Compression and File Sharing of Facial Images via Generative Networks

1 Background

Applications such as Snapchat and Facebook provide streamlined platforms for sending facial images between users. However, due to the high volume of images transferred between users on these platforms, images sent through these applications are either downsampled or captured at a lower resolution than the resolution of a user's camera. This facilitates lower transfer times for the sender, but results in inferior-quality photos for the recipient.

2 Aim

To address the limitations presented above, the use of a Laplacian Pyramid of Generative Adversarial Networks (LAPGAN) is proposed to quickly upsample a downsampled or natively low-resolution facial image for a receiving user. Allowing images to be transferred at an extremely low resolution will reduce the cost of data transfer and improve latency for users.

3 Methodology

The methods that will be utilized to complete this project are outlined in this section.

3.1 Components

The components for this project, inluding libraries, hardware, and the LAPGAN network are listed below. All code will be implemented in Python.

1. A LAPGAN coded in the Pytorch deep learning library
2. Two computers connected to the same wireless network
3. The sockets module to facilitate communication between the two computers
4. The PyQt library to create a user interface for the application

3.2 Methods

The first step of this project is to create a working LAPGAN in Pytorch. The discriminators in the LAPGAN will use the Scaled Exponential Linear Unit (SELU) activation function as these are reported to increase the stability of GANs. The training data for the network will be the CelebFaces Attributes Dataset (CelebA). The images will be split into training and testing sets, downsampled, centered, and fed into the network. The output will then be compared to the ground truth images. The image upscaling process is illustrated in Figure 1.

After the code for the LAPGAN is complete, a separate script will be created to center input images on the face that will be fed into the network. This will mitigate the problems caused

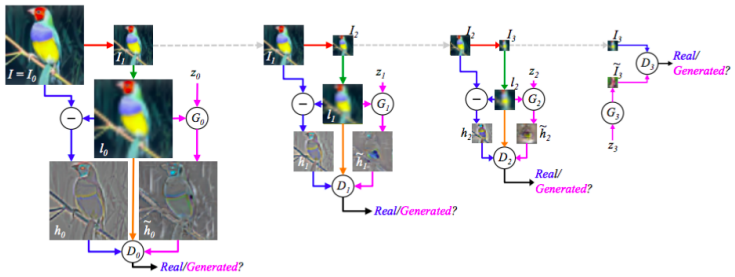

Figure 1: The image upscaling process carried out by the LAPGAN. An input is fed into the rightmost GAN. The output image from that GAN (at the top of the Laplacian Pyramid) is fed into the next network. This concomitant upsampling yields an image of significantly higher quality than one produced by a vanilla GAN and is illustrated in the top left. A process similar to the one illustrated here will be used in the proposed application.

by the poor translational invariance of Convolutional Neural Networks used in the LAPGAN's discriminators.

After the scripts for interfacing with the LAPGAN are complete, an application will be created that facilitates communication between machines. The application will utilize PyQt and sockets to create the user interface and communicate with other computers respectively.

If time permits, a script will be created to interface with the built-in cameras on computers running the application. If completed, an attempt will be made to extend the application to mobile devices.

4 Term Project Summary

During the weeks leading up to the end of the semester, there are four aspects of the project that need to be started and then completed: creating a LAPGAN in pytorch, creating scripts for data preprecessing, writing the software so that the application can communicate with other computers, and writing the software for the user interface.