CSC 84020 Neural Networks & Deep Learning

Hw3: Color Image Classification with Convolutional Neural Networks (Keras and/or Tensorflow and/or PyTorch)

Problem 1: Explore the implementation of CNN with a **color image dataset** of your choice. Datasets as **MNIST, CIFAR-10, CIFAR-100 or MNIST-like and CIFAR-like are excluded.** As the datasets are usually bigger than you can afford, you may decide use only part of it.

All teams explore different color image datasets. Send me an e-mail to have you dataset approved.

a) Download the original dataset (cut it in advance if necessary). You need not less than 5 - 6 classes with not less than 1000 images for each of the classes in the training set, not less than 300 images for each of the classes in the validation set, and not less than 60 images for each of the classes in the test set. These numbers are approximate and you should follow the recommended distribution ratio.

- Split the dataset into **three subsets**: training, validation and a test set following the ratio (70% - 20% - 10%).

- Use dataset analysis and preprocessing as normalization or standardization etc.

- Represent the classes in training and validation sets in the following two ways **(graphics and text – Fig. 1)** shown below for training set. Do the same for validation set.

- Make conclusions about the imbalance of the training and validation sets and pick-up the way to fix the imbalance problem of both datasets: training and validation. **(Help: Lec8 CNN Part2, SMOTE.ipynb)**

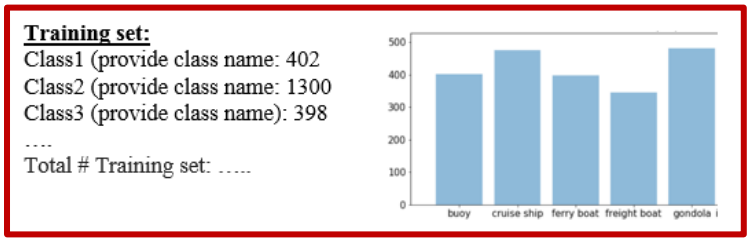


Fig. 1

For testing set provide the same information (graphics and text – Fig. 1) without considering the issue of imbalance.

b) Use Sequential model and build the topology of your **CNN** implementing: **imbalanced dataset as well the balanced training and validation datasets.**

- Your CNN topology should include not less than four convolutional layers, where the number of neurons, the number of filters and their size and the type of pooling layers and where to include the last one is up to you.

- Do experiments with 3 different models (the difference could be in the choice of activation function, optimizer, and number of convolutional layers, batch size or any combination of these, **learning rate change**, stride etc.). Include as a first choice ReLU, Adam and batch size 32. Do not forget batch normalization and shuffling. Describe each model. Apply regularization and dropout to the model to prevent overfitting if necessary.

- The above experiments with the specified 3 models must be performed for the imbalanced and balanced training and validation datasets.

c) Plot training and validation accuracy as well as training and validation loss for each of the above models using combinations like the one shown below. Make appropriate notations having details concerning model, optimizer and activation function

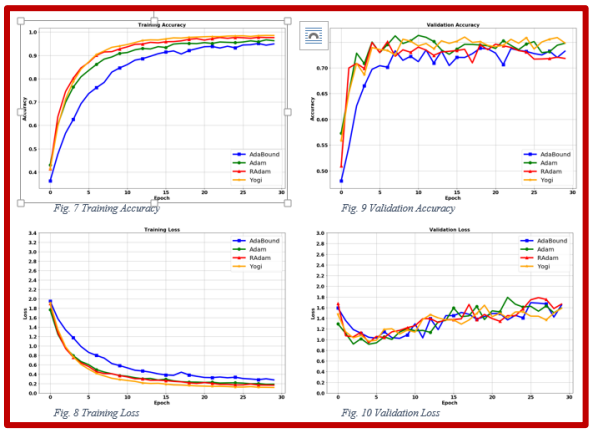


Fig. 7 Training Accuracy *Fig. 9 Validation Accuracy*

Fig. 8 Training Loss *Fig. 10 Validation Loss*

d) Display some channels in every intermediate activations and explain why this is useful. **(Check Help)**

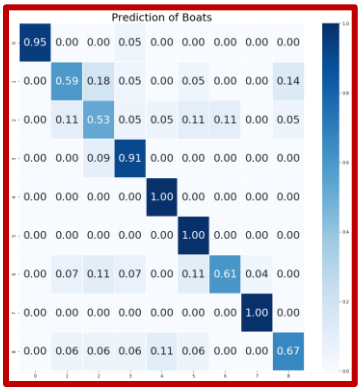
e) Visualize (display) convolutional filters: get the gradient of the loss with regard to the input, include a code for filter visualizations and generate a grid of some filter response patterns in a layer. **(Check Help)**

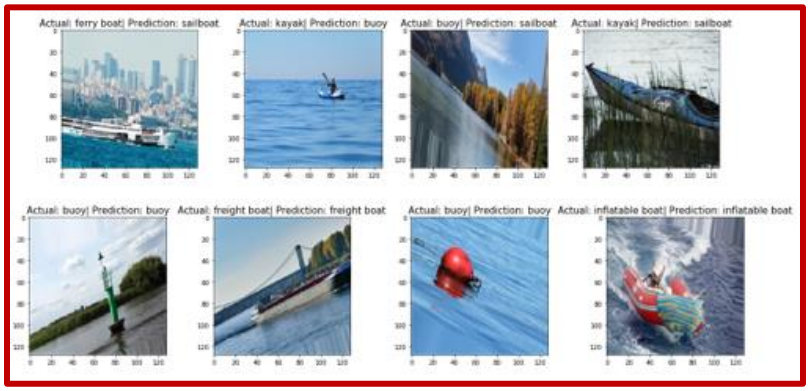
Grading for full completion of Problem 1: for each part a, b, c, d, e – max 20 points
(Max total for Problem 1 = 100 points)

Submission 1:

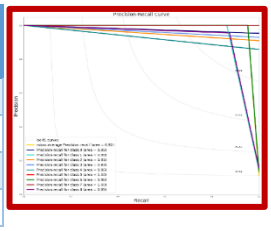
- **All source files in ipynb and pdf with respective graphs and plottings which contain the steps a, b, c, d, e of Problem 1 described above.**
- **Report 1 which includes the description, analysis and conclusions of Problem 1 a) and b) in doc or pdf format (not less than 1 full page with text, figures and tables come additionally, Times New Roman 11, single space).**

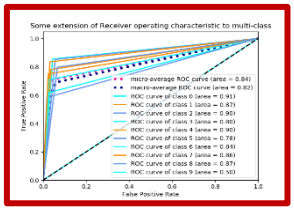
Problem 2. Perform testing. Get the predicted class – show some correctly and not correctly predicted images of testing set as below. **Evaluate test results using confusion matrix, classification report (precision, recall, f1-score, support, ROC/AUC and Precision-Recall curves). Except tables for this evaluation provide graphics of normalized confusion matrix as shown below, ROC/AUC curves as the one below. (HELP: Adabound_Decay.html)**





Model	Precision	Recall	F-1 Score	AUC
AdaBound	0.7396	0.7136	0.7200	0.8427
Adam	0.7361	0.7107	0.7158	0.8412
RAdam	0.7470	0.7083	0.7156	0.8402
Yogi	0.7587	0.7420	0.7450	0.8585





Implement any ENSEMBLING you want and present the results in a table as shown below as well as confusion matrix of your ensemble:

Ensemble Model	Precision	Recall	F-1 Score	AUC
Adam, AdaGrad, and Adabound	0.8600	0.8800	0.8600	0.9294
Adam, AdaDelta, and Adabound	0.9000	0.8800	0.8800	0.9329

Grading for full completion of Problem 2: (Max total for Problem 2 = 60 points)

Problem 3: Report 2. Result Analysis of Problem 1 and Problem 2 (not less than two full pages with text (pictures and tables are not counted into the two full pages), single space, Times New Roman 11) (max 40 points).

Submit on the Blackboard:

a) **Upload your dataset as well as the portion you use, all the details like description of the full dataset and link to it (you can use csv or zip files for the complete datasets and doc or pdf file for dataset description).**

b) **Upload your Python Hw3 Problem1 and Problem 2 file in py, pdf and ipynb format (Solution Problem 1&2) which combine your program code and all your outputs. Make sure to include comments. Make sure that figures appear in reasonable and readable size.**

c) **Upload your REPORT 1 which includes the description, analysis and conclusions of Problem 1 a) and b) in doc or pdf format (not less than 1 - 2 full page text (pictures and tables are not counted) , Times New Roman 11, single space).**

d) **Upload your REPORT 2: Result Analysis of Problem 1 and Problem 2 (two - three full pages text (pictures and tables are not counted), single space, Times New Roman 11).**

e) **List all members of your team and make sure that the submission is done before the expiration of due date and time.**

The max number of points for Hw3 is 200 points.

Your e-mail submissions will be ignored.