Project Proposal

CPSC 599, Deep Learning For Vision

Deliar Mohammadi, 30072994
John Zheng,30125258
Xinzhou Li, 30066080

1. **The goal of the project (10%) - Clearly state what problem it aims to solve**

The goal of this project is to be able to convert images that are black & white to color images. This process is known as “colorization”. When this project is complete, we will be able to take any black & white image and successfully transform that image into a fairly accurate colored version using a convolutional neural network.

2. **Importance of the problem/project (10%) - Explain why this problem or project is important and relevant.**

This project is important because it has numerous applications. People have old images of family and friends that they would like to see in color, and there are many historical images and videos that are also in black & white, given a fully trained model it would be possible to bring these historic memories back to life.

3. **Data sources (10%) - Outline the data sources that will be used for model building. Consider using an existing image/video dataset, or justify the decision to develop your own dataset, ensuring that it can be generated by the end of February and that you have the necessary capacity to collect the data, including personnel, hardware, and permission. Also, address any potential ethical issues associated with data collection and use**

We will be using data from several sources, one of them being this image colorization dataset: https://www.kaggle.com/datasets/shravankumar9892/image-colorization
We will look towards adding additional data points as we see fit, these data points may come from our own images or other online images.

4. **High-level approach (20%) - Provide a high-level overview of how you plan to solve the problem, for example, "We build an image classification model to distinguish between X and Y, using a 2D CNN model ...."**

The high level approach for this problem is as follows:

a. Build a CNN, specifically following a U-Net architecture
b. Feed the model black&white images, and use the colorized version of those images as the correct output, so loss will be computed against the real colors.
c. Do step c for a large number of images of all kinds such as animals, humans, landscapes and other objects.

With these steps, we should be able to achieve the project goals. Keeping in mind that we may have many iterations in the model, and we may update hyperparameters such as our loss function or learning rate.

5. **Sample size and manual labeling (10%) - Justify that you have enough samples in the dataset and whether any manual labeling is required. Provide evidence such as similar tasks performed with similar data set sizes or initial results/experiments as good justification**

The image dataset that we referred to in question 3 has over 25,000 images. This should be enough data to properly train the model. In the case that this is not enough, there are several other open online image datasets that are specifically built for this task, and we can sample from those at our own discretion. The paper cited in question 6 was in the same area as our project, and they only used 5000 images from the ImageNet dataset, and received very good results. No manual labeling of images are required as we are not detecting objects, we are simply segmenting the image into colors. We may have to manually convert images to grayscale using opencv or other python libraries, but this task is very fast and will not take up a considerable amount of time.

6. **Peer-reviewed article (30%) - Cite at least one peer-reviewed article related to your project and explain how it is relevant to your work. Describe the approach used in the paper, but note that you do not need to follow it completely.**
   https://www.mdpi.com/2073-8994/14/11/2295

The article linked above is very similar to what we are trying to achieve, they used a modified U-Net architecture to successfully colorize images of humans. The dataset they used is 500 images, taken from the ImageNet dataset. One thing to note is that this paper tripled the up-sampling and down-sampling of the original U-net architecture, and they replaced RELU with sigmoid activations. We will be doing a very similar process to the group in the referenced paper, however we will have a larger dataset and attempt to color all objects, not just humans.

Paper Link

This paper above introduced a new architecture of the colorization, The model mostly used the CNN network, and it will consider both local information and global priors. Moreover, the model trained by a large scale of dataset, at the same time the model was added a classification function, based on that function, it can have a better understanding of global priors from the input data, which will lead the model to choose a more accurate and appropriate color.

Paper link

This paper also provides a way to colorize a grayscale picture, they used user-guide mode which provides some color choice for the user for a particular object in the picture, they also implemented a dramatic, high-level modification of the visual appearance. The main progress of the model is the local hint network, it has the functionality that simulated user action, and based on that given user several colors for colorization. By introducing the local hint network, the output images have a great improvement and they look more realistic.

7. **Timeline (10%) - Include a timeline for different steps of the project, including data collection, model training, model evaluation, and deliverables**

Feb 20	Submit Project Proposal
Feb 27	Begin Data collection and organization process
Mar 1	Begin building the U-Net architecture in pytorch. Following the original model. Later adjustments can be made, or we can create multiple models of different sizes.
Mar 6 to Mar 15 (Deliverable 1)	Begin training our model(s) on the training data, using several different hyper parameters and models made previously. This process can take several weeks due to the size of our training data.
Mar 15 to Mar 27	Finalize the model, and select the model with the best overall performance. Prepare a presentation that covers all of our work so far, including data collection, building the model, why we chose this model, sample input/output images and discussion on hyper parameters.
Apr 1	Project Presentation.
Apr 7	Submit Project Report. Finalize the report that we have been working on all term and submit to D2L dropbox.