

# CMU Based Aggressive Approach to Yoga Pose Estimation

Prof.Sonali Muley<sup>1</sup>, Prateek Mahajan<sup>2</sup>, Pratik Medidar<sup>3</sup>, Harish Chavan<sup>4</sup>, Prashant Patil<sup>5</sup>

Assistant Professor, Department of Computer Engineering, Marathwada Mitra Mandal's Institute of Technology,  
Pune, India<sup>1</sup>

Student, Department of Computer Engineering, Marathwada Mitra Mandal's Institute of Technology, Pune, India<sup>2</sup>

Student, Department of Computer Engineering, Marathwada Mitra Mandal's Institute of Technology, Pune, India<sup>3</sup>

Student, Department of Computer Engineering, Marathwada Mitra Mandal's Institute of Technology, Pune, India<sup>4</sup>

Student, Department of Computer Engineering, Marathwada Mitra Mandal's Institute of Technology, Pune, India<sup>5</sup>

**ABSTRACT:** We propose a technique to almost accurately correct the human pose while performing yoga asanas. This proposed system maintains high accuracy while achieving real-time performance in almost cases. Yoga has a wide variety of asanas and here comes the role of the angle between body parts. The cosine similarity technique is used to consider the variance of the angle with original values. Critical angles come with a critical combination of angles and hence multiple dimensions need to be considered. This system finds the variance between actual and target positions and corrects the user by providing proper guidance with text to speech output in real-time to correct the selected pose. In this project we use human pose estimation for estimating the Yoga pose of the individual using computer vision technique and Open pose (open source library). All human joints will get highlighted and will be joined using greedy techniques. Euclidian angle between the body parts will be calculated and matched with the template angles, i.e. the Individual will be indicated about wrong pose using text to speech.

**KEYWORDS:** Open pose, Pose net, CNN.

## I. INTRODUCTION

In computer vision field localization of human body joints can be referred to as a problem of human pose estimation. For estimation of human pose there could be some of the challenges as joints could be small or barely visible and to differentiate between adjacent joints, or if there are multiple persons present in the image or video frame. The work in the field of human pose estimation is focused on detecting all possible poses, for this a variety of models have been proposed. Pose estimation could have used in animation or in fitness, it could be used in interactive installation in augmented reality. In this work, we tried to implement a pose estimation for the correction of yoga poses. many pose detection systems are open source which requires system setup and cameras. The various model for pose estimation can detect only one person and other versions could detect multiple people. The proposed version of the model could detect a single person which is simpler.

## II. RELATED WORK

[1]"Real-time multi-person 2D pose estimation using part affinity fields", CVPR 2017, They proposed a nonparametric approach called as part affinity fields. In the approach key-points of the human body are referred to as 'parts' which focused on locating body joints. The approach is used where multiple persons are present in the image and use confidence maps for identification of parts in any 2D.

[2] "Convolutional pose machines", CVPR APR 2016, Convolutional pose machines are developed by inheriting the advantages of the pose machines. CPM developed employing convolutional networks to produce belief maps. Belief maps can be used to identify the location of a particular part. Images with body joints are compared based on belief maps and key-point localization was carried out. The process was carried out on the Leeds sports pose dataset.

[3]"Deep-pose: human pose estimation via deep neural networks", IEEE Int. Automatic Face & Gesture Recognition, Aug 2014, Proposed a method using deep neural networks for identifying human body joints. They used DNN (Deep Neural Network) based regression model for the learning process. It uses a DNN based novel algorithm that having increased performance on the classification task.

[4]“Stacked Hourglass networks For Human Pose Estimation”, CVPR Jul 2016, Proposed a network architecture using novel convolutional neural network for estimating human pose. The network uses hourglass design to capture key-points at every scale. Local evidence is calculated to identify the features of body joints. The architecture uses layered implementation to achieve flexibility in the process.

[5]“Towards 3D Human Pose Estimation in the Wild: a Weakly-supervised Approach”, IEEE Int. Automatic Face & Gesture Recognition, Aug 2014, The proposed work is carried out in the field of 3D human pose estimation. The main challenge of work is the lack of training data due to which a weakly-supervised approach is used for training the model. The architecture uses a depth regression module on top of the 2D pose estimation module to process 3D wild images.

### III. METHODOLOGY

At significant level posture estimation occurs in two stages:

1. An information RGB picture is taken care of through a convolutional neural system.
2. Either a solitary posture or multi-present translating calculation is utilized to interpret presents, present certainty scores , key-point positions, and key-point certainty scores from the model yields.

At the most significant level, Pose Net will restore a posture object that contains a rundown of key-points and an example level certainty score for each identified individual. Posture certainty score decides the general trust in the estimation of a posture. It goes somewhere in the range of 0.0 and 1.0. It very well may be utilized to shroud represents that are not esteemed sufficient. key-point is a piece of an individual's represent that is assessed, for example, the nose, right ear, left knee, right foot, and so on. It contains both a position and a key-point certainty score. Pose Net at present recognizes 17 key points. Key-point certainty score decides the certainty that an expected key-point position is precise. It goes somewhere in the range of 0.0 and 1.0. It tends to be utilized to stow away key-points that are not esteemed sufficient. key-point position is x and y facilitates in the first info picture where a key-point has been distinguished. To get the key-points of the posture:

1. Image is convoluted to a smaller dimension for faster processing in the 'cmu' model for giving a higher performance in low hardware devices.
2. The confidence score of each joint is returned from the model. If the confidence score is greater than a threshold of  $> 0.5$  then it is considered as a validate joint else no joint detected is marked at that part. The confidence score ensures the presence of that part at that location in the image.
3. After inference of image, the joints are stored in the 'human' object. “humans.body\_parts[i]” gives the x, y coordinates also the confidence score of the particular joints if required.
4. Heat map is used to determine conflict in the joint. Conflict occurs due to more than one possible location of the joints in the image. The heat map area with a higher confidence score is considered as the correct joint position in the original image.
5. Finally, for plotting a joint in the original image its x, y coordinates are obtained from the body part object in the human object and also confidence score for considering joints importance for future process.

To Find Angle between the parts in the pose: Optimized matching strategy: Cosine distance Fig 1 shows different approaches of the same line which puts us in confusion and also the model to interpret further. To resolve this we put constraints in this as “angle  $< 270$ ”. Since no joint angle between parts goes beyond 270 degrees in poses this helps us to get over angle identification.

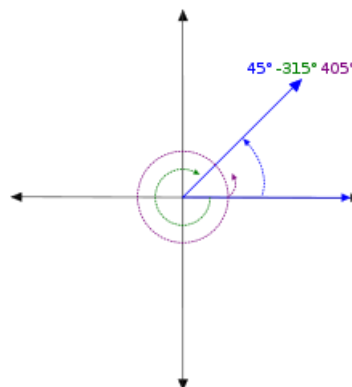


Fig1.Angle Identification

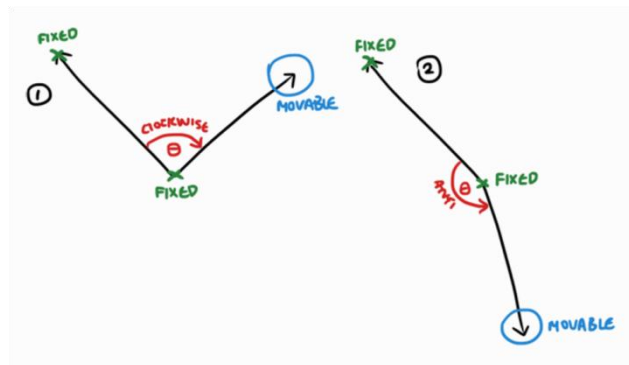


Fig2.Clockwise and anti-clockwise angles for joint

Clockwise and anti-clockwise are the two major factors to be considered which defining angle for a joint refer Fig 2. Therefore right and left part angles of the body are considered alternatively as required.

Confusion Matrix:

The confusion matrix will tell us the performance of the model for test data using which we can interpret certain decisions on the model.

- Positive (P): Output pose is correct.
- Negative (N): Output pose is wrong. □ True Positive (TP) : Output pose is correct and input pose is also correct.
- False Negative (FN) : Output pose is wrong, but input image correct.
- True Negative (TN) : Output pose is wrong and input pose is also wrong.
- False Positive (FP) : Output pose is correct, but input pose is wrong.
- Accuracy=(TP+TN)/(TP+TN+FP+FN).

The entire system architecture can be shown in Fig.3.

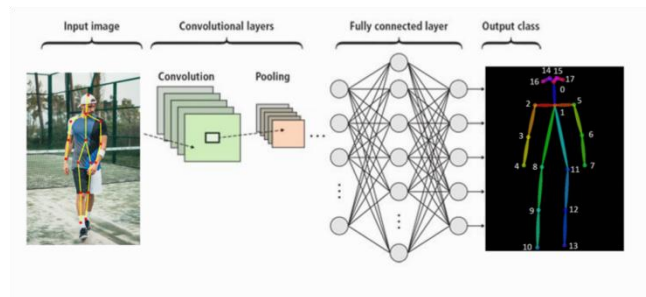


Fig3.System Architecture

#### IV. EXPERIMENTAL RESULTS

Figures shows the results of detection on points and correctness .In fig 4 and fig 5 the points are detected and connected.

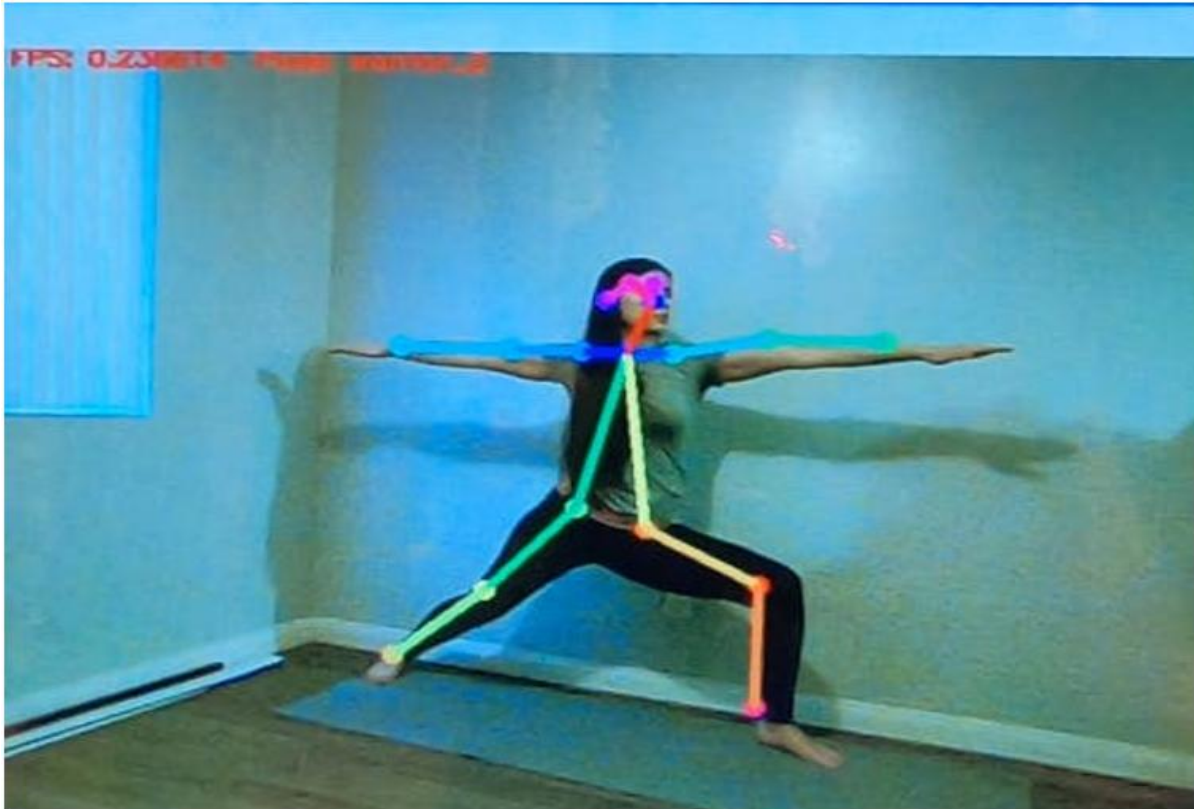


Fig 4. Plotting and connecting joints



Fig 5. Plotting and connecting joints

Now Below Fig shows the Pose correct or not. Following fig 6 is Correct and Fig 7 is wrong.



Fig 6. Identifying the pose

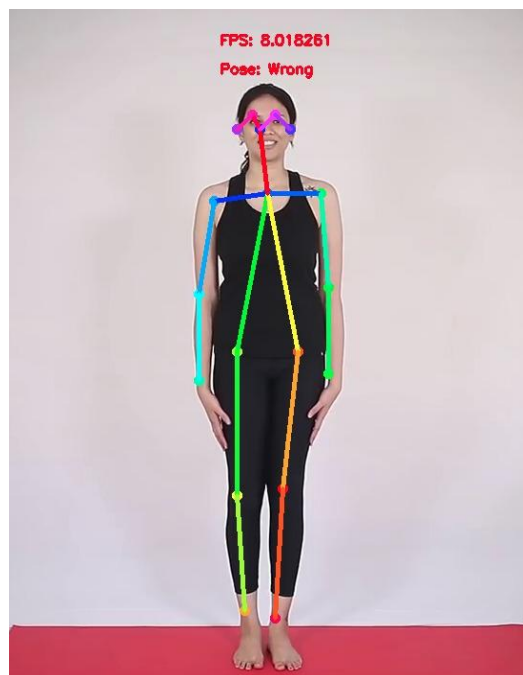


Fig 7. Identifying the pose



## V. CONCLUSION

Using Human pose estimation methods we implemented a real time pose corrector for yoga. System Classification accuracy varies between 75 percent to 96 percent.

## REFERENCES

- [1] Zhe Cao, Shih-En Wei and Yaser Sheikh. Real-time Multi-Person 2D Pose Estimation using Part Affinity Fields. In Proc. CVPR 2017.
- [2] Shih-En Wei and Varun Ramakrishna. Convolutional Pose Machines. In CVPR, Apr 2016.
- [3] Alexander Toshev and Christian Szegedy. Depose: Human Pose Estimation via Deep Neural Networks. IEEE Int. Automatic Face & Gesture Recognition, Aug 2014
- [4] Yu Chen, Chunhua Shen, Xiu-Shen Wei. Adversarial PoseNet: A Structure-aware Convolutional Network for Human Pose Estimation\*. 2D human pose estimation: new benchmark and state of the art analysis. In CVPR, 2015.
- [5] Alejandro Newell, Kaiyu Yang and Jia Deng. Stacked Hourglass Networks for Human Pose Estimation. In CVPR Jul, 2016.