**Related Work**

LeCun et. al. proposed a convolutional neural network (CNN) method for deep learning and successfully applied it to OCR \cite{zhang1998} and generic object recognition tasks. Recently, with greater availability of large labelled datasets and computational power, CNNs have proven to be one of the most accurate methods for generic object classification \cite{krizhevsky2012}. However, CNNs perform poorly on smaller datasets. This issue has been addressed by using unsupervised learning methods on large amounts of unlabelled data \cite{ranzato2007}. Zhang et. al. \cite{PANDA}, use a different approach, called PANDA, which pre-processes the data using part-based methods, which subdivide the image into important subregions, and then single CNN is trained on each of them. Specifically, it uses poselet based part detector, which is trained on positive examples using key-point annotations. \cite{panda} use PANDA for attribute classification, which is a binary classification problem. Pose-estimation is another problem with multiple classes, which has been addressed using DPM (deformable parts model) \cite{paper}. We, in this project, instead apply PANDA based approach to this problem and study its performance.

**Approach**

Pose detection will be performed using a part-based model called poselets. Poselets will be used as part localization scheme for fine-grained categorization of poses. An image is divided into multiple poselets based on every specific pose/action. Features are then extracted from every single poselet using Convolutional Neural Networks (CNNs). It is hard to find the kind of features obtained through CNNs, hence we will try to visualize the extracted features to know the nature of most important features (HOGs, Intensity etc.). Once discriminative image representation for each poselet is obtained then features from all the poselets will be normalized and concatenated to obtain a single representation of the image. By this time images and their convoluted features will be present, which later will be used for training different classifier. Since, CNNs consider the non-linearity within the features therefore a non-linear classifier is needed to do this job. SVM, Deep Belief Networks and k-NN will be used for classification of images for different poses.