Journal of Innovative Image Processing (ISSN: 2582-4252)
www.irojournals.com/iroiip/

PoseNet based Model for Estimation of Karate Poses

Abin Aju[1], Christa Mathew[2], O. S. Gnana Prakasi[3]

Department of Computer Science and Engineering, Christ (Deemed to be University), Bengaluru, India

E-mail: [1]abin.aju@btech.christuniversity.in, [2]christa.mathew@btech.christuniversity.in, [3]gnana.prakasi@btech.christuniversity.in

Abstract

In the domain of computer vision, human pose estimation is becoming increasingly significant. It's one of the most compelling areas of research, and it's gaining a lot of interest due to its usefulness and flexibility in a variety of fields, including healthcare, gaming, augmented reality, virtual trainings and sports. Human pose estimation has opened a door of opportunities. This paper proposes a model for estimation and classification of karate poses which can be used in virtual karate posture correction and trainings. A pretrained model, PoseNet has been used for pose estimation using the results of which the angles between specific joints are calculated and fed into a K-Nearest Neighbors Classifier to classify the poses. The results obtained show that the model achieves an accuracy of 98.75%.

Keywords: Human Pose Estimation, PoseNet, Computer Vision, KNN, Karate, Virtual Trainer

1. Introduction

Karate is a martial art that requires persistent practice to master. To carry out proper and successful training on a regular basis, one must have a thorough understanding of the training. Beginner training is especially important and without proper guidance and practice, perfection would be difficult to achieve. It is possible to train by referring to video tutorials and other materials, but it is challenging for them to assess the correctness of a particular pose. Getting each posture and movement correct is necessary as precision is the key in karate. It is critical to have a skilled trainer and a karate school, also called dojo to develop karate skills. However, once a karateka is familiarized with the fundamentals, it's just as crucial to continue practicing the techniques outside of the dojo. With the advancement in technology especially in the field of artificial intelligence, training without a trainer has

DOI: https://doi.org/10.36548/jiip.2022.1.002
Received: 21.03.2022, received in revised form: 24.04.2022, accepted: 05.05.2022, published: 16.05.2022

become a reality. Today there are models that can detect human poses without the use of specialized cameras or gadgets. These models can detect human poses that are captured through a webcam. PoseNet, BlazePose and OpenPose are few of the existing pretrained human pose estimation models.

Human Pose Estimation is a method of recognizing and classifying the joints of a human body. It allows us to capture a set of coordinates for each joint of the human body, known as keypoints, which can be used to define a person's pose. These points are then used to create a skeleton-like depiction of a human body, which can further be processed for task-specific applications. Unlike Microsoft Kinect, they do not require the users to wear body markers. Instead, anybody with an ordinary webcam and an internet connection can use them. Most of the existing algorithms for human pose estimation are designed for two dimensional images. This, along with variable lighting and camera configurations, is one of the challenges of these models. Human pose estimation can be used in karate as well to build models that can judge if a posture of a karateka is correct or not. In this paper, our aim is to introduce a model that karatekas can easily use to practice karate stances right at the comfort of their homes. PoseNet is used for real-time karate pose estimation in a browser. The user needs to stand in front of their webcam and imitate a karate stance. The webcam captures 17 keypoints of the user's body and classifies the pose, following which the accuracy of the user's pose as compared to the correct pose is displayed.

The remainder of this paper is split into four sections. Section II contains the literature review. Section III elaborates the proposed model and Section IV contains the analysis of the results obtained. The conclusion and future scope are discussed in Section V.

2. Literature Review

Although many works based on human pose estimation have been carried out, there are no published works in the field of karate. Most of them are designed for yoga, healthcare and fitness.

Elmoogy et al. [1] used ResNet50 and ridge regression to create single image indoor localization systems and for outdoor scenes, three fully connected layers were used on top of pretrained ResNet50 features. Down-sampling these features using principal component analysis saved storage space and training time. Chua et al. [2] proposed a model for in-home rehabilitation telehealth service using a PoseNet-based system. The angles and the Euclidean distances between different body parts were calculated to analyse the correctness of a pose. The same approach was used in another work based on the OpenPose model [3]. These models are said to perform well and detect all movements of a subject. In another work for fitness training[4], COCO model and MPII model were used for pose estimation and training was performed using Deep learning and Convolutional Neural Network. The angles between keypoints were calculated for better accuracy. Graphs were plotted on the basis of the angles of different joints and their movements to make comparisons between the ideal athlete's graph and the user's graph. Yamao et al. [5] implemented a model on Raspberry Pi using the PoseNet model for pose estimation in the field of sports. The keypoints of a user are compared with the ideal keypoints registered beforehand to get the accuracy of the user's pose. Experiments conducted to validate the proposed system shows that the model is effective. OpenPose was used in another work [6] in the field of telerehabilitation for detecting body vertices in key frames and in order to judge the correctness of a pose, a novel vertex tracking algorithm was used. A precision camera system was employed, comprising of two cameras precisely positioned in the frontal and lateral directions with the help of a laser cross. They synchronized and aligned the two camera frames in order to independently apply the OpenPose algorithm to each of the two channels and measure movement from two varying body projections. In another work for self-practice yoga [7], the OpenPose model was used to extract the skeleton structure of a pose. The angles between selected keypoints and the difference between the user's vector and the ideal instructor' vector was calculated based on the centre points. The experimental results show that the model was able to detect posture differences between the instructor and the user effectively.

Chaudhari et al. [8] used convolutional neural networks and a human joints localization model for yoga pose identification. It was followed by a process for identifying faults in the pose. This model was found to be 95% accurate. Using OpenPose and FCN, Yan et al. [9] presented an effective method for real-time continuous human rehabilitative action recognition. To monitor human targets and generate 2D poses action sequences from RGB videos, the proposed method first blends OpenPose with a Kalman filter. Then, by sliding the window, the segmented action sequence is extracted from each frame of the human skeleton and the rectangular coordinates are transformed to relative coordinates. To extract spatial-temporal properties and identify activities, a 1D fully convolutional network was created. The experimental results indicate that the model achieved an accuracy rate of 85.6% and can identify continuous rehabilitation actions.

Flores et al. [10] proposed a model for fitness training that skeletonizes the user's video recorded through a smartphone camera and further extracts the angle between specific joints which are then fed to a Fuzzy Inference System to classify workout performance and to identify whether it is likely to cause harm and how to improve. The model gave a training and testing accuracy of 80.42% and 71.67% respectively. Agrawal et al. [11] created a dataset of 5500 images of ten different yoga poses and used TensorFlow pose estimation algorithm to draw the skeletons. Further 12 different angles were calculated and fed into different machine learning algorithms namely Logistic Regression, KNN, Naïve Bayes, SVM, Random Forest and Decision Tree. The results show that the Random Forest Classifier achieved the highest accuracy of 99.04%. In another work for yoga pose classification [12], OpenPose model was used for pose estimation. Pose classification was done using a deep learning model that comprised of time-distributed Convolutional Neural Network, Long Short-Term Memory, and SoftMax regression. This model can classify 6 different yoga poses and achieved an accuracy of 99.91%.

3. Proposed System

The main focus of this paper is to make karate practicing easy for beginners. The objective is to capture the pose of a user through an easily accessible device and judge the correctness of the pose.

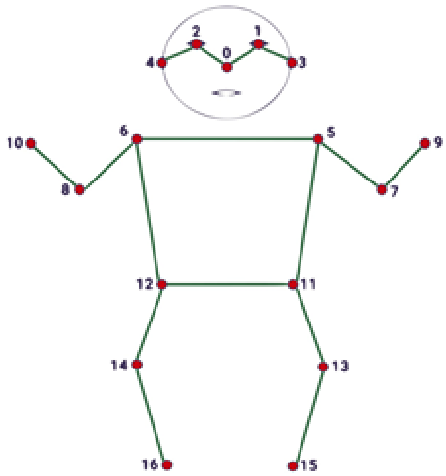


Figure 1. Keypoints of PoseNet model

PoseNet is a deep learning TensorFlow model that can detect various human body parts such as eyes, shoulders, wrist, knees, ankles, elbows, etc. It supports real-time pose estimation in the browser for both single person as well as multiple persons. PoseNet is a two- dimensional pose estimation model which returns a set of coordinates as well as a confidence score for each of the joints of a human body. These coordinates are called keypoints and it forms a skeleton structure of a pose by joining these points, depicted in Fig. 1. A confidence score defines the overall degree of confidence in estimating a pose and keypoints have x and y coordinate values as well as an associated keypoint confidence score. PoseNet can detect 17 keypoints in total for various body parts as shown in Table 1. Because the model was trained using full-size human images, it works well when the user provides full-size human input.

Table 1. Body Joints Detected by POSENET

PoseNet ID	Body Joint
0	Nose
1	Left Eye
2	Right Eye
3	Left Ear
4	Right Ear
5	Left Shoulder
6	Right Shoulder
7	Left Elbow
8	Right Elbow
9	Left Wrist
10	rightWrist
11	Left Hip
12	Right Hip
13	Left Knee
14	Right Knee
15	Left Ankle
16	Right Ankle

In this model, PoseNet is used to detect the keypoints of a human pose, calculated eight different angles between body joints and further used KNN classifier to classify these poses. Transfer learning is used to re-train the PoseNet model and make it effective for karate poses. Transfer learning allows us to use the output of a pretrained model as inputs to another model and is trained to recognize new features or poses, as in our case.

To create the dataset, karate poses demonstrated by authorized karate trainers were captured using an ordinary webcam. The keypoints of each pose returned by the PoseNet model was stored in a JSON file. The dataset contains 100 samples each of four different karate stances namely Heiko Dachi, Zenkutsu Dachi, Shiko Dachi and Neko-Ashi Dachi. The total number of samples in the dataset is 400 out of which 80% data was used for training and the rest 20% was used for testing. The keypoints detected by the PoseNet model are used to compute angles between different joints. Based on these angles, a karate pose will be classified.

In order to classify the poses, K Nearest Neighbors classifier was used. It is one of the most basic and widely used classification algorithms, in which a new data is classified based on the similarity measure of the earlier stored data points and is effective on small datasets. KNN works by calculating the Euclidean distances between a given data point and all the other examples in the data, then selecting k number of examples closest to the that data point and voting for the most recurring label.

The calculated angles were fed as input to the KNN Classifier. Fig. 2 shows the architecture of the proposed model. The model was trained and deployed using ml5 library for JavaScript which is built on top of tensorflow.js.

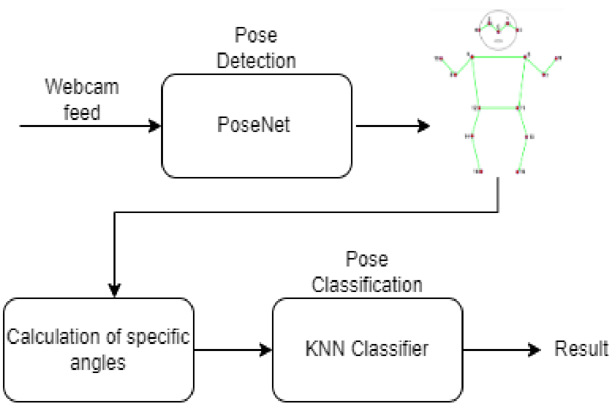
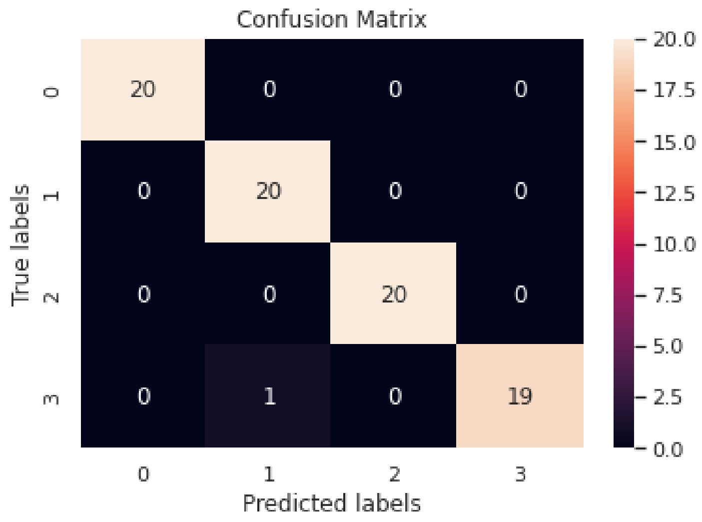


Figure 2. Block Diagram of Proposed Model

4. Results

For KNN classification, setting the value of k as 17 gave the highest accuracy of 98.75% on the testing data. Fig. 3 shows the confusion matrix obtained after classification on the testing dataset. Out of 80 samples, only 1 sample was incorrectly predicted. This indicates that our transfer learning model performs well.



Fgure 3. Confusion Matrix for KNN classification (k=17)

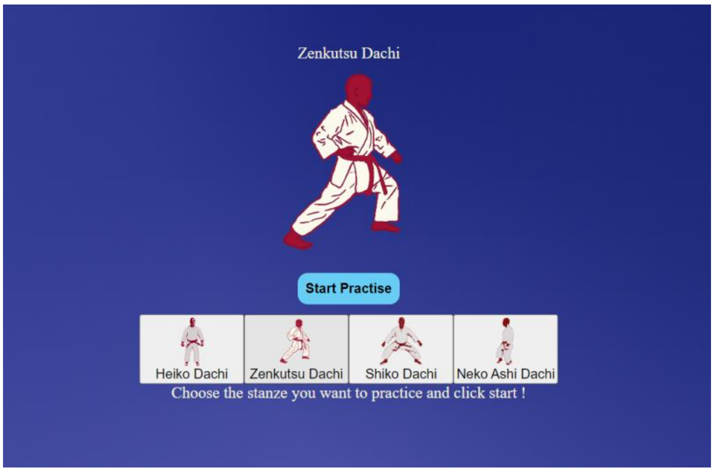


Figure 4. GUI where the correct karate stance is displayed

Figure 5. Trainee's stance where acuuracy is 100%

Figure 6. Trainee's stance where acuuracy is 94%

Figure 7. Incorrect stance of a trainee

A graphical user interface(GUI) was developed to make it easier for the user to interact with the model. The user needs to choose one out of the four karate stances of the trainer to practice (Fig. 4) and once the PoseNet model is loaded, the user will imitate that stance. The keypoints of the user's pose are captured via the webcam using PoseNet model. The angles between specific body parts are calculated and these angles are fed as input to the KNN classifier to classify the pose. The label of the stance selected by the user to practice is compared with the predicted label returned by the classifier. The results are then displayed to the user. The percentage of accuracy is displayed if the pose accuracy is above 90% (Fig. 5 and Fig. 6), otherwise the pose is judged as incorrect (Fig. 7).

5. Conclusion and Future Scope

This paper presents a PoseNet based karate pose estimation and classification model. It is useful for karate practitioners, especially for beginners to learn and practice karate poses outside their dojos. The model will help them to correct their postures which is very significant in karate. The proposed system is able to effectively detect four different karate stances with an accuracy of 98.75%.

In the future, the model can be trained better with more data and will be more accurate if three-dimensional pose estimation models are used. Audio comments can be added so that the user does not have to look at the screen each time to check if the pose imitated is correct or not.

References

[1] Elmoogy, A., Dong, X., Lu, T., Westendorp, R., & Reddy, K. (2020). Linear-PoseNet: A Real-Time Camera Pose Estimation System Using Linear Regression and Principal Component Analysis. In 2020 IEEE 92nd Vehicular Technology Conference (VTC2020-Fall) (pp. 1-6). IEEE.

[2] Chua, J., Ong, L. Y., & Leow, M. C. (2021). Telehealth using PoseNet-based system for in-home rehabilitation. Future Internet, 13(7), 173.

[3] Adolf, J., Dolezal, J., Kutilek, P., Hourova, M., Hejda, J., Milerska, I., & Lhotska, L. (2020). Automatic Telerehabilitation System in a Home Environment Using Computer Vision. Stud. Health Technol. Inform, 273, 142-148.

[4] Dsouza, G., Maurya, D., & Patel, A. (2020, November). Smart gym trainer using Human pose estimation. In 2020 IEEE International Conference for Innovation in Technology (INOCON) (pp. 1-4). IEEE.

[5] Yamao, K., & Kubota, R. (2021, October). Development of Human Pose Recognition System by Using Raspberry Pi and PoseNet Model. In 2021 20th International Symposium on Communications and Information Technologies (ISCIT) (pp. 41-44). IEEE.

[6] Segal, Y., Yona, Y., Danan, O., Birman, R., Hadar, O., Kutilek, P., ... & Adolf, J. (2021, November). Camera Setup and OpenPose Software without GPU for Calibration and Recording in Telerehabilitation Use. In 2021 International Conference on e-Health and Bioengineering (EHB) (pp. 1-4). IEEE.

[7] Lin, C. H., Shen, S. W., Anggraini, I. T., Funabiki, N., & Fan, C. P. (2021, October). An OpenPose-Based Exercise and Performance Learning Assistant Design for Self-Practice Yoga. In 2021 IEEE 10th Global Conference on Consumer Electronics (GCCE) (pp. 456-457). IEEE.

[8] Chaudhari, A., Dalvi, O., Ramade, O., & Ambawade, D. (2021, June). Yog-Guru: Real-Time Yoga Pose Correction System Using Deep Learning Methods. In 2021 International Conference on Communication information and Computing Technology (ICCICT) (pp. 1-6). IEEE.

[9] Yan, H., Hu, B., Chen, G., & Zhengyuan, E. (2020, April). Real-time continuous human rehabilitation action recognition using OpenPose and FCN. In 2020 3rd International Conference on Advanced Electronic Materials, Computers and Software Engineering (AEMCSE) (pp. 239-242). IEEE.

[10] Flores, A., Hall, B., Carter, L., Lanum, M., Narahari, R., & Goodman, G. (2021, November). Verum Fitness: An AI Powered Mobile Fitness Safety and Improvement Application. In 2021 IEEE 33rd International Conference on Tools with Artificial Intelligence (ICTAI) (pp. 980-984). IEEE.

[11] Agrawal, Y., Shah, Y., & Sharma, A. (2020, April). Implementation of machine learning technique for identification of yoga poses. In 2020 IEEE 9th International Conference on Communication Systems and Network Technologies (CSNT) (pp. 40-43). IEEE.

[12] Rishan, F., De Silva, B., Alawathugoda, S., Nijabdeen, S., Rupasinghe, L., & Liyanapathirana, C. (2020, December). Infinity yoga tutor: Yoga posture detection and correction system. In 2020 5th International Conference on Information Technology Research (ICITR) (pp. 1-6). IEEE.