IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

CONVOLUTIONAL YOGA POSE ESTIMATOR

[1]Vijayalaxmi Yalavagi, [2]Impana S, [3]Chethan M, [4]Chandan L, [5]Geetha Tadasad B

[1] Assistant Professor, [2] Student, [3] Student, [4] Student, [5] Student

Department of Information Science and Engineering,

Cambridge Institute of Technology, Bangalore, India

Abstract: Yoga pose estimation is a computer vision technique used to predict the position/ pose of a part of the human body**.** This paper presents the framework to analyse and assess yoga postures by designing, developing, and implementing a Yoga Posture Detection System using computer vision and deep learning (DL) models such as CNN and VGG16. The study employs advanced image processing algorithms to extract information from images or videos of individuals performing yoga poses. We employed several postures which includes camel, downdog, goddess, plank, tree, and warrior2. With the use of a deep learning model that has been built, the system is able to precisely recognise and categorise different positions while providing instantaneous feedback on proper alignment, balance, and posture.

I. INTRODUCTION

Yoga has become a regular part of many people's lives all across the world in recent years. This makes a scientific investigation of yoga postures necessary. Yoga posture estimate is a computer vision technology that anticipates a human body part's position or pose. Pose detection algorithms have been shown to be useful for both pose identification and improving the accuracy of yoga poses. In today's modern- day, ML and DL techniques have been proven to be important for object discovery tasks. we can effectively use the model to recognize different important body parts and estimate the user pose in real time. To accomplish this, we train a model with different yoga pose images. When an image is fed into a pose estimation model, it analyzes the image and recognized body part by performing feature extraction, indicating their positions on the screen. Additionally, the model provides a confidence value for each detection, indicating the likelihood that the given image as a input is correctly identified. we used different Yoga poses like camel pose, downdog pose, goddess pose, plank pose, tree pose, warrior2 pose to train the model, which makes it highly accurate at identifying various poses. The primary goal of this study is to use this detection technology to assist people in identifying which yoga posture they are performing.

Additionally, we also go over the shortcomings of current systems, like their poor accuracy, high processing costs, and restricted applicability to various body shapes and styles of yoga. In this paper, we proposed a Convolutional Neural Network (CNN) based approach for creating a yoga stance detection system. The suggested technique seeks to improve upon the shortcomings of current systems by offering more precise, effective, and broadly applicable solutions for the identification of yoga poses and the generation of feedback. Overall, the application of technology-based tools in our study may help to design more customised and successful yoga practices. Our findings can help build applications like virtual assistants and smart yoga mats that improve accessibility and personalisation of yoga practice.

II. LITERATURE SURVEY

The following literature survey provides an overview of some of the recent advances in Yoga Posture Recognition techniques:

[1] R. Dutta and S. Mukherjee," Smart Yoga Assistant A Mobile Application for Correcting Yoga Posture Using disguise Estimation," 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT), Kharagpur, India, 2020.

This paper introduces Smart Yoga Assistant, a smartphone operation that employs disguise estimation to deliver real-time feedback on the stoner's yoga posture. The suggested system uses a pre-trained deep literacy model to prognosticate the stoner's station and compare it to the proper posture. The authors tested their operation on a dataset of yoga pictures and set up that it was largely accurate in relating indecorous postures. The proposed operation's ease of use and efficacity demonstrate its pledge as a tool for yoga interpreters and preceptors.

[2] S. Sathya and S. Krishnan," Yoga Gesture Recognition using Convolutional Neural Networks," 2020 Fourth International Conference on Inventive Communication and Computational Technologies (ICICCT), Coimbatore, India, 2020.

The authors proposed a new system for recognising yoga movements with Convolutional Neural Networks. The proposed system recognised six yoga gestures with 94.6 delicacy. The authors meliorated the pre-trained VGG16 network by transfer literacy and tested the system on a 1,500- image dataset. The suggested system beat being algorithms for recognising yoga movements, showing its eventuality for real- world use.

[3] S. Saha and M.K. Kundu," Automated Recognition of Yoga Asanas Using Convolutional Neural Networks," 2020 International Conference on Computational Intelligence and Communication Systems (ICCICom), Kolkata, India, 2020. The authors suggested an automated approach for recognising yoga asanas with Convolutional Neural Networks (CNN). The authors suggested an automated approach for recognising yoga asanas with Convolutional Neural Networks (CNN). The suggested system included two stages picture preprocessing and point birth and bracket. The authors employed transfer literacy to fine- tune the pre-trained VGG16 network and estimated the system on a dataset of 20 yoga asanas, attaining an delicacy of 92.5 percent. The suggested system successfully recognised yoga asanas, with implicit operations in health and fitness shadowing.

[4] Narayanan, S. Sankara and Misra, Devendra Kumar and Arora, Kartik and Rai, Harsh, Yoga disguise Discovery Using Deep Learning ways (May 10, 2021). Proceedings of the International Conference on Innovative Computing & Communication ICICC) 2021.

This paper uses OpenPose to construct a model that assesses mortal acts. The experimenters examined the model's results for2D and 3D picture points to assess whether adding fresh characteristics to the dataset improves model delicacy. The authors propose a introductory neural network model for analysing input images and determining correct acts. For the original phase, TensorFlow was used to import pre-trained weights. OpenPose was also used to describe 17 body factors, including the left and right elbows. To make a shell image, a model was created utilising live camera to identify body sections and save the values of 17 points. The model was trained to honour and store correct Yoga acts. The camera- captured body corridor was compared to the original values of the instructed Yoga acts, yielding a result that included the posture name and accurate prosecution.

III. METHODOLOGY

3.1.Pre-processing:

This study proposes a deep learning-based approach to estimate yoga poses, similar to algorithm 1, for the purpose of identifying correct positions and offering feedback to enhance the posture. The proposed approach was done using CNN and VGG16, i.e., one of the pre-trained models in CNN, and comprises of three major steps.

1. **Classification:** The final fully-connected layers (Dense) perform the classification task based upon the extracted features. The convolutional layers provide the flattened matrix of features to the model. Dense layers learn to combine these features in a way that separates different classes. With softmax activation, the final output layer creates probabilities for every class, indicating the model's confidence in each prediction.

2. **Feature Extraction:** The convolutional layers (Conv2D) act as the feature extractors in the CNN. They learn filters that detect specific patterns and features within the images. These filters can identify edges, shapes, textures, & additional visual components that are important for differentiating between image types.
3. **Data Preprocessing:** ImageDataGenerator -This creates data augmentation for training data. It performs operations like:

- width_shift_range: Randomly shifts the image horizontally within a specified range (0.1 in our case).
- horizontal_flip: Randomly flips the image horizontally.
- Rescaling: Normalizes pixel values from a range of[0, 255] to [0, 1] using rescale=1./255.

3.2.Random rotation: rotation_range=10.

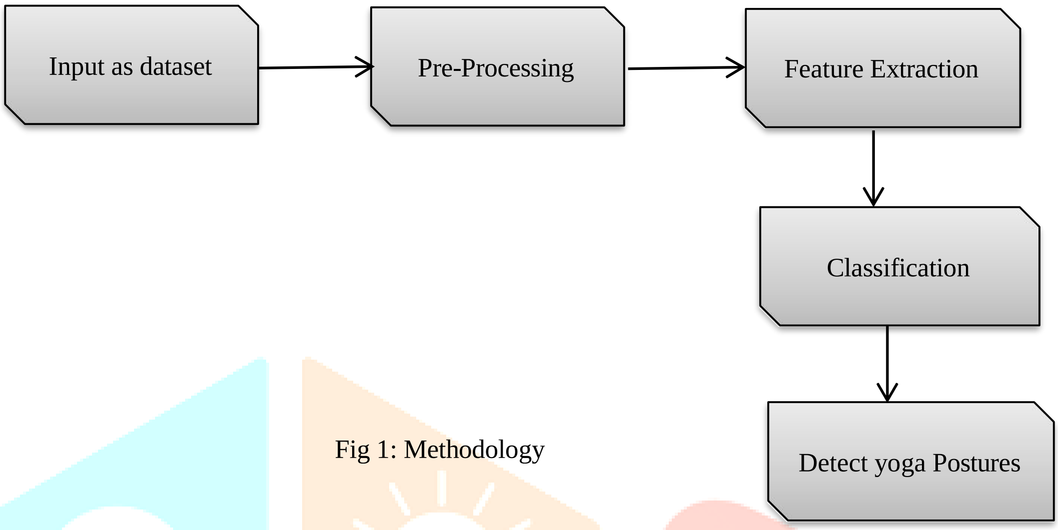

Fig 1: Methodology

3.3.Dataset Description:

The photos in the dataset were gathered from a variety of online resources, such as Google and Kaggle. Images were collected to form a complete dataset for recognising yoga positions. To ensure robustness, the dataset comprises photographs taken from a variety of angles and lighting situations. The dataset contains images of yoga practitioners performing six different yoga poses:

- Camel Pose, Downward Facing Dog (Downdog), Goddess Pose, Plank Pose, Tree Pose, Warrior II Pose.
- A total of 1693 images were used belonging to six classes.

Tabel 1: Dataset Description

TRAINING	1396
TESTING	297

Due to a lack of photos for some classes, data augmentation techniques were used to improve the dataset's diversity and size. Rotation, flipping, scaling, and translation were employed to maintain the integrity of the yoga poses while producing additional training examples.Each image is labelled with the corresponding yoga stance to aid in supervised learning tasks like pose identification or categorization.

Usage:

- Classification of yoga poses in images.
- Pose estimation and tracking in videos.
- Human-computer interaction applications related to yoga practice. Below are sample images representing each class in dataset:



Camel pose Downdog pose Goddess pose



Plank pose Tree pose Warrior II pose

Fig 2: Dataset Images

3.4. CNN Architecture

Convolutional Neural Networks (CNNs) are a powerful type of DL models that excel at extracting characteristics from spatial data, such as photographs. They excel at tasks such as picture classification, object detection, and stance estimation, making them ideal for yoga pose prediction.

Key Components:

- Convolutional layers: Extract picture features such as edges, contours, and body part placements.
- Pooling Layers: Increase efficiency by reducing the size of the photograph while retaining all relevant information.
- Dropout Layers: During training, activations are dropped at random to prevent overfitting.
- Flatten Layer: Converts multidimensional feature maps to a single vector.
- Dense Layers: Perform classification using extracted features.
- Softmax Activation: predicts the most likely yoga pose by transforming output into class probabilities.

Training:

As it trains, the CNN gains the ability to:

- Make a note of characteristics that are crucial for identifying between different yoga positions. (e.g., body part placements, orientation).
- To classify poses, integrate these features into the fully-connected layers.

Prediction:

When presented with a new image:

- The CNN extracts features using the trained convolutional layers.
- Based on learned relationships, the fully-connected layers predict the most likely yoga pose.

3.5. VGG16 Architecture

VGG16 is a convolutional neural network (CNN) architecture widely used for image classification. Its approach combines many small convolutional filters with a 3x3 kernel size to achieve excellent accuracy while prioritising simplicity. With this strategy, VGG16 can efficiently learn basic characteristics (lines, edges) in the top layers and progressively capture more intricate information (shapes, object components) in the lower layers.

.5.1. Data Preprocessing:

The initial step for our model probably is receiving image data from folders and transforming it, maybe by scaling and normalising it.

.5.2. Feature Extraction with VGG16:

.5.2.1. Our method uses the include_top=False parameter to use a pre-trained VGG16 model. This suggests that we are just utilising VGG16's convolutional layers, up to the last fully-connected layer.

.5.2.2. From the input pics, these convolutional layers extract features. With the use of substantial data sets such as ImageNet, VGG16 is trained to acquire general picture features that can be used for a range of classification tasks.

.5.3. Freezing VGG16 Layers(optional):

The VGG16 layers are iterated through and their trainability is set to false (layer.trainable = False). Using pre-trained models for transfer learning is a widespread approach. The learning process is directed towards the recently added layers as it stops the pre-trained weights from being overwritten during training.

.5.4. Classification Layers:

.5.4.1. The pre-trained VGG16 model serves as the foundation for a sequential model. Using the features extracted by VGG16 as input, this new model applies them to your particular classification problem.

.5.4.2. The layers include:

.2.1. Flatten: Creates a 1D vector from the 3D feature maps from VGG16.

.2.2. Dense layers: These are fully-connected layers used for classification.

.2.3. For multi-class classification, the final Dense layer has a softmax activation function with 6 units, which correspond to the 6 classes.

Training:

- The model receives snapshots of yoga poses along with labels while undergoing training.
- The pre-trained VGG16 layers extract features from photos. The classification layers learn how to associate these attributes with specific yoga poses by altering their internal weights in response to the comparison of predicted and actual labels.
- The model's ability to differentiate distinct yoga positions improves with time based on the features collected by VGG16.

Prediction:

- The pre-trained VGG16 layers extract features from the image.
- The classification layers utilize the learned relationships between features and yoga poses to predict the class (yoga pose) with the highest probability.
- The softmax activation function outputs probabilities for each yoga pose class, indicating the model's confidence level in its predictions.

3.6. User-friendly Interface (UI):

User-friendly platform is implemented and this user interface enables easy yoga pose prediction using image requests. Users can select an image and one of two models: VGG16 or a custom CNN model. When you click the "Submit" button, predictions are generated and the uploaded image is displayed alongside the anticipated pose and confidence score. This interface allows for simple interaction and clear visualisation of pose detection findings.

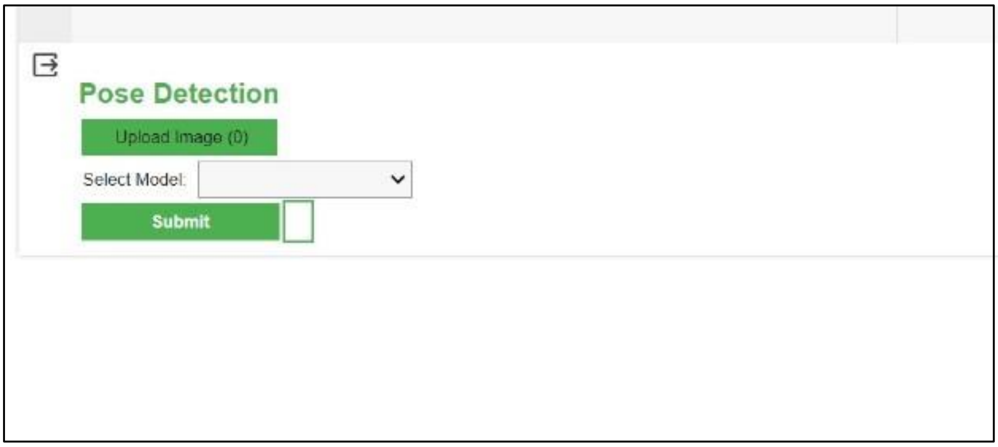

Fig 3: UI Input

Fig 4: UI Output

IV. RESULT

The results obtained from our model's is evaluated by its training and testing accuracy.

able 2: Accuracy

MODEL	TRAINING ACCURACY	TESTING ACCURACY
Convolutional Neural Network	**95%**	**84%**
VGG16	**99%**	**95%**

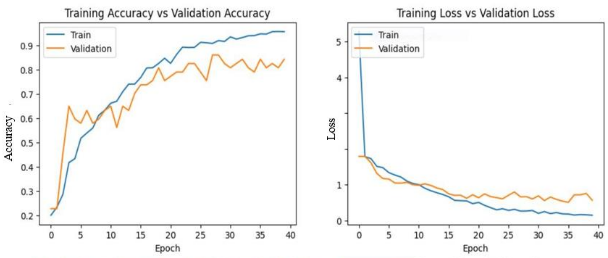

Fig 5: Accuracy and Loss Curve for CNN

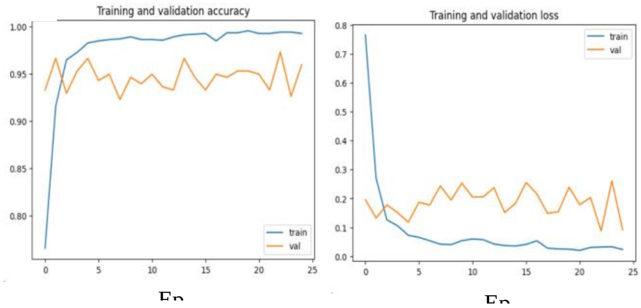

Fig 6: Accuracy and Loss Curve for VGG16

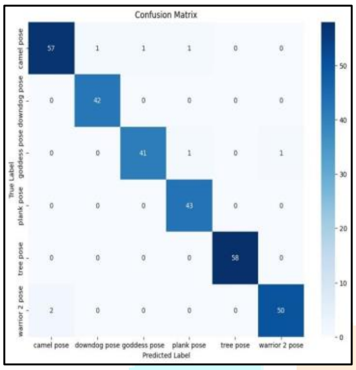

Fig 7: Confusion matrix for CNN

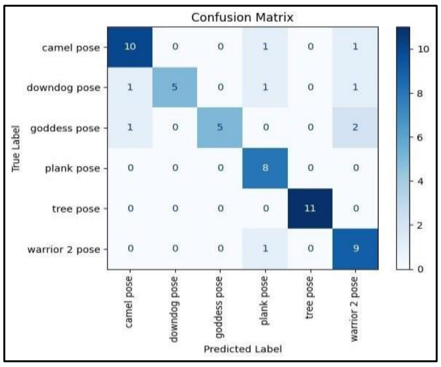

Fig 8: Confusion matrix for VGG16

V. CONCLUSION

To summarise, the established interactive system for classifying images for yoga postures is a valuable resource for yoga practitioners, instructors, and aficionados alike. Users may precisely recognise and correct positions using DL models, which improves their yoga practice experience. The project uses pre-processing techniques like rescaling, normalizing and many more. It employed CNN and VGG16 with 95% and 99% respectively, to identify yoga poses. After being trained and tested using the accuracy measures, the results demonstrated that the models were capable of correctly identifying yoga poses. The system's user- friendly interface and accurate forecasts make it accessible and useful to people of all skill levels. Overall, this project advances yoga practice and promotes well-being through technological solutions.

VI. REFERENCES

[1.] R. Dutta and S. Mukherjee, "Smart Yoga Assistant: A Mobile Application for Correcting Yoga Posture Using Pose Estimation," 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT), Kharagpur, India, 2020.

[2.] S. Sathya and S. Krishnan, "Yoga Gesture Recognition using Convolutional Neural Networks," 2020 Fourth International Conference on Inventive Communication and Computational Technologies (ICICCT), Coimbatore, India, 2020.

[3.] S. Saha and M. K. Kundu, "Automated Recognition of Yoga Asanas Using Convolutional Neural Networks," 2020 International Conference on Computational Intelligence and Communication Systems (ICCICom), Kolkata, India, 2020.

[4.] Narayanan, S. Sankara and Misra, 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.

[5.] Thoutam, Vivek Anand, et al. "Yoga pose estimation and feedback generation using deep learning." Computational Intelligence and Neuroscience 2022 (2022).

[6.] Anilkumar, Ardra, et al. "Pose estimated yoga monitoring system." Proceedings of the International Conference on IoT Based Control Networks & Intelligent Systems-ICICNIS. 2021.

[7.] Palanimeera, J., and K. Ponmozhi. "Classification of yoga pose using machine learning techniques." Materials Today: Proceedings 37 (2021): 2930-2933.

[8.] Kothari, Shruti. "Yoga pose classification using deep learning." (2020).

[9.] Talaat, Amira Samy. "Novel deep learning models for yoga pose estimator." SN Applied Sciences 5.12 (2023): 341.