### **Yoga Pose Classifier using BlazePose**

Caleb Kim, Melita Madhurza, Minji Woo

#### 

#### **Abstract**

With the rise of at-home fitness practices, particularly yoga, there is a growing need for accessible, intelligent tools that guide users toward safe and effective outcome. This project presents a system capable of identifying yoga poses from static images by utilizing BlazePose for pose landmark detection and a neural network for classification. The focus is on poses associated with posture improvement and back pain relief. By enabling users to receive posture evaluations without a human instructor, the system paves the way for interactive, real-time feedback tools that support healthier home-based yoga routines.

#### **Introduction**

As home yoga becomes increasingly widespread, many practitioners - especially beginners seek tools to help them maintain proper form without professional supervision. Maintaining accurate posture is key not only to maximizing yoga’s benefits but also to avoiding strain or injury. This project introduces an image-based yoga pose classifier that relies on body landmarks extracted by BlazePose, Google’s open-source pose estimation framework.

The system currently recognizes five yoga poses, selected for their role in improving spinal alignment and reducing back discomfort. It employs a neural network trained on these key points to evaluate the correctness of user poses. Looking ahead, the project aims to evolve into a real-time feedback mechanism that gently corrects user form and supports continued improvement.

#### **Previous Work**

Previous research has explored computer vision-based systems for yoga pose classification and grading, with a particular focus on deep learning models and skeletal keypoint extraction. All of them helped us to move forward with our project. For instance:

* **A Computer Vision-Based Yoga Pose Grading Approach Using Contrastive Skeleton Feature Representations** (Yubin Wu et al.) utilized BlazePose for pose estimation and contrastive learning to enhance pose feature representation. The system aimed to classify and grade yoga poses with superior accuracy.
* **Real-time Pilates Posture Recognition System Using Deep Learning Model** (Hayoung Kim et al.) also implemented a deep learning-based system for posture recognition. This system used BlazePose for pose estimation and classified Pilates poses, optimizing for real-time recognition using a lightweight architecture.

Both of these projects leveraged BlazePose for keypoint extraction, by highlighting its utility for posture classification and real-time feedback systems.

#### **Dataset**

The primary dataset was sourced from the [*Yoga Poses Dataset*](https://www.kaggle.com/datasets/niharika41298/yoga-poses-dataset)from Kaggle. Initially, the dataset was split into training and testing sets: 1081 training set and 470 testing set. The training set had a limited number of images. Therefore, to further expand the dataset and increase variation in our training dataset – such as camera angles and lighting conditions – we used data augmentation techniques. Thus, the final dataset used for our project consists of 6119 training, 277 validation, and 465 test images. We preprocessed the dataset so that each image is represented by a 99-dimensional feature vector using MediaPipe’s BlazePose model, which will be further explained in the next section.

#### **Method**

First, we created a logistic regression model which we used as a baseline model for the yoga pose classification task. This served as a starting point to later compare with our custom deep learning model. The model leveraged the extracted BlazePose landmark features for training. Note that the BlazePose model extracts keypoints (x, y, z coordinates) in the form of 99-dimensional vectors for 33 body landmarks. We chose to use BlazePose for skeletal representations, as our model would then be able to classify poses without relying on the full image, allowing for faster training and better generalization.

Next, we built the deep learning model to improve accuracy and generalization using a neural network. The model’s architecture contained three dense layers with ReLU activations and two dropout regularization, trained on BlazePose landmark features. The input layer was the Blazepose keypoints in the form of 99-dimensional feature vector; hidden layers were two dense layers with 128 and 64 units respectively; output layer was the dense layer with five units (one for each of the five yoga pose classes). We then trained for 30 epochs.

#### **Evaluation**

We evaluated each of our models on its ability to accurately classify the test set images into the five different poses. Our specific evaluation metrics are:

* **Validation Accuracy:** The overall percentage of correctly classified validation images.
* **Test Accuracy**: The overall percentage of correctly classified test images.
* **Precision**: The proportion of true positives out of all predicted positives for each pose.
* **Recall**: The proportion of true positives out of all actual positives for each pose.

#### **Results and Conclusion**

Both our baseline and deep learning model achieved competitive results. The baseline model reached a validation and test accuracy of **85.56%** and **94%**, respectively, while thefinal deep learning model achieved a validation and test accuracy of **94%** and **97%**, respectively. As shown in Figure 1, the deep learning model’s validation loss decreases while its accuracy increases over the course of 30 epochs, indicating strong model performance.

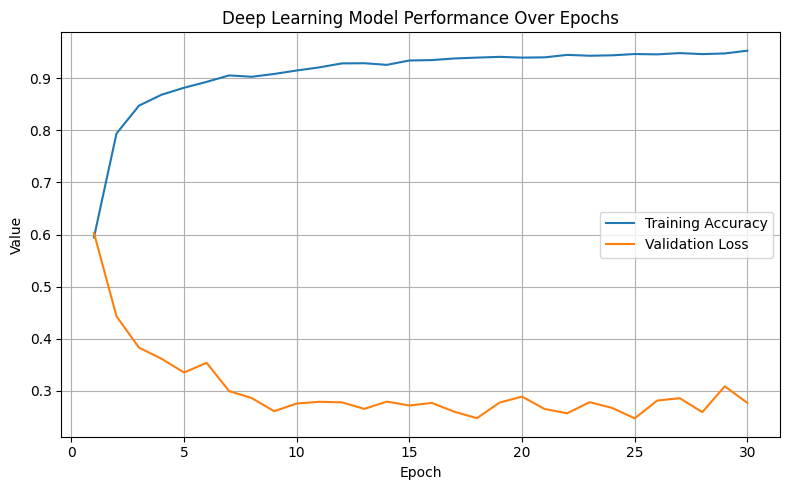
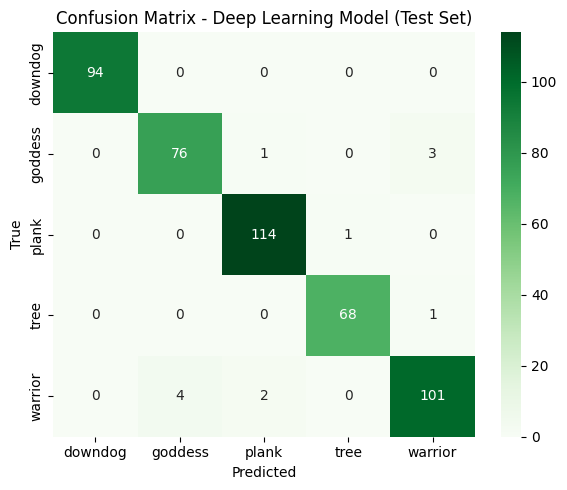
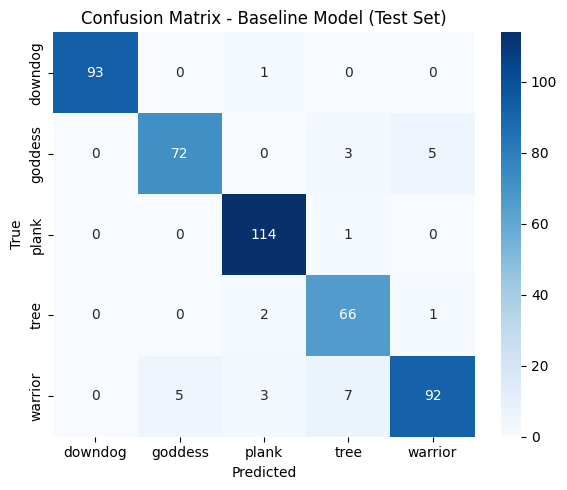


Figure 1

Another pattern we observed was that both models struggled with classifying the “Warrior” pose. Figure 2 and 3, below, presents two confusion matrices that illustrate this, highlighting an infrequent but non-trivial misclassification of the “Warrior” pose with other complex poses such as “Tree” and “Goddess”. However, the deep learning model was able to significantly improve this, decreasing misclassifications across all poses.

  
Figure 2 Figure 3

Thus, our final deep learning model achieved more competitive results when compared with the baseline model. However, the decision to utilize **Blazepose** to transform our dataset proved to be extremely effective for both of the models’ accuracy. By representing each image with 99-dimension feature vectors—extracted via a pre-trained model—we were able to capture spatial data that is both rich and compact.

#### **Discussion**

While our deep learning model showed strong performance, there are areas for improvement. Some of the major limitations and challenges that we faced in this project include:

* **Limited Dataset Size**: Even with our data-augmentation process, our dataset's size was relatively small, especially for certain poses, which led to overfitting in some cases. With more time, expanding beyond the Kaggle dataset and applying more diverse augmentation techniques can help mitigate this issue. Additionally, applying real-world images to our dataset can further avoid overfitting.
* **Limited Training:** Due to time constraints, we were not able to experiment as much as we would like with different model architectures, longer training schedules/epochs, and various learning rates. As a result, the final model may not reflect its full potential performance. To address this, we plan to explore various learning rates and optimization strategies to maximize the accuracy and strength of our deep learning model.

**Summary**

Overall, our long-term goal is to develop a **real-time feedback system** that can assist yoga enthusiasts and beginners alike with their alignment during practice. By utilizing larger datasets, fine-tuning, and an optimized architecture, we are confident in our model’s ability to utilize real-time performance and posture grading.

**Citations**

Wu, Yubin, et al. “A Computer Vision-Based Yoga Pose Grading Approach Using Contrastive Skeleton Feature Representations.” *Healthcare*, vol. 10, no. 1, 25 Dec. 2021, p. 36, https://doi.org/10.3390/healthcare10010036. Accessed 18 May 2022.

Kim, Hayoung, et al. “Real-Time Pilates Posture Recognition System Using Deep Learning Model.” *Lecture Notes in Computer Science*, 1 Jan. 2023, pp. 3–15, https://doi.org/10.1007/978-3-031-43950-6\_1. Accessed 25 Sept. 2024.

‌

‌