International Journal of Innovative Research in Engineering
Volume 4, Issue 3(May-June 2023), PP: 467-472.
www.theijire.com ***ISSN No: 2582-8746***

Yoga Pose Detection System

Sadhana Maurya[1], Pratibha Sharma[2], Sanjal Gupta[3], Sadhana Singh[4], Sunil Yadav[5]
[1,2,3,4]Dept. Of Computer Science, ITM Gorakhpur, UP, India.
[5]Asst.Professor, Dept. of Computer Science, ITM Gorakhpur, UP, India.

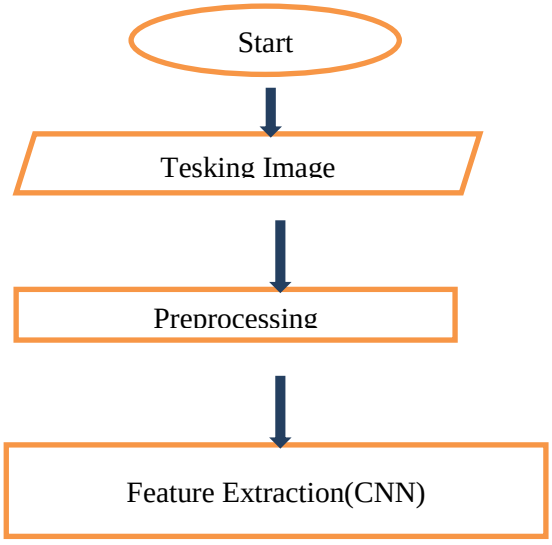
How to cite this paper:
Sadhana Maurya[1] ,Pratibha Sharma[2] , Sanjal Gupta[3], Sadhana Singh[4], Sunil Yadav[5], "Yoga Pose Detection System", IJIRE-V4I03-467-472.

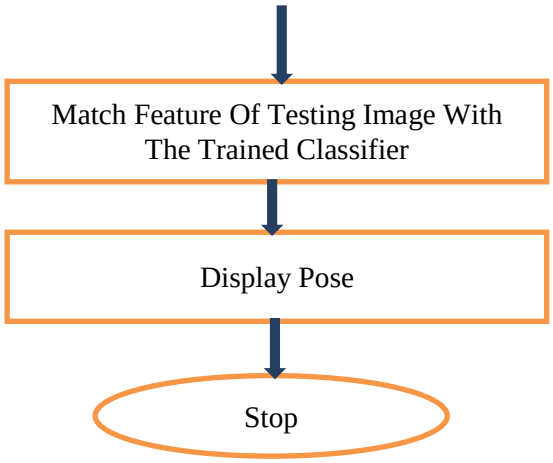
Abstract: *Yoga is a practice that unifies the mind, body, and spirit with the goal of developing a well-rounded personality. A person's body and brain might suffer major harm from incorrect yoga poses and techniques. It was only carried out under the direction of a skilled Guru in the past (teacher). Finding a skilled Guru is challenging in this hectic world with time and location restrictions. It is crucial to start practicing yoga the appropriate way from the beginning. Deep Learning models can be taught to recognize yoga positions and to offer feedback or adjustments as necessary. The first real-time multi-person system that completely changed pose estimation was called Open Pose. In this study, we developed a model that assesses a given human pose using Open Pose. . In order to test whether or not adding more features to the dataset actually improves the model's accuracy, we compared the model's output for 2D and 3D points of the image. Also, we have proposed a straightforward neural network model that accurately analyses the input image and communicates if the stance taken in the image is correct.*

Key Word: *Machine Learning, Deep learning, Pose estimation, Open Pose, Convolutional Neural network.*

I.INTRODUCTION

There is no denying the value of sports and exercise in human lives. Exercise has numerous advantages that are both physical and mental. As a result, many researchers have become quite interested in this field. People demand models and systems that give them the finest prospects for job progression as the fields of computer vision and deep learning have experienced significant growth. In order to grow their business, major sports leagues including the NHL, MLB, NBA, NFL, and NASCAR are now using deep learning in the United States. Wearable technology, automated journalism, and incorporating computer vision are a few areas in the sports sector where deep learning has made a significant contribution. One workout that calls for difficult positions is yoga. Yoga, which originated in ancient India, was previously thought of as a senior citizen practice. But due to yoga's numerous spiritual, bodily, and mental advantages, it gained popularity among people of all ages. The proper posture in which Yoga and other exercises must be conducted is one of the main issues. Even a small posture mistake can undermine the effects of exercise and increase the risk of accidents or structural deformities. This highlights the need for an instructor who can provide instructions on how to perform the exercise properly in order to obtain the most benefit possible from it. Yet, not everyone has access to a trainer who can watch over them while they exercise and correct their posture . This necessitates the need for a model who can recognize the poses and produce an output that will satisfy the needs of a trainer. So, we developed a model that not only identifies the workout but also notifies the user whether they are performing it correctly or not. The goal of this study is to identify various yoga positions and utilize an accurate model to address the issue of poor posture.





II.LITERATURE SURVEY

2.1 Satyam Goyal Human body pose identification is still a difficult task, despite extensive research and attention in the fields of computer vision and artificial intelligence. Human pose detection has several uses, including public security and health monitoring. This essay focuses on yoga, a discipline that has been practiced for more than 2,000 years. Yoga has become a popular form of exercise in contemporary society, which has led to a desire for guidelines on how to practice yoga correctly. The presence of a trainer is crucial since performing some yoga postures incorrectly might result in tiredness and injuries. Since many people lack the means to hire a yoga instructor or guide, artificial intelligence can step in and offer advice on postures. The classification of yogic poses is currently the main topic of study on pose estimation for yoga. In this work, we offer a method that enables real-time posture estimation to find a person's pose fault and help them rectify it using the Tensorflow Move Net Thunder model.

2.2 Shruti Kothari Over the past few years, many studies have been conducted on human pose estimation. Human pose estimate differs from other computer vision challenges in that it must locate and assemble human body parts based on the human body's already known structure. Pose estimation in sports and fitness can help people work out more effectively and reduce the risk of injury. Yoga self-instruction systems have the ability to both ensure that it is practiced properly and make yoga popular, as [3] shows. . The extensive research being done in this area makes deep learning techniques appealing. The classification of all 6 yoga poses using a hybrid CNN and LSTM model using OpenPose data is proven to be quite successful. Basic CNNs and SVMs also outperform our predictions in impressive ways. Performance of SVM shows that pose estimation or activity identification issues can also be solved with ML techniques. Additionally, compared to a neural network, SVM is significantly simpler, lighter, and needs less training time.

2.3 Shailesh S2 and Josvin Jose Yoga is a beneficial exercise that has its roots in India and can revitalize a man's physical, mental, and spiritual wellbeing. As technology develops quickly, there are numerous opportunities for computational probing across all social domains. However, it is currently difficult to use artificial intelligence and machine learning techniques in a multidisciplinary field like yoga. This work uses deep learning methods like convolutional neural networks (CNN) and transfer learning to create a system that can identify a yoga position from an image or frame of a video.

III.DETECTION OF HUMAN BODY USING OPEN POSE AND CNN

Open Pose is deep learning based computer vision algorithm that can be detect and track human body key point in real time. It uses convolutional neural network (CNN) to estimate the location of body joints such as head, shoulders, elbows, wrists, hipsknees, and ankles. The algorithm works by first detecting the person in the image or video frame using a body part segmentation network. Then, it estimates the 2D location of the body key points using a pose estimation network. The CNN is trained on large dataset of annotated image and learns to recognize patters in the data that correspond to different body parts. The network consists of multiple layer of convolutional and pooling operators that extract features from the input image. These features are then passed through fully connected layer that output the 2D location of the body key points. The CNN is trained using loss function that measures the difference between the predicted key points and the ground truth key points. During training, the weights of network are adjusted to minimize the loss function, resulting in the model that accurately estimates the location of body key points in new image or video frames.

The network consists of several stages, each of which performs a specific task in the pose estimation process.

The first stage of the network is a feature extractor that takes an input image and extracts a set of feature maps that capture the important visual information in the image. The feature maps are then passed through a series of convolutional layers that perform spatial filtering and feature extraction.

The second stage of the network is a set of part affinity fields (PAFs) that encode the spatial relationships between body parts. The PAFs are generated by a set of convolutional layers that predict the likelihood of each pixel belonging to a particular body part and the direction of the connection between body parts.

The third stage of the network is a set of key point confidence maps that predict the likelihood of each pixel belonging to a particular body key point. The confidence maps are generated by a set of convolutional layers that predict the probability of each pixel belonging to a particular key point.

Finally, the PAFs and key point confidence maps are combined to generate the final pose estimation output. This is

done by first grouping the key points that belong to the same person based on the PAFs, and then refining the keypoint locations based on the confidence maps.

Overall, the Open Pose architecture is a complex and powerful deep learning system that is capable of accurately detecting and tracking human body key points in real-time. It has many applications in fields such as sports analysis, healthcare, and robotics.

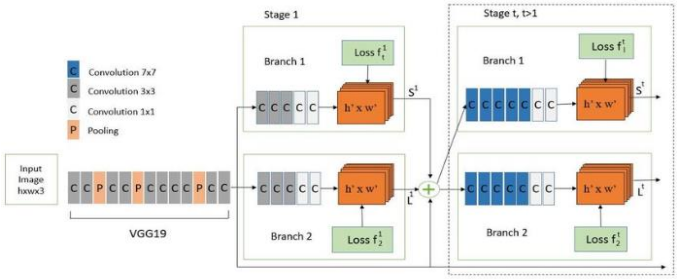


Fig 2.Architecture of Open Pose

IV. CONFIDENCE MAPS

The confidence map is a key component of the Open Pose model, which is a convolutional neural network (CNN) designed for human pose estimation. The confidence map is essentially a heat map that represents the likelihood of a body part being present at a particular location in an image.

In the Open Pose model, the confidence map is generated by passing the input image through a series of convolutional layers, which extract features from the image. These features are then used to predict the location and confidence score of each body part.

The confidence score is represented as a value between 0 and 1, with higher values indicating a higher likelihood of the body part being present at that location. The confidence map is then used to estimate the pose of the person in the image by identifying the locations of the body parts with the highest confidence scores.

Overall, the confidence map is a critical component of the Open Pose model, as it allows the model to accurately estimate the pose of a person in an image by predicting the likelihood of each body part being present at a particular location.

Part Affinity Fields

The part affinity field is another key component of the Open Pose model, which is a convolutional neural network designed for human pose estimation. The part affinity field is essentially a set of vectors that represent the spatial relationships between different body parts in an image.

In the Open Pose model, the part affinity field is generated by passing the input image through a series of convolutional layers, which extract features from the image. These features are then used to predict the direction and magnitude of the vectors that connect different body parts.

The part affinity field is used in conjunction with the confidence map to estimate the pose of the person in the image. Specifically, the part affinity field is used to identify the connections between different body parts, such as the arms and shoulders or the legs and hips. By identifying these connections, the model can accurately estimate the pose of the person in the image.

Overall, the part affinity field is a critical component of the Open Pose model, as it allows the model to accurately estimate the pose of a person in an image by identifying the spatial relationships between different body parts.

Loss Function

The loss function is a key component of the Open Pose model, which is a convolutional neural network designed for human pose estimation. The loss function is used to measure the difference between the predicted pose and the ground truth pose for a given image.

In the Open Pose model, the loss function is a combination of two different loss functions: the confidence map loss and the part affinity field loss. The confidence map loss measures the difference between the predicted confidence map and the ground truth confidence map, which represents the likelihood of each body part being present at a particular location in the image. The part affinity field loss measures the difference between the predicted part affinity field and the ground truth part affinity field, which represents the spatial relationships between different body parts in the image.

The loss function is used during the training phase of the model to adjust the weights of the convolutional layers in order to minimize the difference between the predicted pose and the ground truth pose. The goal is to find the set of weights that results in the most accurate pose estimation for a given image.

Overall, the loss function is a critical component of the Open Pose model, as it allows the model to learn from the ground truth poses and improve its accuracy over time.

V.DATASET OF YOGA POSE

A dataset of yoga poses typically includes a collection of images or videos of individuals performing various yoga postures.

These images or videos are often labeled with the name of the pose. On a publicly accessible, open-source

collection dataset that is available online, the proposed methodology is tested. Three yoga poses—Tree pose, Warrior 2 pose, and T pose—are included in this dataset.

The Open Pose CNN model is trained on large dataset of image and videos of people in various poses and positions.

Fig 3.Dataset of different yoga poses

Tree pose

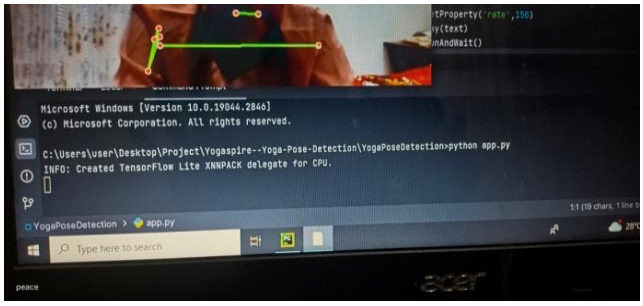
T pose

warrior2 pose

Fig 4.

VI.RUNNING STATUS OF THE PROGRAM

To run the program



- Open Command Prompt (PyCharm Community Edition 2023)
- Enter python app.py

VII.RESULT

The Open Pose model architecture consists of two main components: the feature extraction network and the key point detection network.

The feature extraction network is a series of convolutional layers that extract features from the input image. The output of this network is a set of feature maps that represent different levels of abstraction in the image.

The key point detection network takes the feature maps as input and predicts the location of body key points. This network consists of several stages, each of which refines the key point predictions based on the previous stage. The final output of the network is a set of confidence maps that represent the likelihood of a key point being present at each location in the image.

In addition to the key point detection network, Open Pose also uses a Part Affinity Field (PAF) network to estimate the association between key points and body parts. The PAF network predicts a set of heat maps that represent the likelihood of a connection between two key points. These heat maps are used to estimate the orientation and direction of body parts, such as arms and legs.

Overall, the Open Pose model uses a combination of feature extraction, key point detection, and part affinity

estimation to accurately detect and track human body key points in real-time. It give the overall result with the accuracy of 72%

VIII.OUTPUT

Tree Pose

T Pose

IX.FUTURE WORK

There are more than 80 different yoga postures, yet the suggested system is limited to just three. The proposed dataset can be increased by including the critical locations for the necessary yoga poses. It's difficult enough to include many postures and to have model work on many poses (classifying various poses). When body components are overlapping or missing, actions should be done to obtain important points in order to improve outcomes. Keras pose estimate affects how well the model performs. Other uses for this method of extracting angles as features include activity detection and sports activity monitoring.

X.CONCLUSION

In conclusion, the use of Open Pose CNN-based model for yoga pose detection has shown promising results. The model is able to accurately detect and track human body key points in real-time, which is essential for analyzing yoga poses. By using a combination of feature extraction, key point detection, and part affinity estimation, the model is able to estimate the orientation and direction of body parts, such as arms and legs, which is crucial for identifying different yoga poses. This technology has the potential to revolutionize the way yoga is taught and practiced, by providing real-time feedback and guidance to practitioners. Overall, the Open Pose CNN-based model is a powerful tool for analyzing human movement and behavior in a variety of applications, including yoga.

References

1. *A. Agarwal and B. Triggs. (2004). "3D human pose from silhouettes by relevance vector regression". Intl Conf. on Computer Vision & Pattern Recogn.pp.882–888.*
2. *A. Gupta, T. Chen, F. Chen, and D. Kimber. (2011). "Systems and methods for human body pose estimation". U.S. patent, 7,925,081 B2.*
3. *A. Kendall, M. Grimes, R. Cipolla. (2015). "PoseNet: a convolutional network for real-time 6DOF camera relocalization". IEEE Intl. Conf. Computer Vision.*
4. *Chen HT, He YZ, Hsu CC et al. (2014). Yoga posture recognition for self-training. pp 496–505.*
5. *Connaghan D, Kelly P, O'Connor NE et al . (2011). Multi-sensor classification of tennis strokes. . Retrieved from Proc IEEE Sens.: https://doi.org/ 10.1109/icsens.2011.6127084*
6. *DatasetOnline. (n.d.). Retrieved from YogaVidCollected: https://archive.org/details*
7. *E. Trejo, P. Yuan. (2018). "Recognition of yoga poses through an interactive system with kinect device". Intl. Conf. Robotics and Automation Science.*
8. *G. Ning, P. Liu, X. Fan and C. Zhan. (2018). "A top-down approach to articulated human pose estimation and tracking". ECCV Workshops.*
9. *Gao Z, Zhang H, Liu AA et al. (2016). Human action recognition on depth dataset. . Neural Comput Appl 27:2047–2054..*
10. *H. Sidenbladh, M. Black, and D. Fleet. (2000). "Stochastic tracking of 3D human figures using 2D image motion". Proc 6th European Conf. Computer Vision.*
11. *L.Sigal. (2011). "Human Pose Estimation". Springer: Ency. of Comput.Vision.*
12. *M. Islam, H. Mahmud, F. Ashraf, I. Hossain and M. Hasan,. (2017). "Yoga posture recognition by detecting human joint points in real time using microsoft kinect". IEEE Region 10 Humanit. Tech. Conf., pp. 668-67.*

13. *M. Dantone, J. G. (2013). "Human pose estimation using body parts dependent joint regressors". Proc. IEEE Conf. Computer Vision Pattern Recogn.*
14. *P. Dar. (2018). "AI guardman – a machine learning application that uses pose estimation to detect shoplifters".*
15. *P. Szczuko. (2019). "Deep neural networks for human pose estimation from a very low resolution depth image" . Multimedia Tools and Applications, vol. 78, no. 20, pp. 29357–29377.*
16. *S. Kreiss, L. Bertoni, and A. Alahi. (2019). "PifPaf: composite fields for human pose estimation". IEEE Conf. Computer Vision and Pattern Recogn.*
17. *S. Haque, A. Rabby, M. Laboni, N. Neehal, and S. Hossain,. (2019). "ExNET: deep neural network for exercise pose detection". Recent Trends in Image Process. and Pattern Recog..*
18. *S.Yadav, A.Singh,A.Gupta and J.Raheja. (May,2019). "Real Time Yoga Recognition Using Deep Learning". Neural Comput. and Appl.*
19. *U. Rafi , B. Leibe , J. Gall , and I. Kostrikov. (2016). "An efficient convolutional network for human pose estimation". British Mach. Vision Conf.*
20. *Z. Cao, T. Simon, S. Wei, and Y. Sheikh. (2017). "OpenPose: realtime multi-person 2D pose estimation using part affinity fields" . Proc. 30th IEEE Conf. Computer Vision and Pattern Recogn.*