Introduction & Motivation

With yoga growing in popularity as a form of exercise, stress relief, and mindfulness, there's a rising need for technology that helps users achieve proper form—especially for those without access to in-person instruction. Incorrect poses can cause discomfort, injury, or reduce effectiveness, making real-time feedback essential.

Recent advances in computer vision and pose estimation, like MediaPipe Pose Landmarker, OpenPose, and DeepLabCut, enable the mapping of human posture and body angles. Building on this, we aim to develop an AI-powered system that recognizes yoga poses, provides real-time alignment suggestions, and recommends logical next poses to create a fluid, engaging experience.

By incorporating angle invariance, the system delivers accurate feedback regardless of camera angle. A recommendation engine will also guide users through smoother, more dynamic sequences, avoiding repetitive routines. The result: a personalized, safer, and more effective yoga practice for all skill levels.

Experimental Design

Classifier

The first step in completing our pipeline was creating a model to classify yoga poses by their given name. When a subject is doing a pose, the AI model should be able to identify the specific pose and where the subject doing this pose is located. To accomplish this, we will implement a pose estimation model in Python and train a classification model. While we have already experimented with Google's MediaPipe Pose Landmarker, the exact model we will use may change depending on our needs. The Pose Landmarker works well for mapping skeletal points on an image, but often has issues when translating this to a 3-Dimensional map. Should we need 3-D representations, we could switch to a model that better handles this task such as OpenPose, DeepLabCut, or some other model.

Once the pose estimator is up and running, we will run it on images from the Yoga-82 dataset to transform image data into skeletal landmarks. This will include the location of wrists, elbows, shoulders, heads, hip joints, knees, ankles, and more. Reducing image data into numerical features will allow for more efficient and generalized learning. This new dataset will be used to train a Neural Network that will perform the pose prediction. An important note about the dataset and the training: The dataset includes pictures of the same yoga pose from different angles, primarily from head on or perpendicular to the camera. We can then utilize the idea of angle invariance, which states that angles and ratios of distances are invariant under scalings, rotations, translations and reflections. Thus, the idea is that a pose at 45° to the camera will be similar enough to the same pose at either 0° or 90° such that the model can still correctly classify it. Looking ahead, angles being invariant under rotation will also allow us to still reliably calculate necessary angles of skeletal features across a wide range of subject to camera angles.

Live Identification & Next Pose Recommendations

Once we have a trained model on yoga images that is capable of classifying poses, we will extend the model to work with a live video feed such that it can identify poses in real time. Ideally, we will also support multiple subjects within the frame. We do not expect this to be too difficult, and will have the classifier operate with some frequency such that it is both responsive but not overactive. Included in our AI response to a pose will be a 'Next Step' pose. This will be done through comparing relevant angles between poses, such as the angles that elbows, knees, torsos, and hips are bent at. Using some graph structure, the AI can then recommend a next pose that is not too distant from the pose it is currently detecting. To prevent repeated and highly deterministic pose-flows, our graph will most likely include some weights and stochastic nature such that multiple poses may be recommended.

Subject Specific Suggestions

Lastly, our project will culminate in the most ambitious task of having the AI give out subject specific pointers to improve a pose and get closer to an ideal version of it. This presents many challenges, as we must have a calculated ideal pose and be able to properly track angles and lines through rotation. This will allow us to calculate relative angles of the subject and have the AI give out suggestions such as "Straighten your right arm," or "Widen your stance." Due to the nature of angles and ratios being invariant under rotation, regardless of a subject's angle to the camera, we should be able to compare the learned skeletal form's angles and ratios against the subjects and decide which pointers to apply simply based on case statements.