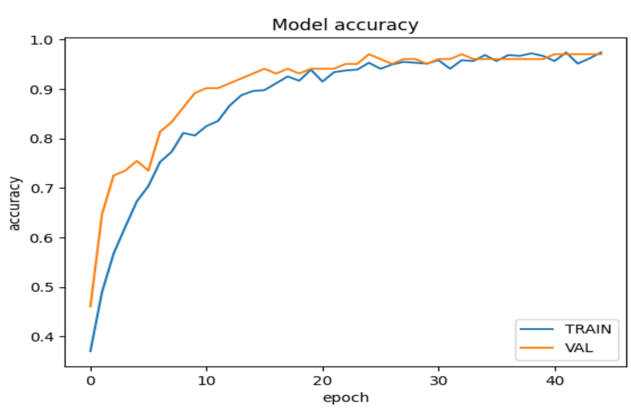


Fig. 4 Accuracy vs epochs graph

Max accuracy during training reached was 0.9740

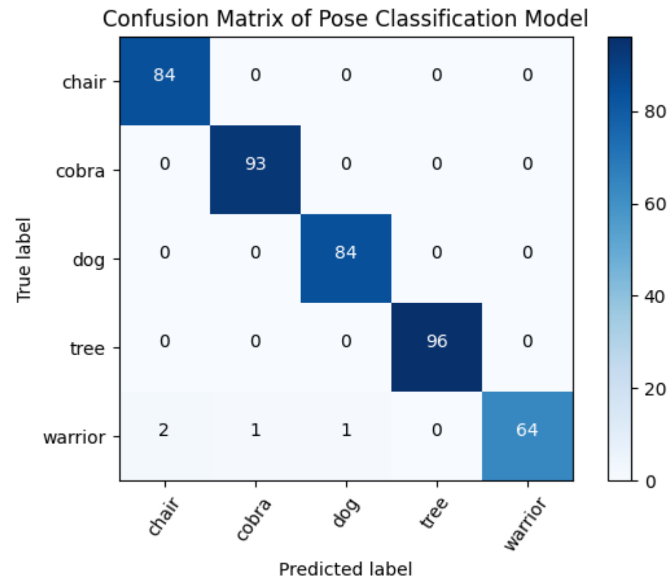


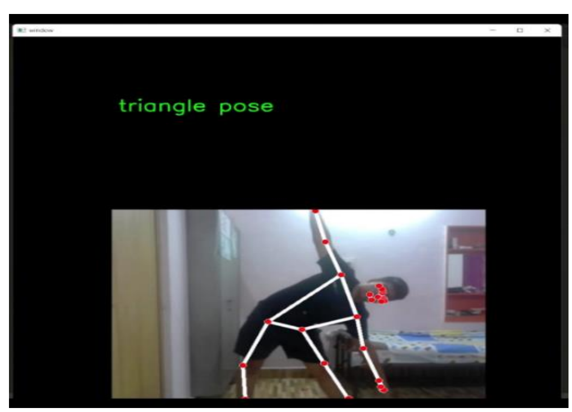
Fig. 5 Confusion Matrix

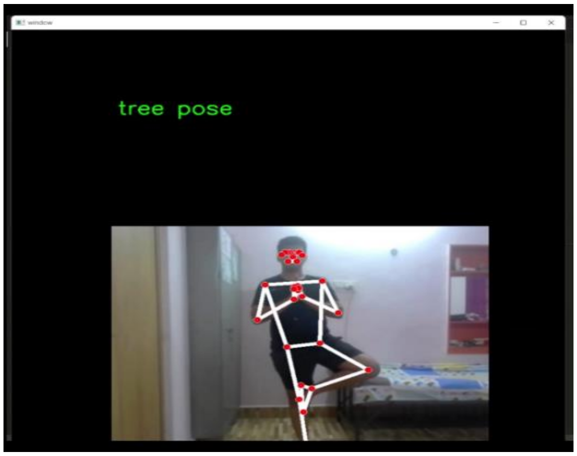
Result

On evaluating the model on test dataset:
loss: 0.0424 - accuracy: 0.9906

SN.	Machine Learning Model	Accuracy	Precision	Recall	F-1 Score
1	Logistic regression	0.94	0.94	0.94	0.94
2	SVM classifier	0.93	0.94	0.93	0.93
3	Random Forest classifier	0.89	0.90	0.89	0.89
4	Knn classifier	0.93	0.94	0.93	0.93
5	Naïve Bayes classifier	0.91	0.93	0.89	0.91

Fig. 6 List of earlier used models for classification.





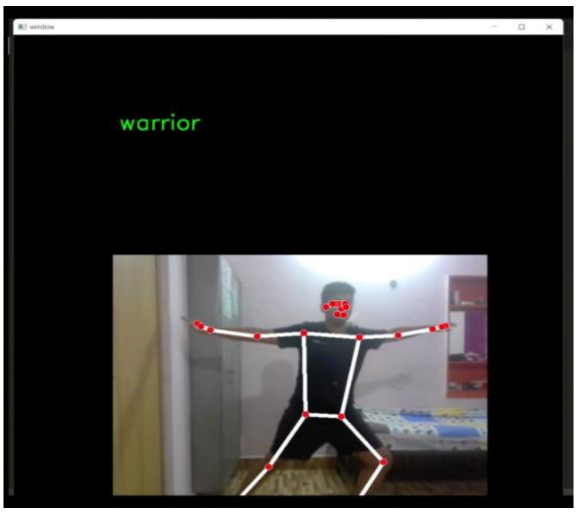


Fig. 7 Correctly detected Poses

Reference

1. Guddeti, R.R., Dang, G. Williams, M.A. Alla, V.M. (2018). Role of Yoga in cardiac disease and rehabilitation. J Cardiopulm Rehabil Prev. https://doi.org/10.1097/hcr.0000000000000037.
2. Dr. Gaba, Importance of Yoga in the Pandemic,https://sarvodayahospital.com/blog/importance-ofyoga-during-this-pandemic.
3. Palanimeera, J and Ponmozhi, K. (2020). Classification of yoga pose using machine learning techniques Author: https://doi.org/10.1016/j.matpr.2020.08.700.
4. Ajeet Ram Pathak, Manjusha Pandey, Siddharth Rautaray. (2018). Application of Deep Learning for Object Detection, Procedia Computer Science, Volume 132, 2018, Pages 1706-1717, ISSN: 1877-0509, https://doi.org/10.1016/j.procs.2018.05.144. (https://www.sciencedirect.com/science/article/pii/S1877050918308767).
5. Shruti Kothari. (2020). Yoga Pose Classification Using Deep Learning Ph.D. thesis SAN JOSE STATE UNIVERSITY Modernization.
6. Agrawal, Y., Shah, Y and Sharma, A.(2020). "Implementation of Machine Learning Technique for Identification of Yoga Pose," 2020 IEEE 9th International Conference on Communication Systems and Network Technologies (CSNT), 2020, pp. 40-43, DOI: 10.1109/CSNT48778.2020.9115758.
7. Narayanan, S., Sankara and Misra. (2021). Devendra Kumar and Arora, Kartik and Rai, Harsh, Yoga Pose Detection Using Deep Learning Techniques (May 10, 2021). Proceedings of the International Conference on Innovative Computing & Communication (ICICC) 2021, Available at SSRN: https://ssrn.com/abstract=3842656 or http://dx.doi.org/10.2139/ssrn.3842656.
8. Kumar., Deepak and Sinha, Anurag. (2020). Yoga Pose Detection and Classification Using Deep Learning. International Journal of Scientific Research in Computer Science Engineering and Information Technology. 10.32628/CSEIT206623.
9. Nandakishore Joshi. (2020), August 17, REALTIME 2D YOGA POSE ESTIMATION WITH CODE WALKTHROUGH, Medium, https://nandakishorej8.medium.com/realtime-2d-yoga-pose-estimationwith-code-walk-through-cfd69262d356.
10. Syed Abdul Gaffar Shakhadri. (2021). May 18. Pose Estimation using OpenCV, Analytics Vidhya, https://www.analyticsvidhya.com/blog/2021/05/pose-estimation-using-opencv/ Cramer, Jan Salomon. "The origins of logistic regression." (2002): 4.
11. Bazarevsky. Valentin. (2020). "BlazePose: On-device Real-time Body Pose tracking." arXiv preprint arXiv:2006.10204 (2020).
12. Pauzi, A.S.B. (2021) Movement Estimation Using Mediapipe BlazePose. In: Badioze

Zaman H. et al. (eds) Advances in Visual Informatics. IVIC 2021. Lecture Notes in Computer Science, vol 13051. Springer, Cham. https://doi.org/10.1007/978-3-030-90235-3_49

*****