Binary Classification of Yelp Restaurant Images

Team-members:
1) Burman, Nishith (burman.ni@husky.neu.edu)
2) Akbar, Md Navid (akbar.m@husky.neu.edu)

Objective: Imagine a time when you wanted to view how that steak looks after cooking, in a restaurant you are considering visiting. For reasons such as this, it is quite essential that an option exists in the reviews of restaurants page (Yelp, in our case) that groups images by such features. The primary objective of this project is to distinguish between two classes of images: food and drink so that users can view them by category.

Approach: We want to build a binary image classifier using a convolutional neural network (CNN), with a depth of about three (convolution, activation, and pooling combination) layers. First, we will resize all the images to a common resolution (e.g. 150 by 150 pixels). Then, we will train the network using the training set (e.g. 6000 images), which maintains the ratio of both the classes, food, and drink, as in the original dataset. Training will be done using Keras in Python. Finally, we will test it on the out-of-sample images (e.g. remaining 1000 images). This same procedure will be repeated for *k*-fold cross-validation. In addition, we want to visualize to intermediate layers in Python, to deliver an explainable CNN model. If time permits, we want to apply an intelligent image cropping technique [3rd link above] reported recently and observe the corresponding change in performance accuracy.

Dataset to be used: The dataset made available by Yelp on Kaggle contains about 200,000 images. They belong to either of the five classes: indoor, outdoor, food, drink, and others. For our project, we plan on using about 7000 images in total, from the two categories: food and drink.

Progress-timeline: We plan to mine, clean and wrangle the image dataset based on our model requirement by the next week. Then we will be working on the architecture of our model this would take another week or so. Finally, we plan to train our model using the selected data and once the model is trained, we plan to test the accuracy of our model by the end of this month.

Deliverables:
- A complete report, using Tensorboard diagrams, on the train-test performance using the baseline CNN model.
- Intermediate visualization of the layers and possible interpretation.
- A performance analysis using the new image cropping technique.

Splitting of Work [Tentative]:

Nishith Burman
- Choosing optimal tuning parameters for the binary classifier
- Designing of Tensorboard and incorporating in the model
- Preparation of report

Md Navid Akbar
- Separating images from Yelp dataset into two categories: food and drink
- Designing CNN in Keras with TensorFlow as backend
- Preparation of report

References:

1. Yunpeng Li, David Crandall, and Daniel Huttenlocher "*Landmark Classification in Largescale Image Collections. 2009, IEEE. 1957-1964*".
2. Li, F. (2017, January 5). *Convolutional Neural Networks (CNNs / ConvNets).* Retrieved November 21, 2017, from http://cs231n.github.io/convolutionalnetworks/
3. Andrew Shaw, Yaroslav Bulatov, Jeremy Howard "*progressive resizing for classification, and rectangular image validation; NVIDIA's NCCL with PyTorch's all-reduce; Tencent's weight decay tuning*"
4. Rifkin, R. (2008, February 25*). Multiclass Classification*. Retrieved November 20, 2017, from http://www.mit.edu/ 9.520/spring09/Classes/multiclass.pdf
5. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., ... & Rabinovich, A. (2015). *Going deeper with convolutions. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1-9*).
6. He, K., Zhang, X., Ren, S., & Sun, J. (2016). *Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778)*.