**18794 Project Feasibility Report:**

**Image Colorization via Ensemble of Convolutional Neural Networks**

Sohil Shah (sohils@andrew.cmu.edu)

Rohit Banerjee (rohitban@andrew.cmu.edu)

Brandon Perez (bmperez@andrew.cmu.edu)

**Background and Project Overview:**

Image colorization seeks to generate a full color image from an input grayscale scene. The approaches to this problem can be divided into two broad categories: automatic and non-automatic methods.

Scribble image colorization and Matrix completion rank among the most well studies non-automatic methods [1, 2, 4]. Scribble colorization requires the user to annotate the grayscale image with colored lines (“scribbles”). These colors are then propagated in space based on the principle that areas of the image with similar intensities have similar colors [2]. On a similar vein, matrix completion requires both the grayscale image and a list of pixels with color labels.

A shortcoming of non-automatic methods is that different user input is required for each distinct image. Automatic colorization methods overcome this limitation by first training a CNN on an image dataset and then using the trained network to map luminosity values to color values in the image. However, automatic methods tend to be biased towards desaturated background colors [4].

We propose a novel method for image colorization by learning an ensemble of identical CNNs with using AdaBoost to effectively reduce bias of the reconstruction while keeping the method automatic. We will use augmented data to increase the diversity of the estimators generated by the training. We hope to have each CNN performing better on a different subset of tasks, such as one for landscapes and one for people. This will allow us to overcome issues faced in previous research.

**Data:**

The network is tested and trained on a subset of the ImageNet database [3]. The images are rescaled to 256x256 for consistency and first converted to grayscale. We perform data augmentation by creating duplicates of the images at higher and lower contrast, different amounts of blur, and other label preserving transformations. This increases diversity of the hypotheses learned by the ensemble. A potential bottleneck may be how many images we can train on, as the previous work uses powerful GPU based clusters to train on millions of images [5, 6]. The test images will be taken from the same synsets of the ImageNet database as was trained on to keep consistency and prevent us from needing to train on too large a number of images.

**Method:**

We use an ensemble of *n* identical convolutional neural networks, following similar design to that seen in [5, 6]. In this case, *n* will be determined experimentally in order to get a good tradeoff between bias and variance. The augmented data is placed in a sample pool of size *m*. We then bootstrap *n* sets of *m* training examples randomly with replacement from the sample pool. This method diversifies the concepts learned by each of the CNNs, allowing them each to converge to different solutions. Then, the AdaBoost method is used to train the ensemble and reweight the CNN coefficients. The final reconstruction will consist of a linear combination of the *n* CNNs.

We will evaluate the prediction by calculating the mean square error of the reconstructed output and the original image. Training finishes when MSE is less than a threshold . The bias of the individual CNNs is expected to be high. With a diverse enough set of generated hypotheses, boosting provably reduces both bias and variance.

**Timeline:**

Tasks Completed So Far:

1. We have read over a significant number of research papers to determine the various ways to tackle colorization, and determine a novel approach to the problem.

Tasks To Be Completed:

1. Generate our training data set by augmenting a subset of ImageNet, to be completed by the mid-progress report (before Thanksgiving break, 11/23).
2. Implement and train a CNN for colorization on the data set, to be completed by the mid-progress report (before Thanksgiving break, 11/23).
3. Implement the AdaBoost algorithm for dynamically reweighting the CNN coefficients, to be completed by the mid-progress report (before Thanksgiving break, 11/23).
4. Tune the CNN to achieve desirable results, to be done the last two weeks of class after Thanksgiving break.

**References:**

**[1]** Charpiat G, Bezrukov I, Hofmann M, Altun Y, Schölkopf B. “Machine learning methods for automatic image colorization”. *Lukac R (ed) Computational photography: methods and applications*, 2010.

**[2]** A. Levin, D. Lischinski, and Y. Weiss**. “**Colorization using Optimization”. *ACM Transactions on Graphics,* 2004*.*

**[3]** O. Russakovsky, J. Deng, H. Su, et al. “ImageNet Large Scale Visual Recognition Challenge”. *IJCV,* 2015.

**[4]** G. Charpiat, M. Hofmann, B. Schölkopf. “Automatic Image Colorization Via Multimodal Predictions”. *Proceedings of the 10th European Conference on Computer Vision: Part III*, 2008.

**[5]** B. Hariharan, P. Arbelaez, R. Girshick, and J. Malik. “Hypercolumns for object segmentation and fine-grained localization”. *Computer Vision and Pattern Recognition*, 2015.

**[6]** Z. Cheng, Q. Yang, B. Sheng. “Deep colorization”. *Proceedings of the IEEE International Conference on Computer Vision*, 2015.