**And Introduction for VGG-like neural networks:**

VGG, stands for Visual Geometry Group, a group of scientists and researchers in Oxford University, which can be considered as one the most influential group of scientist in the field of Computer vision. Their proposed network is consist of 16 layers, the basics structure of their neural network is as follows:

* Convolutions layers (used only 3\*3 size )
* Max pooling layers (used only 2\*2 size)
* Fully connected layers at end

The more description of the networks and many other useful images can be found at their website:

<http://www.robots.ox.ac.uk/~vgg/>

**Long Description for VGG-MC\_SL: (Multi Class Single Label)**

The Current version of the classifier contains 7 classes. The Images are downloaded by sending some queries to Google image search engine and store them locally. Each query list is a set of synonyms or noun plus adjective for each member of label list. The label of a photo is located in the first word of a query list.

Some pre-processing tricks as like as rotating and distorting image, translating it and etc. are implemented to increase the accuracy of the classifier and also to increase the number of training set.

The classifier accepts image storage path of any size but just in JPG (JPEG) format. It resizes the input image, applies the model onto it, and outputs the top-1 prediction label and the corresponding probability \_\_The extension of it to top-k is also possible.

The Classes are: Screwdriver, drill, hammer, Pliers, Doctor, Nurse, and Dentist.

The Top-1 Accuracy of the classifier is approximately 80.0%.

**Long Description for VGG-MC\_ML: (Multi Class Multi Labels)**

The Current version of the classifier contains 8 classes. The Images are downloaded by sending some queries to Google image search engine and store them locally. Each query list is a set of synonyms or nouns plus adjectives for each member of label list. The label list is a set of labels per each photo which is located in the first word of a query list.

Some pre-processing tricks as like as rotating and distorting image, translating it and etc. are implemented to increase the accuracy of the classifier and also to increase the number of training set.

The classifier accepts image storage path of any size but just in JPG (JPEG) format. It resizes the input image, applies the model onto it, and outputs the top-1 prediction label and the corresponding probability \_\_The extension of it to top-k is also possible.

The Classes are: black-dress, black-scarf, black-watch, red-dress, red-watch, white-dress, and white-watch.

The Top-1 Accuracy of the classifier is approximately 90.0%.

**Long Description for INCEPTION-V3-MC\_SL:**

Inception-v3 is trained for the [ImageNet](ImageNet:%20http://www.image-net.org/) Large Visual Recognition Challenge using the data from 2012. This is a standard task in computer vision, where models try to classify entire images into 1000 classes, like "Zebra", "Dalmatian", and "Dishwasher". Basically, the inception models are based on combinations of max average pooling and convolutional filters in different size.

The more details can be found here:

<https://arxiv.org/abs/1512.00567>

The classifier accepts image storage path of any size but just in JPG (JPEG) format. It resizes the input image, applies the model onto it, and outputs the top-5 prediction label and the corresponding probability.

According to Tensor Flow’s reports, Inception-V3 outperforms the other algorithms with the average error rate of 3.46% whereas Andrej Karpathy who attempted to measure his own performance on this dataset reached 5.1% top-5 error rate. The more details can be found on the links below:

<https://www.tensorflow.org/tutorials/images/image_recognition>