Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org
© 2024 SCPE. Volume 25, Issues 2, pp. 739–750, DOI 10.12694/scpe.v25i2.2590

A METHOD FOR SPECIFYING YOGA POSES BASED ON DEEP LEARNING, UTILIZING OPENCV AND MEDIA PIPE TECHNOLOGIES

ANURADHA T,[*] N. KRISHNAMOORTHY,[†] C S PAVAN KUMAR,[‡] L V NARASIMHA PRASAD [§] CHUNDURU ANILKUMAR ,[¶] AND USHA MOORTHY[*‖]

Abstract. Yoga is a years discipline that calls for physical postures, mental focus, and deep breathing. Yoga practice can enhance stamina, power, serenity, flexibility, and well-being. Yoga is currently a well-liked type of exercise worldwide. The foundation of yoga is good posture. Even though yoga offers many health advantages, poor posture can lead to issues including muscle sprains and pains. People have become more interested in working online than in person during the last few years. People who are accustomed to internet life and find it difficult to find the time to visit yoga studios benefit from our strategy. Using the web cameras in our system, the model categorizes the yoga poses, and the image is used as input. However, the media pipe library first skeletonizes that image. Utilizing a variety of deep learning models, the input obtained from the yoga postures is improved to improve the asana. The algorithms like VGG16 (Visual Geometric Group), VGG19, Convo2d, CNN.

Key words: Deep Learning, Media Pipe, OpenCV, VGG16, VGG19, Skeletonization.

1. Introduction. Yoga is a type of exercise is not just those looking to stay fit and healthy but for those who seek harmony in their lives. An admixture of asanas and postures, yoga is aimed to help with inflexibility, attention, and calmness. From bending and stretching to contemplation, yoga helps to regulate ails, reduce weight, and attack issues like stress and anxiety. Yoga is the ideal drug for those who like to lead healthy lives and make a connection with body, mind, and spirit Mortal posture assessment is a delicate issue in the control of PC vision. It manages confinement of mortal joints in a picture or videotape to shape a cadaverous depiction. To accordingly fete an existent's posture in a picture is a worrisome errand as it relies upon colorful perspectives, for illustration, scale and thing of the picture, enlightenment variety, foundation mess, dress kinds, environmental factors, and connection of people with the environmental factors. The external gadget with a built-in algorithm that recognizes body corridors and markers as a result in each depth image. The approach is based on arbitrary decision trees that were trained from a sizable database of artificial body part marker images. Several mortal body models and a database of mortal stir prisoner data are used to create the synthetic training images. The diversity and unpredictability of the stir prisoner data reveal the efficiency of the body part recognition system. For instance, the training data set's unbalanced distribution of data in one stir delicacy bloody compared to the other orders will have a poor impact on the recognition performance. It should be carefully designed to collect the training data somewhat from the area of mortal activities. In real life, negotiating this is difficult. Many locomotion's, gestures, and standing behaviors can be easily acquired using marker-grounded stir training and are included in stir databases that are open to the public. However, because of marker occlusion, landing behaviors that need the person to thicken, sit, kneel, or bend are more taxing and do not yield as many data as standing behaviors. A sizable collection of stir data from various stir orders is used to train the recognition system, however only some stir orders may have enough diversity. The foundation for posture evaluation and recognition is the information-gathering of the entire body. In order to collect posture information using colorful detectors, numerous biases have been devised. The most popular detectors are inertial dimension units, wearable entrance markers, picture detectors, and electromyogram detectors (IMUs).

[*]Department of Information Technology, VR Siddhartha Engineering College, Vijayawada, Andhra Pradesh, India
[†]Department of Computer Science and Engineering, Kongu Engineering College, Perundurai, Erode, Tamilnadu 638060, India
[‡]Department of Information Technology, VR Siddhartha Engineering College, Vijayawada, Andhra Pradesh, India.
[§]Department of Computer Science and Engineering, Institute of Aeronautical Engineering, Hyderabad, Telangana, India.
[¶]Department of Information Technology, GMR Institute of Technology, Rajam, Srikakulam, Andhra Pradesh, India.
[‖]Department of Computer Science and Engineering, REVA University, Bengaluru, India (usha.m@reva.edu.in)

In general, people do not need to wear redundant equipment while using image detectors, picture other side be easy body pose should not cause image occlusion

The contribution of the paper is summarized below: -

1. A useful system for categorizing yoga positions that can be practiced at home unaccompanied. The general pose discovery model is replaced by the precise and reliable Google Media Pipe library for body key point recognition.
2. To make the system less amenable to different conclusions, the model employs training and testing images with different backgrounds. The system allows for background changes and disturbances, such as exposure to bright light, because the background in each image is the same.
3. It makes use of a sizable dataset made up of pictures of yoga poses for testing, training, and validation.
4. Future research on yoga pose detection will be made possible by comparative analysis of the various models.
5. A precise and low-latency model is provided for real-time settings for practitioners who are housebound.

2. Related Work. The 3D CNN architecture's planned structural elements. The 3D CNN armature's suggested structural elements. also shows the developed network's configuration data, including the large count of parameters corresponding to the different layers, kernel sizes, and the overall size of the point maps. As depicted in the mentioned figure, the input to the specially constructed containing 3D CNN as an image cell made up of sixteen frames with a resolution of containing 112 x 112 x 3 pixels that were taken from the video's inputs of yoga poses. Neural network, which is made up of different layers 8 Conv3D layers ,5 MaxPool3D layers and AvgPool3D subcaste, reuses the input picture cell after classifying one of the ten Yoga actions using a SoftMax classifier. Each of the 3 x 3 x 3 kernels used in the convolutional layers corresponds to aimlessly initialized pollutants. During convolution, the pollutants apply a stride of one along the spatial and temporal boundaries with border mode set as valid. For the network to increasingly prioritize discriminating features from the input video of the Yoga positions, the number of pollutants in the convolutional layers is gradually raised from 64 to 512 [12]. Our slice technique produces stir data, which is used to train arbitrary decision trees that automatically annotate input depth photos. The bodily portion to which each pixel in depth pictures corresponds is identified by a label. Reliable offers for the locations of 3D cadaverous joints are the outcome of the algorithm for recognizing mortal disguises. A decision tree labels pixels in a rather noisy manner. A noisy labelled image built on mean shift offers the cypher common [23]. the use kinematic shell limitations to enhance the delicacy and robustness of typical offers. The severity of the bones and their fixed connection are used to infer the kinematic restrictions. A femur, for instance, connects a hip and knee joint and places restrictions on the range of motion for each joint. The depth values between the connected joints should fluctuate linearly from the knee to the hipsterism within an error threshold since they are intended to measure on the face of the ham. We can recognize body corridor mislabeling thanks to these restrictions [21]. The model was written in Python using the Keras Successional API. The shape of the input case is 45*9*18, which represents 45 success frames with 18 key points that each include X and Y coordinates. ReLU activation is having to key spots of every frame for point birth in time-distributed CNN subcaste with 16 contaminants of size 3* 9* 3. CNNs are better able to value scale- and gyration-steady spatial characteristics. The CNN subcaste can value spatial characteristics like the angles and distances between the vibrant focal points in a frame. The CNN affair is subjected to batch normalization for quicker convergence. This is a powerhouse subcaste that prevents overfitting by dropping some weights at random. The CNN data is applied to each of the 45 frames and smoothed before being fed with 20 units and a 0.5 unit forget bias to the LSTM sub-caste. The CNN subcaste uproots certain features, and the temporal changes in the features are recognized by LSTM [22]. A body posture modeling and evaluation system to celebrate and assess yoga poses and suggest instruction to learners is proposed in this study. A two-stage classifier was built using FCM and BP-ANN to model and celebrate full-body poses. Because of its crucial nonlinear processing, BP-ANN Detectors recommended as the first classifier to classify yoga stations into distinct orders capability, and FCM was espoused as the successive classifier to categorize the postures in a order for the flexible fuzzy partition. The two-stage classifier could handle the variations in subject position while improving recognition outcomes at a reasonable computational cost. Extension: To further improve the recognition delicacy, we also suggested a recognition technique to reject insulated honored products and noisy results with accretive probability [8]. The article suggests a system for correcting posture or poses. With

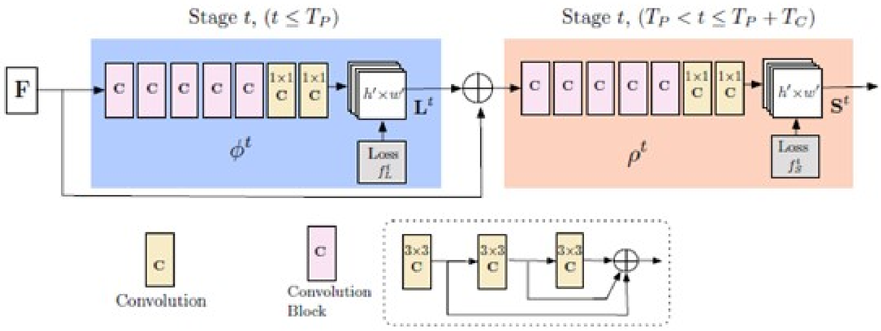

Fig. 2.1: Open Pose Architecture Model

YogaST, three categories of poses have been created using the kinect technology: Downward Dog, Tree Pose, and Warrior Pose are examples. A yoga expert system has been created using texts and graphics to direct the practitioner. the approximately 5500 photos in the YOGI collection [14]. Angles and joints from these photo frames were fed into various machine learning models, including Logistic Regression, Random Forest, and Support Vector Machine, using the ”tf-posture” technique (SVM). In the experiment, skeletal characteristics of the human posture in the photo frame were extracted using Open Pose architecture, and they were then stored in a NumPy array. Its model is a straightforward neural network architecture with two hidden layers and an output layer [5]. The posture is estimated with a help of web camera that helps us in capturing the real time image. They took into consideration 12 sun salutation asana poses that were recorded on a webcam. The method creates a skeletonized image and on that image feature extraction is applied and through ML model a sun salutation is classified along with Support Vector Machine (SVM), k-nearest neighbors (KNN), Logistic regression and Naïve Bayes producing 96% of the result [9].

The VGG 19 convolutional neural network is the model utilized for the experiment. It comprises 19 deep layers that are made up of a successional composite of complication, corrected direct unit activation, and Max Pooling layers. illustrates the VGG-19 model's connected layers. The size of the input subcaste is 224 x 224 x 3, which corresponds to the height, range, and RGB (red, green, and blue) channel separately [2]. the interconnected successional layers that make up our final experimental model. Figure 8 shows the Transfer learning VGG 19 model's armature. A SoftMax activation function that conforms to five class markers for each of the a fore mentioned yoga acts makes up the affair subcaste. The Adam optimizer, which combines grade descent with incitement and root mean square propagation, is used in this model. Since it is a multiclass bracket model, the Adam optimizer's literacy rate is set to 0.001 and its loss function is categorical cross entropy [1]. The model's performance is covered with a batch size of 16 and several ages or passes set to 25. demonstrates the VGG19 model's performance standards on the train and test datasets. We notice a more suggestive change in the model's delicacy after five ages or passes, and it plateaus at 98 percent delicacy after twenty-five [13].

The suggested model's DCNN armature is like the models. The three layers (convolutional layers) along with the max-pooling layer with a 2x2 window size and of 2 strides make up the feature extraction portion. The first three convolutional layers comprise 32(first layer), 128(second layer), and 256(third layer) each with a size of 3x3 and 1 stride. Because there are three maximum pooling layers used in the armature, the point feature extraction creates 256-point maps that are three times smaller (3*1) than the size of the input image (24*8). likewise, the fully interconnected layers receive these qualities (the bracket part). The first, alternate, and third layers, which make up the bracket portion, respectively have 128, 64, and 32 neurons each and use the ReLU activation function. However, the final subcaste uses the SoftMax activation function and contains 26 neurons per labor. Since we used a tenfold cross-validation approach during the training procedure, ten DCNN models were trained for 100 ages [11]. This study employs Convolutional Neural Networks (CNNs) for the identification of various poses. Utilizing Kinect, the study investigates a range of image combinations, including RGB, Depth, RGB-D, and background-subtracted RGB-D [4].

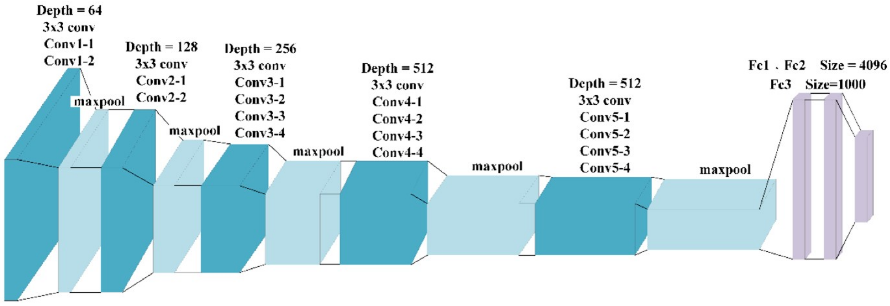

Fig. 2.2: VGG19 Architecture

A custom dataset was created by capturing diverse activities carried out by various individuals in various indoor environments. The findings indicate that the most effective approach for indoor video-based fall monitoring involves combining RGB background-subtracted and Depth with CNNs [3]. In this research paper, we introduce a system designed for recognizing Yoga postures. This system not only identifies the specific Yoga pose being performed by a practitioner but also retrieves relevant Yoga training information from the Internet to provide guidance on correct posture [15]. To achieve this, we initially employ a Kinect device to capture the user's body map and extract the body contour. Subsequently, we utilize 'star skeleton,' a rapid skeletonization technique that connects from the centroid of the target object to the contour extremes. This star skeleton serves as a representative descriptor for recognizing the practitioner's Yoga posture.

Lastly, the system retrieves Yoga training information associated with the recognized posture from the Internet, assisting the practitioner in maintaining the correct posture during their practice [20]. This paper introduces a system designed to track and assess the precision of various yoga poses, thereby assisting users in their yoga practice. The system utilizes Microsoft Kinect to capture real-time data on the movement of different body parts, allowing for the calculation of angles based on the detected joint points. These angles are then employed to evaluate the accuracy of specific yoga poses for the user. The system has demonstrated the capability to promptly identify and assess a variety of yoga poses in real-time [7]. This paper aims to construct a model for training data that aligns with human posture characteristics, thereby breaking down this complex problem to reduce computational complexity and enhance system performance in practical applications. Through real-world experimentation, the model identifies distinct body movement postures by observing sequences of human postures, employing matching, identification, and classification processes [19]. The results affirm the feasibility and effectiveness of the proposed method for human posture recognition. Furthermore, for detecting human movement targets, the paper introduces a method based on Gaussian mixture modeling, and for extracting human object contours, it proposes an approach based on the Sobel edge detection operator.

Additionally, this study includes an experiment on human posture recognition and an assessment of our cloud-based monitoring system for elderly individuals utilizing our methodology [18]. The proposed system begins by detecting humans before conducting posture analysis, restricting the posture recognition process to human silhouettes. The human detection technique is intentionally designed to exhibit resilience to diverse environmental factors. As a result, posture analysis relies on straightforward and efficient features that are intended for characterizing human silhouettes without considering environmental constraints [10]. The posture recognition method, employing fuzzy logic, can identify four static postures and is robust against variations in the camera-to-person distance and variations in a person's body shape. With a satisfactory posture recognition accuracy of 74.29%, this approach can effectively identify emergency situations, such as a fall, within a health-oriented smart home [17]. Visual Gesture Builder, a data-driven machine-learning tool for detecting gestures, was employed to accurately capture essential yoga movements. This gesture analysis technology is currently under investigation for potential integration into exergames designed for personalized medical interventions. The research objective is to evaluate whether a machine learning algorithm embedded in a basic computer-based video exergame can evaluate the acquisition of yoga skills in specific target populations, thereby promoting

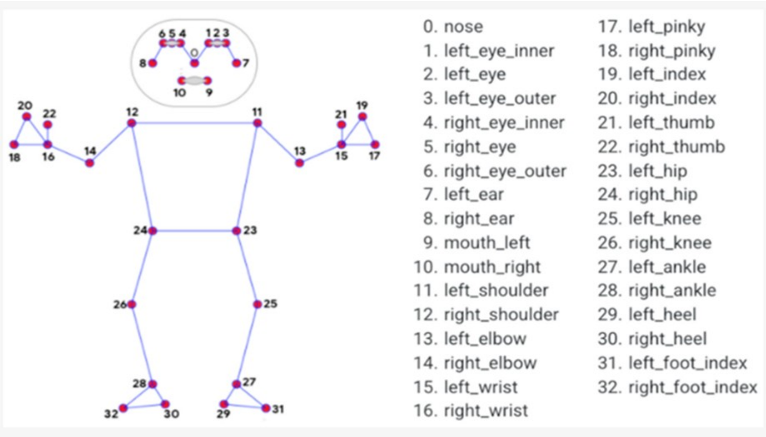

Fig. 3.1: Land Marks in Media Pipe Pose

healthy physical activity [6]. This paper introduces an alternative, computationally efficient method for recognizing Yoga poses in complex real-world settings using deep learning. In pursuit of this goal, a Yoga pose dataset was curated, featuring the participation of 27 individuals (comprising 8 males and 19 females) and encompassing ten Yoga poses: Malasana, Ananda Balasana, Janu Sirsasana, Anjaneyasana, Tadasana, Kumbhakasana, Hasta Uttanasana, Paschimottanasana, Uttanasana, and Dandasana. The videos were recorded using smartphone cameras with a 4K resolution and a 30-fps frame rate. For real-time Yoga pose recognition, a three-dimensional convolutional neural network (3D CNN) architecture was devised and implemented. This architecture is a modified version of the C3D architecture initially designed for human action recognition. In the proposed modified C3D architecture, computationally intensive fully connected layers were pruned, and additional layers like batch normalization and average pooling were introduced to enhance computational efficiency. To the best of our knowledge, this study is among the first to leverage the inherent spatial-temporal relationships among Yoga poses for recognition. The designed 3D CNN architecture achieved a test recognition accuracy of 91.15% on the in-house curated Yoga pose dataset containing ten distinct Yoga poses [16].

3. Data Collection. All the videos contain amassed for greater than 40 seconds in an enclosed space at a frame rate of 35 frames every second. An overall of 88 training videos are 1.5 hour, 5 minutes, and 5 seconds long at 35 frames every second, or roughly 750 frames. For every terminal state of each pose, roughly a hundred samples should be collected in order to create an effective classifier. It is critical that the samples used in the collection reflect a range of camera perspectives, landscape features, body types, and position variations. 33 pose landmarks are predicted by the MediaPipe Pose data set landmark model.

The data set body mark model in Media Pipe Pose predicts the area of overall pose landmarks nose, left_eye_inner, left_eye, left_eye_outer, right_eye_inner, right_eye, right_eye_outer, left_ear, right_ear, mouth_left, mouth_right, left_sholder, right_shoulder, le- ft_elbow, right_elbow, left_wrist, right_wrist, left_pinky, right_pinky, left_index, left_thumb, right_thumb, left_hip, right_hip, left_knee, right_knee, left_ankle, right_ankle, left_heel, right _heel, left_foot_index, right_foot_index A full-body segmentation mask displayed as a two-class segmentation can be predicted by Media Pipe Pose (human or background).

4. Methodology.

4.1. Neural Network Architecture. This model majorly works with neural network where the model tries to imitate the human brain to recognize the yoga posture. Initially, after the skeletonization of the image or the real time image, the model grasps the random inputs from the input nodes and passes through the hidden layer providing the output. But before resulting in the output the model tries to recognize the detected model is relevant to its posture or not through the other proposed algorithms and goes through backpropagation method. As the human brain tries to learn from the errors, the model imitates the same format of human learning from

Table 3.1: Key points Used

No	Key Point	No	Key Point
0	Nose	17	left_pinky
1	left_eye_inner	18	right_pinky
2	left_eye	19	left_index
3	left_eye_outer	20	right_index
4	right_eye_inner	21	left_thumb
5	right_eye	22	right_thumb
6	right_eye_outer	23	left_hip
7	left_ear	24	right_hip
8	right_ear	25	left_knee
9	mouth_left	26	right_knee
10	mouth_right	27	left_ankle
11	left_shoulder	28	right_ankle
12	right_shoulder	29	left_heel
13	left_elbow	30	right_heel
14	right_elbow	31	left_foot_index
15	left_wrist	32	right_foot_index
16	right_wrist		

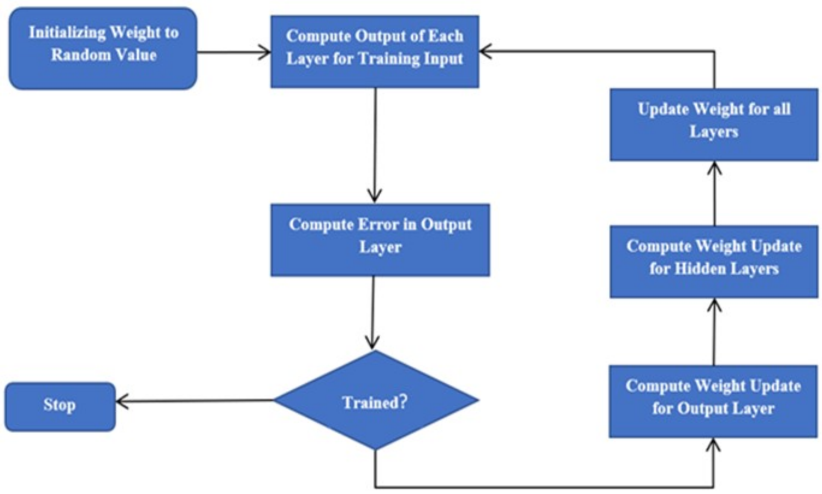

Fig. 4.1: Neural Network Architecture

errors and mistakes. The model goes through all these phases and recognizes the exact yoga posture and using media pipe the points allocated to each phase and part of the body, the model reads the skeletonized image and through the angles between the points allocated to joints and guides us in practicing the right posture.

A major portion of the exploration work on the Yoga Dataset has been carried out using Convolutional Neural Networks (CNN). A complication is to input as a sludge performing in activation. It can learn a huge number of pollutants under constraints specific to a training dataset. The armature substantially comprises colorful layers with pollutants and activation functions. The convolutional subcaste applies pollutants to the input image. Pollutants are vital in armature since they descry the spatial patterns in an image grounded on the change in intensity values of pixels. A pooling subcaste is added to reduce computational costs by

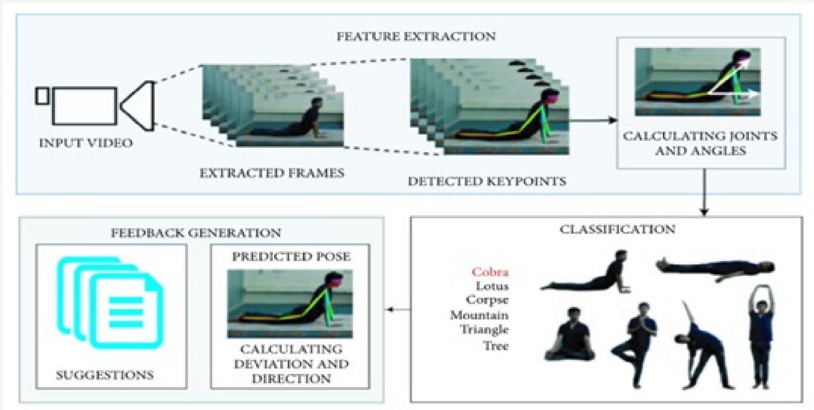

Fig. 4.2: Pose Extraction

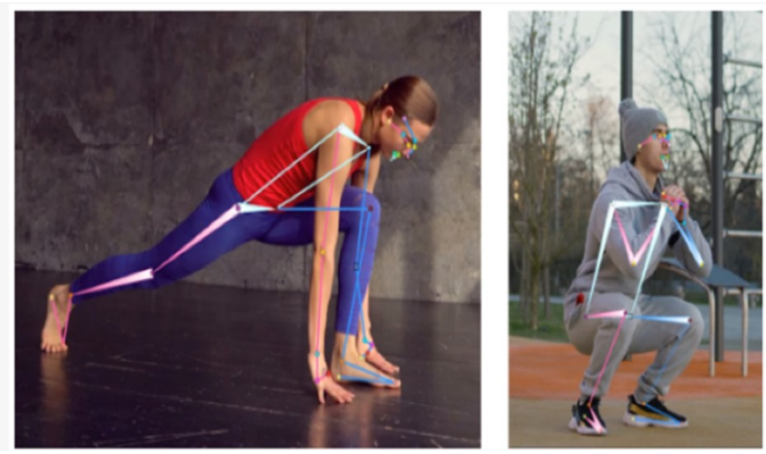

Fig. 4.3: Media Pipe Key points on Actual Humans

dwindling interlayer connections. Weights and impulses are added to completely connected layers, which connect neurons from different layers. Powerhouse layers are added to drop neurons, reducing the size of the model. Experimenters have experimented with different layers and activation functions in their work. Shows the armature of a Convolutional neural network.

From figure 4.2 describes about Keyframes are created from each frame of the videotape in the first phase, which is followed by their storage in JSON format. Key points are various body parts that make up a person and are important for the formation of a yoga disguise; examples include the shoulders, elbows, wrists, knees, etc. For important aspects, we used the Media Pipe library, a cross-platform toolkit created by Google that offers incredible ready-to-use ML results for computer vision applications.

A highly optimized, trained CNN model is utilized in this phase to infer thirty-three 3D milestones and entire body from RGB video tape masks with the background segmentation frames. This phase also uses a CNN model for high-dedication body position shadowing. Three equals are generated by the Media pipe library, with Z denoting the depth of a 2D match. (10). Fig. 3.1 displays the 33 key points provided by the Media Pipe library, whereas Fig. 4.1 shows the outcome of using the key points to root the data.

4.2. Training. In this paper, we have analyzed the yoga pose classification techniques in detail and have also created a model to contrast with the existing approaches. While most of the models exceed expectations

Table 5.1: Accuracy of Various Classes

S.No	Name	Class	Accuracy(%)
1	Validation	Framewise	98.45
2	Training	Framewise	97.56
3	Real Time	polled	99.05
4	Test	Polled	96.34
5	Test	Framewise	97.89

with higher performance metrics, the dataset needs to be more mature. An ideal dataset should be diverse and should reflect real life as much as possible. Overfitting and Underfitting a machine learning model can perform worse on newer data and leads to poorer accuracy and other performance metrics. Data Augmentation techniques can help in producing more data and thereby improve the size and the quality of the dataset For the visualization of asanas, features are carried out using Open Pose and the common position values are recorded in the JSON train. CNN and LSTM (Long Term Short Memory) models are also connected. Due to the combination of the two, we can recognize the fashionable set of data created by CNN (Convolution Neural Networks) and the data dependencies for long term created by LSTM. Theano's backend and Keras are used to collect the model. The performance of the entirely linked subcaste's relationship with SoftMax activation is scaled using the categorical because of its suitability cross-entropy loss function happened. Because of the straightforward inputs and small design, the training is quick at about 22 seconds every time. Over the course of training, both the modify in delicacy and the loss function independently. Initially, the accuracy of both increases fleetly with training accuracy staying below the validation accuracy The variations for training and confirmation accurateness, and training and validation misplacements throughout network training. The both accuracy increases steadily for the first 10 years, as depicted in the figure. however, a marked difference in accuracy can be seen in later ages. Also, it is possible to see that the validation accuracy detects the training accuracy, shows that sufficiently generic model is and will render the chosen Yoga poses coherently on test or unseen Yoga options. Once the network converges for the 25th time, none of the two measures—training accuracy nor validation accuracy—undergoes any substantial change (the model becomes pregnant). At the fifteenth repetition, the errors and misplacements transfer their asymptotic values. Nevertheless, When the model was impregnated, training was stopped after another 18 epochs, with the validation accuracy remaining unchanged. As a result, the Model Checkpoint call back job automatically saves the model of the weights corresponding to the time having good validation accuracy, for conclusion test data of yoga disguise input videos. In conclusion, our developed model successfully fitted the dataset of difficult real-world Yoga poses from our own set during training within 25 epochs.

5. Results and Discussion. The perpetration of other transfer literacy models for image classification similar as ResNet50, Inceptionv3 and Efficient Net. Advanced data addition ways can be used to give added quality- concentrated data which aids the model's interpretation. Although it is computationally precious, tentative GANs and neural style transfer styles transfigure an image from one sphere to another furnishing robust data to feed the neural network models. The count of acts used to develop in the data set to aggrandize the dataset's quality and bring further credibility to the use of the yoga classification model. Table 5.1 represents the accuracy of various classes.

The work provides self-train systems for recognizing and correcting the posture into consideration and uses the RGB webcam. In our data consisting the yoga postures based on the image of the system performed well and giving instructions to correcting the yoga pose includes totally various from each and every appearance, it gains an total accuracy of 90% in feature extraction.

They used homemade point birth and make a distant model for every yoga pose, which performs and gives lower time count devouring to apply and requires a good number of features add . In our route a good order of Yoga is reachable as can append any neuron in every t thick subcaste and again train the model on the substitute data in the dataset. The given model contains 99.38% is also prideful to containing model as shown

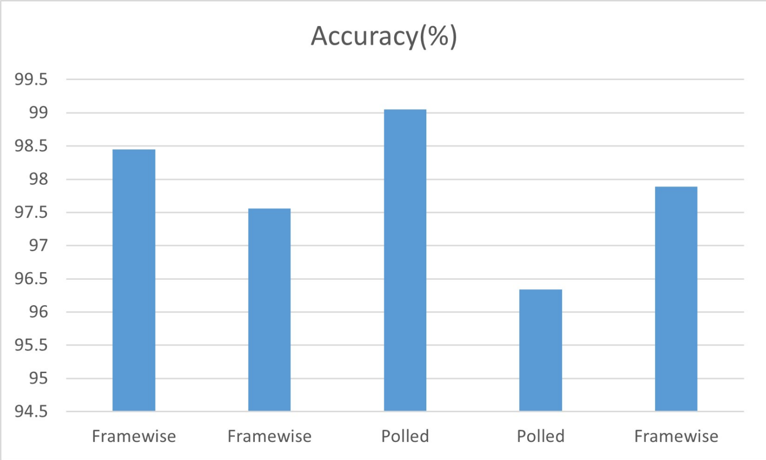

Fig. 5.1: Comparison of Accuracy for various classes

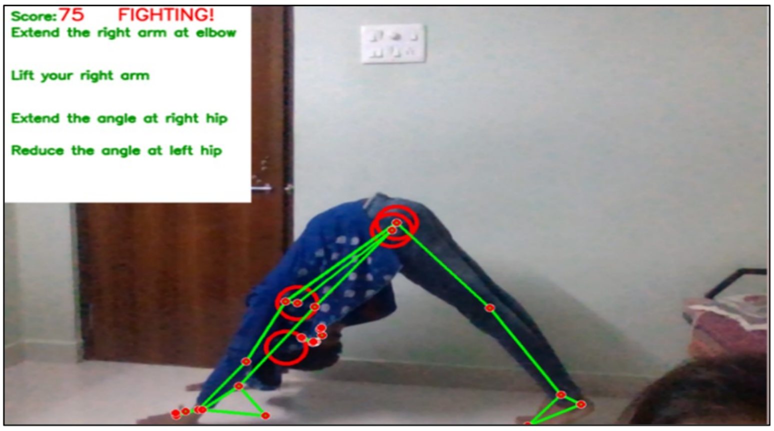

Fig. 5.2: Adho Mukha Svanasana

in figure 5.1.

5.1. Real Time Yoga Posture. From fig 5.2 Based on the trained data, the model suggests the person instructions to make the posture perfect. By comparing the real time practice to the trained practice with the body points and angles between the joints, the model has instructed the person to lift his right arm, extend the angle at right hip, and to extend the right arm at elbow giving the score 75.

From figure 5.3 this posture got some different angles compared to the previous yoga pose. These angles are now to be compared to the real time practice. As we can see the both arms of the person, the left arm is a bit tilted whereas the other arm is straight. The model instructed the person to lift his right arm and his left arm.

From figure 5.4 the person is instructed on lifting his left and right arm and extending and reducing the angle at his right and left hip to make the posture perfect giving the score 77

From figure 5.5 the current posture lacks some details from the actual posture and the person is instructed on extending his both right and left arms reducing the angles at his right hips and knees and extending the right leg more. According to the score 67, the person needs to work on this posture more.

From figure 5.6 the practitioner got advised about lifting his both arms to make the namaskar gesture

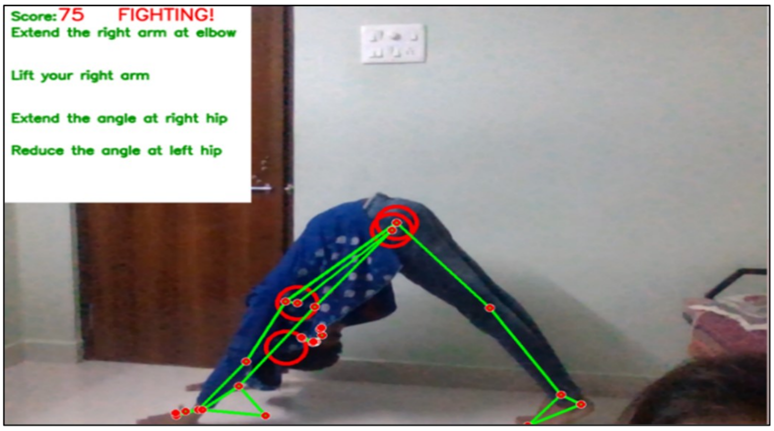

Fig. 5.3: Virabhadra Asana

Fig. 5.4: Adho Mukha Svanasana

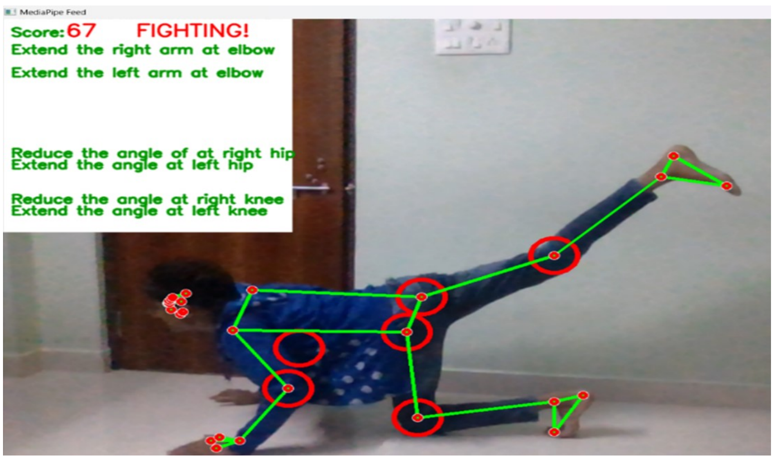

Fig. 5.5: Marjarya Bitil Asanaa

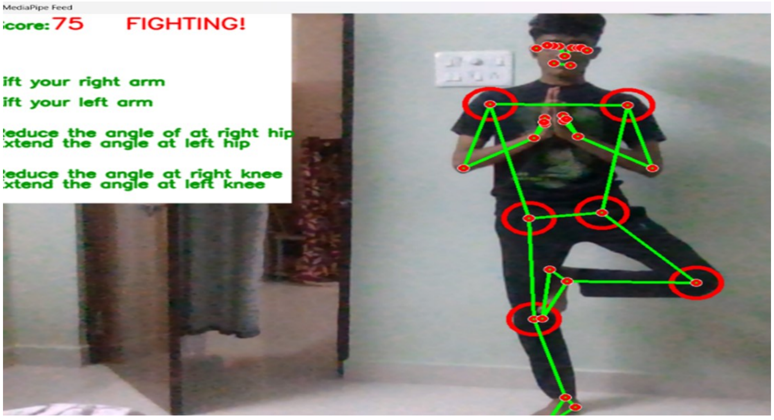

Fig. 5.6: Vrksasana Pose

straight and perfect and to make his right leg straight. And about the left knee, the person still needs to increase the space between the knee that is to increase the angle.

6. Conclusion. We have created a system in this design that consists of a channel for disguise identification, point localization on the mortal body, and an error identification method. This technique attempts to aid individuals in correctly yoga practice on their own and assist with ailments that may result from improper yoga poses. The approach using the algorithms CNN (Convolution Neural Networks) and the algorithm LSTM (Long Short-Term Memory) on disguise data attained to the library of Open Pose for Yoga actions discovery and set up to be largely successive. A device identifies the all poses on recorded videos as of the all activities of real time for every person. The devices gained detects Yoga shows in a videotape with 99.04% for framewise and 99.38% after polling process of each 45 frames. The network device scored 97.92% in real time activity of different people shows its capability to achieve easy more contents and conditions. The approach is suitable for evaluating the stoner's disguise from the front and providing feedback so that they can improve their yoga disguise using deep literacy techniques. The designed model is mounted atop a dashboard that was likewise made with stoners in mind It would be commented that our route predicates the need for any different technical tackle for Yoga pose adjustment and can be enforced on taken input from employing RGB camera. additional yoga postures, more datasets, and real-time image and video activity are all included. The device can be mounted on a mobile for tone-training and real-time forecasting. This work improves as proof of effort identification.

REFERENCES

[1] K. Adhikari, H. Bouchachia, and H. Nait-Charif, *Activity recognition for indoor fall detection using convolutional neural network*, in 2017 Fifteenth IAPR International Conference on Machine Vision Applications (MVA), IEEE, 2017, pp. 81–84.

[2] U. Bahukhandi and S. Gupta, *Yoga pose detection and classification using machine learning techniques*, Int Res J Mod Eng Technol Sci, 3 (2021), pp. 13–15.

[3] D. Brulin, Y. Benezeth, and E. Courtial, *Posture recognition based on fuzzy logic for home monitoring of the elderly*, IEEE transactions on information technology in biomedicine, 16 (2012), pp. 974–982.

[4] X. Cai, Y. Gao, M. Li, and W. Song, *Infrared human posture recognition method for monitoring in smart homes based on hidden markov model*, Sustainability, 8 (2016), p. 892.

[5] A. Chaudhari, O. Dalvi, O. Ramade, and D. Ambawade, *Yog-guru: Real-time yoga pose correction system using deep learning methods*, in 2021 International Conference on Communication information and Computing Technology (ICCICT), IEEE, 2021, pp. 1–6.

[6] H.-T. Chen, Y.-Z. He, C.-C. Hsu, C.-L. Chou, S.-Y. Lee, and B.-S. P. Lin, *Yoga posture recognition for self-training*, in MultiMedia Modeling: 20th Anniversary International Conference, MMM 2014, Dublin, Ireland, January 6-10, 2014, Proceedings, Part I 20, Springer, 2014, pp. 496–505.

[7] B. Donahoe-Fillmore, C. J. Brahler, M. I. Fisher, and K. Beasley, *The effect of yoga postures on balance, flexibility, and strength in healthy high school females*, Journal of Women's & Pelvic Health Physical Therapy, 34 (2010), pp. 10–17.

[8] S. Garg, A. Saxena, and R. Gupta, *Yoga pose classification: a cnn and mediapipe inspired deep learning approach for real-world application*, Journal of Ambient Intelligence and Humanized Computing, (2022), pp. 1–12.

[9] M. Gochoo, T.-H. Tan, S.-C. Huang, T. Batjargal, J.-W. Hsieh, F. S. Alnajjar, and Y.-F. Chen, *Novel iot-based privacy-preserving yoga posture recognition system using low-resolution infrared sensors and deep learning*, IEEE Internet of Things Journal, 6 (2019), pp. 7192–7200.

[10] R. R. Guddeti, G. Dang, M. A. Williams, and V. M. Alla, *Role of yoga in cardiac disease and rehabilitation*, Journal of cardiopulmonary rehabilitation and prevention, 39 (2019), pp. 146–152.

[11] M. U. Islam, H. Mahmud, F. B. Ashraf, I. Hossain, and M. K. Hasan, *Yoga posture recognition by detecting human joint points in real time using microsoft kinect*, in 2017 IEEE Region 10 humanitarian technology conference (R10-HTC), IEEE, 2017, pp. 668–673.

[12] S. Jain, A. Rustagi, S. Saurav, R. Saini, and S. Singh, *Three-dimensional cnn-inspired deep learning architecture for yoga pose recognition in the real-world environment*, Neural Computing and Applications, 33 (2021), pp. 6427–6441.

[13] ———, *Three-dimensional cnn-inspired deep learning architecture for yoga pose recognition in the real-world environment*, Neural Computing and Applications, 33 (2021), pp. 6427–6441.

[14] C. Long, E. Jo, and Y. Nam, *Development of a yoga posture coaching system using an interactive display based on transfer learning*, The Journal of Supercomputing, (2022), pp. 1–16.

[15] P. Pullen and W. Seffens, *Machine learning gesture analysis of yoga for exergame development*, IET Cyber-Physical Systems: Theory & Applications, 3 (2018), pp. 106–110.

[16] G. J. Salem, S. S.-Y. Yu, M.-Y. Wang, S. Samarawickrame, R. Hashish, S. P. Azen, G. A. Greendale, et al., *Physical demand profiles of hatha yoga postures performed by older adults*, Evidence-Based Complementary and Alternative Medicine, 2013 (2013).

[17] G. Sathyanarayanan, A. Vengadavaradan, and B. Bharadwaj, *Role of yoga and mindfulness in severe mental illnesses: A narrative review*, International journal of yoga, 12 (2019), pp. 3–10.

[18] M. B. Schure, J. Christopher, and S. Christopher, *Mind–body medicine and the art of self-care: teaching mindfulness to counseling students through yoga, meditation, and qigong*, Journal of Counseling & Development, 86 (2008), pp. 47–56.

[19] J. K. Sethi, H. Nagendra, and T. S. Ganpat, *Yoga improves attention and self-esteem in underprivileged girl student*, Journal of education and health promotion, 2 (2013).

[20] T.-H. Tan, M. Gochoo, S.-C. Huang, Y.-H. Liu, S.-H. Liu, and Y.-F. Huang, *Multi-resident activity recognition in a smart home using rgb activity image and dcnn*, IEEE Sensors Journal, 18 (2018), pp. 9718–9727.

[21] M. C. Thar, K. Z. N. Winn, and N. Funabiki, *A proposal of yoga pose assessment method using pose detection for self-learning*, in 2019 International conference on advanced information technologies (ICAIT), IEEE, 2019, pp. 137–142.

[22] Z. Wu, J. Zhang, K. Chen, and C. Fu, *Yoga posture recognition and quantitative evaluation with wearable sensors based on two-stage classifier and prior bayesian network*, Sensors, 19 (2019), p. 5129.

[23] K. Yang, K. Youn, K. Lee, and J. Lee, *Controllable data sampling in the space of human poses*, Computer Animation and Virtual Worlds, 26 (2015), pp. 457–467.

Edited by: Sathishkumar V E
Special issue on: Scalability and Sustainability in Distributed Sensor Networks
Received: Aug 31, 2023
Accepted: Nov 4, 2023