IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Impact Factor: 7.53

Volume 4, Issue 5, April 2024

Deep Learning Approach for Yoga Posture Evaluation

Sri Ch. Siva Subramanyam[1], N S S P Kameswar[2], K Yuva Teja[3], K Madhu Sai Varma[4], M Govind[5]

Assistant Professor, Department of Information Technology[1]

Students, Department of Information Technology[2,3,4,5]

S.R.K.R Engineering College, Bhimavaram, Andhra Pradesh, India

nedunurisaikamesh132@gmail.com and sivasubbu22@gmail.com

Abstract*: In the present generation due to the mechanical lifestyle of people, everyone is facing several health issues. To balance the lifestyle, one needs to inculcate habits that improve health. Adopting yoga as a habit can be incredibly beneficial. Holding certain yoga positions that require supporting one's body weight can be challenging and strengthen specific muscles. Yoga consists of various components that can enhance flexibility, strength, balance, and stability.*

Our application is evaluated on different Yoga poses under different scenes. It has been observed that pose detection techniques can be used to identify the postures and to assist people to perform yoga more accurately, for the accurate detection of yoga pose different feature extraction and pre-processing methods are applied to the dataset just by using machine learning algorithms. In this work, a website can be created to link yoga poses with their corresponding asanas and upon correct execution that pose time will be calculated and report will be generated to view pose time he/she performed yoga as an correctly.

Keywords: CNN, Machine learning, Deep learning, Movenet thunder , Artificial intelligence

I. INTRODUCTION

1. Background:

The concept of yoga is deeply rooted in aligning the mind and body to achieve spiritual unity. Originating from Sanskrit, the word "yoga" translates to "union," symbolizing the integration of various aspects of one's being. Initially, yoga was imparted individually and restricted to men of the highest caste, reflecting its sacred and exclusive nature. However, over time, yoga has transcended cultural and societal barriers to become a widely recognized practice renowned for its holistic benefits. complexity and volume of malicious attacks has magnified certain limitations of traditional machine learning methods, such as their focus on processing low-dimensional data and the manual selection of features.

2. Medical Significance and Contemporary Relevance:

In contemporary society, yoga has garnered increasing attention in medical research due to its myriad health benefits. Numerous studies have explored its efficacy in diverse medical applications, ranging from cardiac rehabilitation to mental health interventions. One of yoga's distinguishing features is its accessibility, catering to individuals of all ages and physical capabilities. Its gentle yet effective nature makes it an appealing option for those seeking physical and psychological well-being amidst the demands of modern life.

3.Yoga Pose Evaluation Scope:

The accuracy of yoga practice hinges on correct posture and technique, as improper execution can lead to injury or diminished benefits. To address this challenge, researchers have turned to machine learning techniques, particularly the Random Forest classifier, to assess yoga poses. These approaches pave the way for innovative self-learning systems and practical yoga assistant applications, enabling individuals to receive accurate guidance and feedback in their practice.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 5, April 2024

4. Pose time calculation and report generation:

In the realm of yoga practice, ensuring correct posture and technique is paramount for reaping its full benefits and preventing injury. Leveraging advancements in technology and machine learning, a novel approach has been proposed to enhance the accuracy and efficacy of yoga practice. By utilizing pose detection algorithms, particularly the MoveNet, this system aims to assess the correctness of yoga poses in real-time. The system calculates the duration for which each pose is held accurately, providing users with valuable feedback on their performance. Additionally, the system generates comprehensive reports detailing pose execution times and overall practice metrics. Through this innovative integration of technology and traditional wellness practices, individuals can optimize their yoga practice, fostering holistic health and well-being.

Movenet Algorithm:

MoveNet Thunder is based on a convolutional neural network (CNN) architecture and uses a set of deep learning algorithms to estimate the 2D and 3D coordinates of human joints from video data. The model is trained on a large dataset of labeled human poses, which allows it to accurately recognize a wide range of movements and poses.

The Movenet Thunder utilizes a neural network architecture that is specifically designed for analyzing human poses and movements in real-time. Its key attribute is its ability to operate with precision and speed simultaneously. The model employs depth-wise separable convolutions, which involve breaking down the computation into two stages: depth-wise convolution and pointwise convolution. During depth-wise convolution, a single convolutional filter is applied to each input channel, and pointwise convolution is used to create a linear combination of the output from the depth-wise convolution. This approach allows the model to effectively capture intricate spatial and temporal characteristics of human motion while maintaining high accuracy and efficiency. The model consists of a series of convolutional layers that extract features from the input image or video frames. These features are then fed into a series of depth-wise separable convolutional blocks, which perform spatial filtering to detect and localize key points in the human body. Each depth-wise separable convolutional block is composed of a depth-wise convolutional layer, which applies a separate convolutional filter to each input channel, followed by a point wise convolutional layer, which applies a 1x1 convolutional filter to combine the output channels

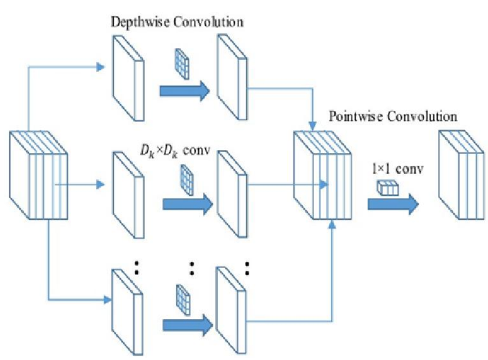


Fig 1 .Depthwise Separable Convolution

MoveNet Architecture:

The MoveNet architecture is a pose estimation model that accurately locates human keypoints using heatmaps. It consists of two main components: a feature extractor and a set of prediction heads.

The feature extractor has a feature pyramid network (FPN) attached to it. This enables highresolution (output stride 4) and semantically rich feature maps. There are four prediction heads connected to the feature extractor, which predict the following:

- Person center heatmap: predicts the geometric center of person instances
- Keypoint regression field: predicts the full set of keypoints for a person, which is used for grouping keypoints into instances.
- Person keypoint heatmap: predicts the location of all keypoints, regardless of person instances.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 5, April 2024

- 2D per-keypoint offset field: predicts local offsets from each output feature map pixel to the precise sub-pixel location of each keypoint.

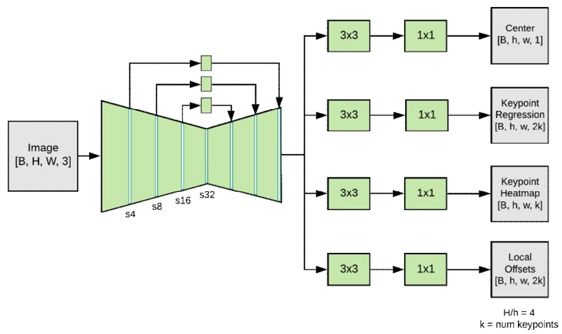


Fig 2 MoveNet Architecture

Pose Extraction Using MoveNet:

MoveNet is a highly efficient and precise model that can detect the 17 key points of the human body, making it ideal for various applications. It is hosted on TensorFlow and has two variants known as "Lightning" and "Thunder". The Lightning variant is designed for use in critical latency applications, while the Thunder variant is suitable for applications that require high accuracy. Both variants are capable of running at more than 30 frames per second (FPS) on most modern devices, including mobile phones and computers. It is a pose estimation model released by Google on May 17, 2021. Compared to conventional pose estimation models, it improves the detection accuracy in videos with intense motion. It is ideal for live fitness and sports applications. The architecture of the AI model is similar to CenterNet, utilizing MobileNetV2 as the feature extractor and adding Feature Pyramid Network (FPN) to enhance the model's performance. By setting the output stride to 4, the model can efficiently process high-resolution feature maps. The output of the AI model includes a person center heatmap, a keypoint regression field, a person keypoint heatmap, and a 2D per-keypoint offset field.

Multi-Layer Perceptron (MLP):

Multi-Layer Perceptron also known as MLP in short. It is a type of artificial neural network used in machine learning. An MLP consists of multiple layers of nodes, where each node is a neuron that performs a computation on its input and produces an output.

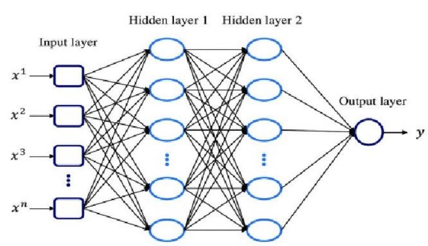


Fig 3 Multi-layer perceptron Architecture

II. LITERATURE SURVEY (RELATED WORK)

1. Implementation of Computer Vision in Detecting Human Poses

Author: Samuel Mahatmaputra Tedjojuwono

Published: 2020

Description:

Authors explained that their review explores the implementation of computer vision in detecting human poses, discussing traditional and modern techniques such as CNNs and OpenPose. It highlights advancements in accuracy and applications across sports analysis, healthcare, and robotics. Challenges like occlusions and real-time processing are

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 5, April 2024

noted, with proposed solutions. Future directions include enhancing algorithms, integrating multimodal data, and exploring 3D pose estimation. Overall, while significant progress has been made, further research is needed to overcome challenges and fully realize the potential of computer vision in human pose detection

2. A Proposal of Yoga Pose Assessment Method Using Pose Detection for Self Learning
Author: Maybel Chan Thar, Khine Zar Ne Winn,
Published: 2019
Description:
The authors determined that OpenPose utilizes Part Confidence Maps and Part Affinity Fields to identify joint locations, facilitating multi-person keypoint detection for human body, hand, and facial keypoints. The process involves bipartite matching and parsing to achieve accurate results. However, comparisons of accuracy with existing literature are hindered by the absence of access to a standardized dataset. Despite this limitation, OpenPose demonstrates robustness and versatility in detecting complex human poses across various applications. Its methodology combines computer vision and deep learning techniques to enable real-time, high-fidelity pose estimation. Further research may focus on addressing dataset accessibility issues and refining algorithms to enhance performance in diverse contexts

3. Infinity Yoga Tutor: Yoga Posture Detection and Correction System
Author: Fazil Rishan, Binali De Silva, Sasmini Alawathugoda, Shakeel Nijabdeen, Lakmal Rupasinghe, Chethana Liyanapathirana
Published: 2020
Description:
The paper likely introduces the system's architecture and methodology, which may involve the use of deep learning algorithms for pose detection and analysis. The system is likely equipped with a camera or other imaging devices to capture practitioners' movements during yoga sessions.
The research may discuss the technical details of the pose detection and correction algorithms employed in the Infinity Yoga Tutor system. These algorithms are likely trained on annotated datasets of yoga poses to learn the features and patterns associated with correct and incorrect postures

4. Deep neural network for exercise pose detection
Author: S. Haque, A. Rabby, M. Laboni, N. Neehal
Published: 2019
Description:
Presenting at the 2019 conference on Recent Trends in Image Processing and Pattern Recognition, S. Haque, A. Rabby, M.
Laboni, N. Neehal, and S. Hossain detailed their approach of employing an automatic Learning Rate reduction method to accelerate convergence and achieve closer alignment with global optimization goals in their model. Their research primarily centers on Human pose detection, Object detection, Deep learning, and Exercise Pose Detection domains. Despite their efforts, they observed that their proposed model encountered challenges in accurately interpreting certain poses, attributed to overfitting phenomena. This issue suggests the need for further refinement in model training methodologies to mitigate overfitting effects and improve overall performance in pose detection tasks.

5. Live Yoga Pose Classification Using Image Processing And LR Algorithm
Author: Yang Xu
Published: 2020
Description:
This paper likely introduces a system that utilizes image processing methods to analyze live video feeds of individuals performing yoga poses. This system may involve techniques such as image segmentation, feature extraction, and pattern recognition to identify and classify different yoga poses in real-time , the research may discuss the use of the Logistic Regression algorithm as a machine learning model for pose classification. Logistic Regression is a statistical method commonly used for binary classification tasks, but it can also be adapted for multi-class classification, making it suitable for categorizing various yoga poses and describe the implementation of the proposed system preprocessing steps applied to the input video data, the training LR model using labeled datasets of yoga poses, and the classification of live yoga poses in real-time

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 5, April 2024

III. METHODOLOGY

Machine learning is a modern innovation and is used widely all around the world. This is mainly helpful in healthcare sectors such as in health and fitness applications. We use MoveNet and MLP models to assess the yoga pose is correctly performed or not.

OpenCV is a computer vision library that is available as open-source software, offering an extensive selection of tools and algorithms for various applications such as image and video processing, object detection and recognition, feature extraction, among others. On the other hand, TensorFlow is a renowned open-source software library designed by Google for building and training machine learning models.

It offers a comprehensive range of functionalities and tools for developing, training, and deploying machine learning models for various applications. The main difference between TensorFlow and TensorFlow.js is the programming language used to build and train the models. TensorFlow is primarily designed to be used in Python, while TensorFlow.js is designed to be used in JavaScript. Additionally, tensorflow.js can run the models in browser.

Here's a breakdown of the methodology into steps:

1. Data Collection:

Data plays a major role for efficient working of any model. There is an open dataset which consists of images of people with different personalities performing different poses. Out of 84 traditional yoga asanas, a subset of 7 asanas is utilized in this work. It consists of 7 yoga asanas having around 3200 images. The whole dataset is divided into two sections: training and testing dataset. Training dataset is utilized for training the model and for testing the testing dataset is utilized.

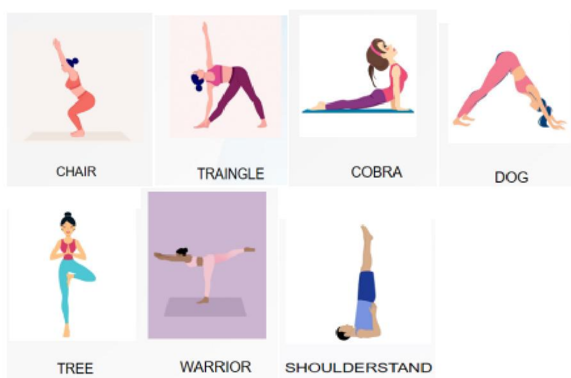


Fig 4 Yoga Asans

2. Data Preprocessing:

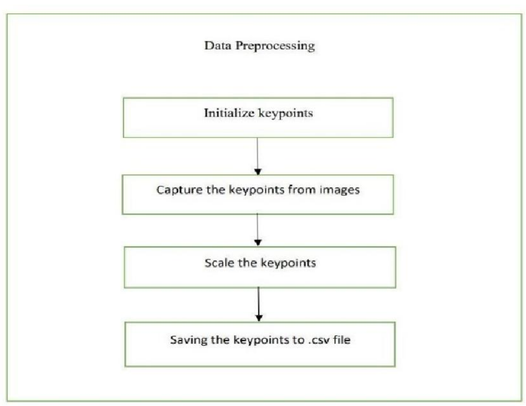


Fig 5 Preprocessing

The collection may consist of various noise and blurry images. In order to remove these inappropriate data, we performed data preprocessing. If the input image is not in the RGB colour space or in the JPEG file format, it will be

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 5, April 2024

discarded. After the collection of right set of images, feature extraction process takes place. In order to classify, the system needs to detect the posture of the Yogi.

Here, the keypoint detection algorithm- MoveNet will be used to extract the keypoints from the image. Each keypoint contains x-coordinate, y-coordinate, and its confidence score. The image having keypoints with threshold confidence score will be taken for assessment. The MoveNet will crop the input image based on the central point of the body. The keypoints of various yoga poses were extracted from a dataset using MoveNet, and the keypoints were stored in a CSV file with corresponding labels for each. This csv file data is used to train our MLP model to assess the yoga asanas in real-time.

Here is the proposed methodology of the system that calculate the pose time upon correct execution of the pose and real time report generation is going totake place by the following methodology

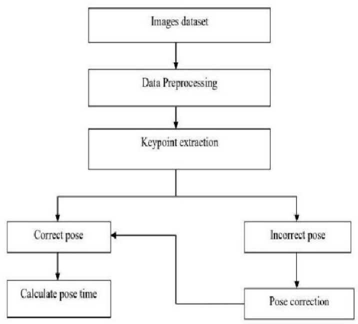


Fig 6 Proposed Methodology.

IV. EXPERIMENTATION & EXECUTION

Created a website for storing data of evaluated yoga postures is a crucial step in developing a comprehensive yoga posture evaluation system. This login page serves as the entry point for users, allowing them to access and input their yoga posture data securely. By implementing user authentication mechanisms, such as username and password verification, the system ensures that only authorized users can store and access their evaluation data. The stored data can include information such as the date and time of evaluation, the specific yoga poses performed, and any additional notes or feedback provided by the user.

After signup and logging into the website then user has to perform the asana and upon correct execution then the coordinates turn to green and the timer get started as in the figure below

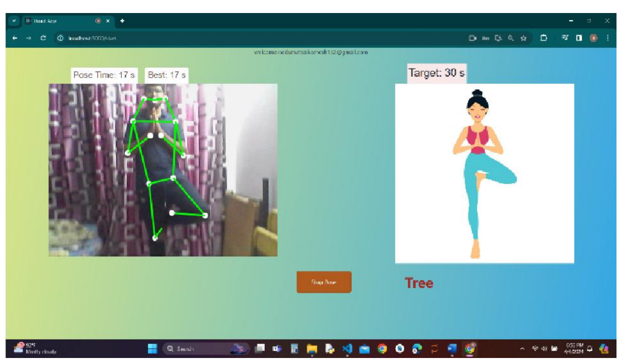


Fig 7 Correct pose execution and timer countin

Whenever the user stopped perform the pose the timer turn back to zero showing that the white color key-coordinates resembles that the pose is not being performed.

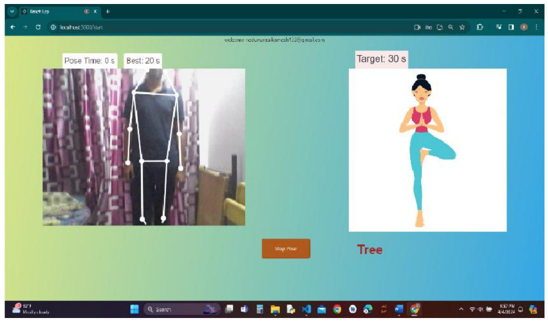
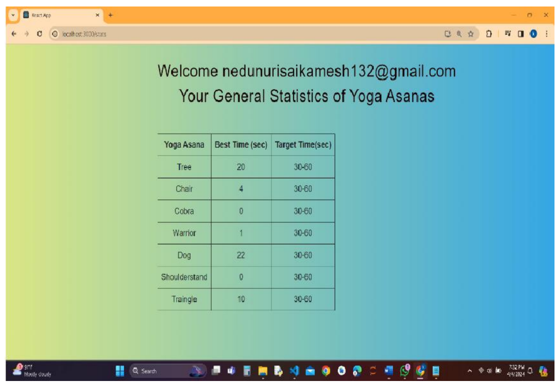


Fig 8 Wrong pose and stopped performing .

After performing all the asans or whenever the practioner want to see the report he have to press the generate report button and the report will be generated and it a below



Yoga Asana	Best Time (sec)	Target Time(sec)
Tree	20	30-60
Chair	4	30-60
Cobra	0	30-60
Warrior	1	30-60
Dog	22	30-60
Shoulderstand	0	30-60
Traingle	10	30-60

Fig 9 Generated Report.

The output is correctly calculating the duration of yoga poses and generating reports is crucial for practitioners to track progress and ensure they are achieving their fitness goals effectively. To achieve this, yoga instructors or practitioners typically use a combination of timing techniques and mindfulness. Each pose may have a recommended duration based on its complexity and the individual's level of proficiency. Utilizing timers or counting breaths can help ensure poses are held for the appropriate amount of time, promoting balance and strength building while avoiding injury. Once a yoga session is completed, practitioners or instructors can compile data regarding the duration of each pose, noting any variations or modifications made. This information can then be used to generate detailed reports that provide insights into progress, areas for improvement, and overall performance trends over time. Such reports can be invaluable tools for refining practice routines, setting goal.

V. RESULT ANALYSIS

Classification Report

	precision	recall	f1-score	support
0	1.00	1.00	1.00	39
1	0.97	1.00	0.99	35
2	0.93	1.00	0.96	38
3	1.00	1.00	1.00	3
4	1.00	1.00	1.00	7
5	1.00	0.88	0.93	8
6	1.00	1.00	1.00	71
7	1.00	0.88	0.94	26
accuracy			0.98	227
macro avg	0.99	0.97	0.98	227
weighted avg	0.98	0.98	0.98	227

Fig 10 Classification Report.

ACCURACY: This metric determines the ratio of accurately classified instances to the total number of instances in the dataset. It reflects the performance of a classification model by indicating how often the model made correct predictions for all instances in the dataset.

Accuracy = (TP + TN) / (TP + TN + FP + FN)

PRECISION: It measures the proportion of true positive predictions among all positive predictions made by the model.

Precision = TP / (TP + FP)

RECALL: It measures the proportion of true positive predictions among all actual positive instances in the dataset.

Recall = TP / (TP + FN)

F1 SCORE: It is the harmonic mean of precision and recall and provides a balance between the two metrics.

F1 Score = 2 * (Precision * Recall) / (Precision + Recall)

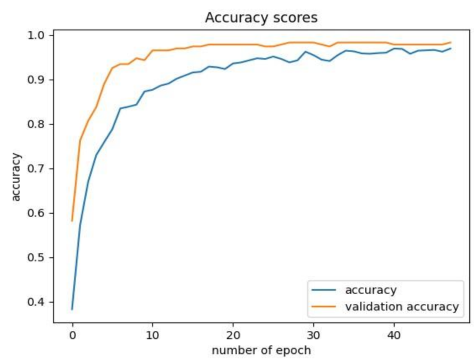


Fig 11 Validation plot

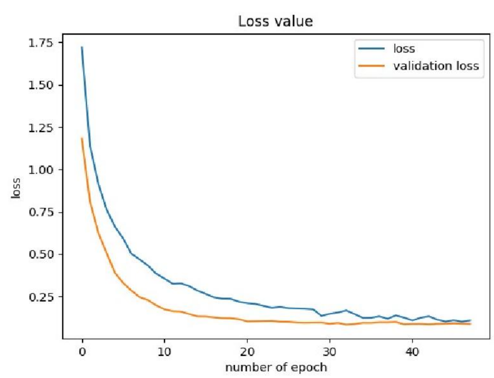


Fig 12 Loss plot

Confusion matrix:

A confusion matrix is a tabular representation that helps in assessing the effectiveness of a classification model by comparing its predictions with the actual values. It allows for a detailed analysis of how well the model performs in differentiating between different classes and can provide insights into the model's strengths and weaknesses.

True Positive (TP): The number of actual positive cases that are correctly classified as positive by the model.

False Positive (FP): The number of actual negative cases that are incorrectly classified as positive by the model.

False Negative (FN): The number of actual positive cases that are incorrectly classified as negative by the model.

True Negative (TN): The number of actual negative cases that are correctly classified as negative by the model.

The confusion matrix helps to calculate important performance metrics of a classification model such as accuracy, precision, recall, and F1-score. These metrics are useful in evaluating the model's effectiveness in identifying positive and negative cases accurately

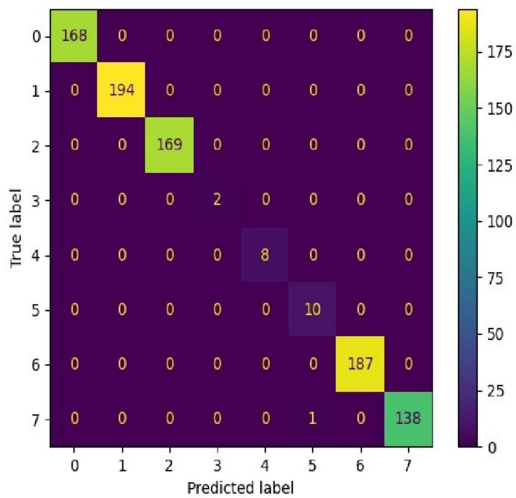


Fig 13 Confusion Matrix

VI. CONCLUSION

The project "Deep Learning approach for Yoga Posture Evaluation " is a valuable tool for individuals practicing yoga. Performing yoga every day is a healthy habit. This application takes in real time video of the person performing the yoga asanas. The system uses machine learning techniques to analyse user's posture and provide feedback on how to improve their form, which helps to improve safety and efficacy in yoga practice. It also indicates the time to perform each pose. The project combines computer vision and machine learning algorithms to accurately identify body landmarks and provide real-time feedback on the user's alignment.

Additionally, this project has the potential to enhance accessibility to yoga instruction and increase the likelihood of individuals adhering to a consistent yoga practice. It is a promising technology that can be further developed and integrated into existing yoga studios or home practice environments. Overall, the " Deep Learning Approach For Yoga Posture Evaluation " project represents an exciting advancement in the intersection of technology and wellness, with potential benefits for both practitioners and instructors.

VII. FUTURE WORK

1. Incorporating more yoga asans and expanding the dataset: The current project is based on recognizing a limited number of yoga poses . To make the model more comprehensive and useful, it could be expanded to recognize a larger vocabulary of asans. This would require collecting more data and training the model on a wider range of yoga poses.
2. Improving accuracy: Even though the current model achieves high accuracy, there is always room for improvement. Tuning the model's hyperparameters or using more advanced deep learning techniques like recurrent neural networks or transformers could help improve the accuracy even further.
3. Deployment on mobile devices: Once the model is developed, it can be deployed on mobile devices. This can be beneficial individuals or yoga practitioners who can use the app to perform yoga more effectively in daily basis.

REFERENCES

[1] Z. Cao, T. Simon, S. Wei and Y. Sheikh - "Realtime Multi-person 2D Pose Estimation using Part Affinity Fields" IEEE, 2017.
[2] O. Patsadu, C. Nukoolkit, and B. Watanapa - “Human gesture recognition using kinect camera” IEEE, 2012.
[3] S. Patil, A. Pawar, and A. Peshave - “Yoga tutor: visualization and analysis using SURF algorithm” IEEE, 2011.

[4] H.-T. Chen, Y.-Z. He and C.-C. Hsu - "Computer Assisted Yoga Training System" IEEE, 2018

[5] S. Jin, X. Ma, Z. Han, Y. Wu, W. Yang, W. Liu, C. Qian and W. Ouyang – "Towards MultiPerson Pose Tracking: Bottom-up and Top-down Methods" ICCV2017, 2017.

[6] Yang Lin, Zhang Longyu and Dong Haiwei - "Evaluating and improving the depth accuracy of Kinect for Windows v2" IEEE, 2015.

[7] M. C. Thar, K. Z. N. Winn, and N. Funabiki - "A proposal of yoga pose assessment method using pose detection for self-learning" ICAIT, 2019.

[8] Kendall, M. Grimes and R. Cipolla - "PoseNet: a convolutional network for real-time 6- DOF camera relocalization" IEEE, 2015.

[9] Fazil Rishan, Binali De Silva, Sasmini Alawathugoda, Shakeel Nijabdeen, Lakmal Rupasinghe, Chethana Liyanapathirana - "Infinity Yoga Tutor: Yoga Posture Detection and Correction System" IEEE, 2020.

[10] Muhammad Usama Islam, Hasan Mahmud, Faisal Bin Ashraf, Iqbal Hossain, Md. Kamrul Hasan- "Yoga Posture Recognition by Detecting Human Joint Points in Real Time Using Microsoft Kinect" RESEARCHGATE.IN, 2019.