Reconstruction of blurry/pixelated facial images

Aarsh Patel & Michael Sadler

March 21, 2018

1 Problem

What is the problem that you will be investigating? Why is it interesting?

Answer:

The high level problem we are trying to solve is the task of transforming a single low resolution input into a high resolution image (super resolution). This project will use Convolutional Neural Networks (CNNs) to reconstruct blurry/pixelated images of human faces. Our CNN will utilize image processing techniques (i.e. edge detection, trend removal, adaptive smoothing filters, gray-scale mapping, etc..) to increase the definition of an image and in hopes to accurately construct a non-blurred version of the image. The goal is for the CNN to identity the details/features in the obfuscated image in order to recover the original image. This could be useful in forensics, for example, when a security camera records at a very low resolution and certain fine-grain details in the recording are not clearly visible.

2 Data

We will be using the following datasets:

- **Labeled Faces in the Wild** - A dataset of more than 13000 images of faces collected from the web.
- **FaceScrub** - A large face dataset with 100,000 Face Images of 530 People.

We will use preprocessing techniques such as **pixelation** and **blurring** inorder to obfuscate the images.

3 Approach

What method or algorithm are you proposing? If there are existing implementations, will you use them and how? How do you plan to improve or modify such implementations?

Answer:

We first want to create our dataset of blurry/pixelated faces. We will use some free open source software inorder to blur and pixelate the faces in both of our datasets. We will also crop the image, while maintaining the center of the image. This will reduce the size of the image and thus prove to help computation of the algorithm.

We are planning to use a Convolutional Neural Network, with different filtering techniques being applied at each layer in the network, to enhance our model's prediction of the image. The input to the CNN will be an obfuscated image and the output will be some reconstruction of that image. We will first consider a simple CNN architecture (baseline model), but will consider more advanced architectures later on (ResNet, Inception, VGG). We will most likely use style transfer in which we can train our model on ImageNet and then finetune it on our dataset. We are planning to use pixel loss as our loss function for the network, which is essentially just the euclidean distance between the network output and the ground truth image.

4 Evaluation

How will you evaluate your results? Qualitatively, what kind of results do you expect (e.g. plots or figures)? Quantitatively, what kind of analysis will you use to evaluate and/or compare your results (e.g. what performance metrics or statistical tests)?

Answer:

Two popular evaluation metrics in the task of super resolution are **PSNR** (peak signal-to-nose ration) and **SSIM** (structural similarity). PSNR measures the quality of the reconstruction in terms of the absolute error. SSIM is a similarity metric that measures the perceived changes in structural information of the images. Since these metrics aren't very much correlated to human perception, we will also manually evaluate the reconstruction of the images.

5 Literature Review

The readings that we will look at inorder to provide some context and background to the problem are the following:

- Reconstructing Obfuscated Human Faces (http://cs231n.stanford.edu/reports/2017/pdfs/223.pdf)
- Defeating Image Obfuscation with Deep Learning (https://www.cs.cornell.edu/ shmat/shmat_imgobfu
- EnhanceNet: Single Image Super-Resolution Through Automated Texture Synthesis (https://arxiv.org/pdf/1612.07919.pdf)

- Fast and Accurate Image Super Resolution by Deep CNN with Skip Connection and Network in Network (https://arxiv.org/pdf/1707.05425.pdf)

  Reading these papers will provide us background necessary to tackle this project and will bring light to some ideas that we can try to implement.