Ypose: Towards Efficient Yoga Pose Recognition in Single Images

***Abstract*—Pose recognition in computer vision attempts to estimate the position and orientation of human postures. It involves locating human body joints and subsequently using them for the pose prediction. The task of locating human body joints becomes challenging especially when complex postures are performed. This paper proposes the pose recognition for yoga asanas that consist of many complex postures involving body rotation and folding the body into multiple manners. It imposes various challenges on the computer vision algorithms like occlusion, inter-class similarity, intra-class variability, viewpoint complexity, *etc*. This paper presents Ypose, a yoga pose recognition network to recognize yoga asanas from single RGB images. Ypose consists of four steps. First, the region of interest (ROI) is extracted using the segmentation and bounding box tracking approach. Then, an EfficientNets backbone architecture having a convolutional neural network (CNN) is applied for the feature extraction. Third, dense blocks of the Densely connected convolutional networks are added. These blocks consist of inherent skip connections that learn to produce relevant features. Fourth, fully connected layers are applied for the yoga poses classification in a multi-level hierarchy. Ypose has been tested on the Yoga-82 dataset, which is a publicly available benchmark dataset for yoga pose recognition. Experimental results show that the proposed model achieves the state-of-the-art on this dataset. Ypose has obtained an accuracy of 91.74% that is a 12% improvement from earlier state-of-the-art (79.35%).**

***Index Terms*—Pose recognition, Yoga, Image classification, Segmentation and classification, Deep CNN architectures.**