

Deep Learning and Support Vector Machine Project

In this project you are asked to create your own 2 deep learning models on the “Mushrooms” dataset from Kaggle. The dataset has 9 classes. We will assume that classes are described in alphabetical order:

```
classes = {Agaricus, Amanita, Boletus, Cortinarius,Entoloma,Hygrocybe,  
Lactarius, Russula, Suillus}.
```

First model will be used for classification using softmax as output layer. Second model is used for feature extraction and the result of those features will be plug into SVM for classification. Both models cannot have more than 10 layers of convolution and dense layer. It is ok to have less than 10 layers.

You may find the following links to be useful:

- <https://www.tensorflow.org/tutorials/images/classification>
- https://www.tensorflow.org/tutorials/images/data_augmentation
- https://www.tensorflow.org/tfx/guide/tft_bestpractices
- <https://scikit-learn.org/stable/modules/svm.html>
- https://www.tensorflow.org/addons/tutorials/losses_triplet

In designing your program you may use all of the functionality of OpenCV, and deep learning under the following conditions:

- You must use tensorflow.
- You can use Scikit-Learn and other visualization related libraries.
- You must **not use pytorch** for this project.
- You must **not use existing backbone models**.
- You cannot use any special libraries or software packages. Please contact the TA if you are not sure if something should be considered “special”.

Provided files

- **mushrooms.zip** file for training data.
- The python script **proj3_test.py** is provided as basic for creating 2 testing programs. Test programs read in test folder and csv file for their labels.
- Sample test folder as **mushrooms_test.zip** and the corresponding csv file **mushrooms_test.csv** files are provided for testing.

Evaluation

- We will use test accuracy as the part of the measure for evaluation.
- We will create a test folder containing test images and their corresponding labels. The images **will not** be the same as those in the provided **mushrooms** folder, but similarly collected from the web.

Here is the detail breakdown of the scoring.

Reading input folders, pre-processing	10 points
data augmentation of training data	10 points
Creating classification model with softmax	10 points
Creating feature extraction model and SVM for classification	10 points
Save and submit 2 trained models and weight files	15 points
Edited evaluation code with data pre-processing	10 points
Report	15 points
Accuracy evaluation for classification model	10 points
Accuracy evaluation for feature extraction model with SVM classifier	10 points

What you need to submit

1. Python or ipython notebook source code of your program. Please name it **proj3.py**.
2. Documentation that explains your approach.
3. Your 2 models, weight files, your testing code and instruction on how to set up environment for testing your code. Please put in any pre-processing and model setup you do in your test code. Please name classification model test code as **proj3_classification_test.py** and your feature extraction model with SVM as **proj3_extractSVM_test.py**
4. Please provide your environment setup which you used to save your model. You can create and upload **requirements.txt** file of your environment. You must provide at least your python and tensorflow version.

These items should be put in a zip file named with your netid and submitted on elearning. For example, if your netid is xyz1234 you should submit a zip file named xyz1234.zip.