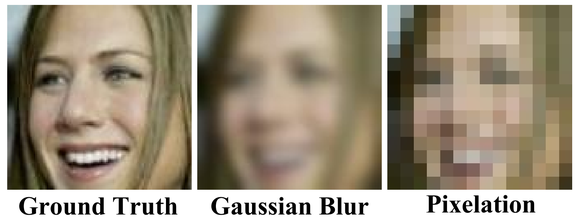
Reconstructing Obfuscated Human Faces

Jacob Conrad Trinidad

Motivation

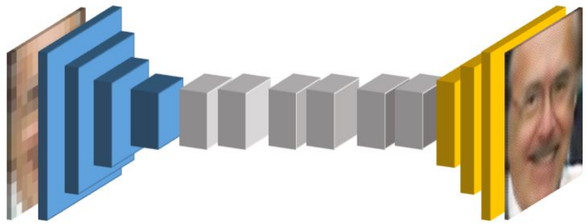
In photographic media, faces are often obfuscated to protect the identity of those pictured. This obfuscation process is done by removing details from the photograph in order to reduce the identifying facial features. In this work, we hope to examine these obfuscated images by attempting to recover these facial details. We will use convolutional neural networks in this image transformation task of super-resolution.

The input to our algorithm is an obfuscated image of a person's face. We then use a convolutional neural network to output a reconstructed image of the person's face. We'll examine two types of loss functions: one based on per pixel loss and one focusing on perceptual loss. This will be applied to the Labeled Faces in the Wild data set using two types of obfuscation techniques on faces: pixelation and Gaussian blurs. Results will be measured using the PSNR and SSIM metrics as well as visual appearance.



Ground Truth | **Gaussian Blur** $\sigma = 4$ | **Pixelation** $n = 8$

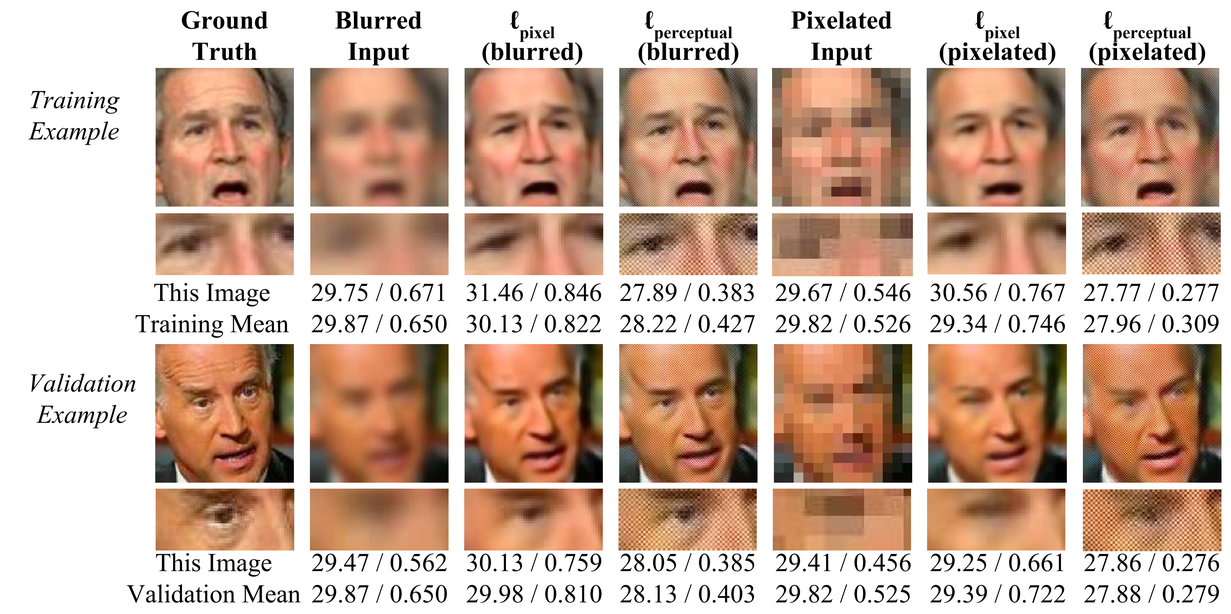
Model



The model is an image transformation network with an input and output shape of 112x112x3. The model initially subtracts the mean training image and divides each channel by 255. Then we send it through 4 convolutional layers, with a stride of 2 on the last 3 layers. Then we send it through 3 residual blocks, with each block is 2 convolutional layers. Then it goes through 4 transpose convolution layers, with a stride of 1/2 on the first 3 layers. All convolutional layers are followed by batch normalization and ReLU activation, exception the last layer which is followed by a tanh activation and then scaled to the range [0,255].

The network uses an Adam optimizer to update parameters. When using perceptual loss, we minimize content loss at the relu2_2 layer from the VGG-16 loss network when fed the image transformation network output. When using pixel loss, we minimize the l2 loss between the ground truth and the network output.

Experiments and Results



We trained with a batch size of 8 for 15k iterations with a learning rate of 1e-3. We evaluated this model architecture using four experiments which compared the results of pixel loss vs perceptual loss on either pixelated input (n=8) or blurred input (σ=4) as training data. The above figure shows the PSNR / SSIM for the given examples and for the given dataset.

The PSNR / SSIM metrics show that ℓ_{pixel} performs better than $\ell_{perceptual}$ but the examples show that this does not necessarily translate to better images. The models trained on perceptual loss generate more details and sharper images despite the obfuscation type. These details can be more clearly seen seen in the white's of the eyes and wrinkles in the zoomed eye regions.

Data Set

We will use the Labeled Faces in the Wild data set. This data set contains 13,233 images collected from the web. A benefit of using this data set is that the images are preprocessed, centering the faces and scaling the image to capture the head and so it's featured at a uniform size. For each image in the data set, we preprocess by first cropping the image from 250x250 to 112x112. This is the facial region we will focus on for obfuscation. This dataset was randomly split into 8000 training samples, 2646 validation samples, and 2587 testing samples.

The first obfuscation method is pixelation. This is the process of dividing up the area of interest into n by n squares. For each square, the average pixel color is calculated and then every pixel in that square is set to that color. We used n=8, 12, and 16.

The second method is using Gaussian blurs. In this process, an image is blurred by convolving the area of interest with a Gaussian kernel, resulting in a less blocky version of the original image in comparison to pixelation. For these blurs, we used σ=4,6, and 8.

Discussion and Analysis

Pixel loss appears to produce higher PSNR / SSIM metrics because these metrics quantify low level features, which is what pixel loss optimizes. However, perceptual loss produces more pleasing images because it focuses on minimizing the differences in higher level content features.

Several common facial details tend to be misrepresented in the reconstruction of faces, such as teeth, hair, and wrinkles, due to the loss of information in the obfuscation process. In particular, wrinkles are often not shown because pixelation and blurring hides this information and the models often struggle to learn it. This can potentially be improved with changing hyperparameters such as the content feature layer to better the model's capacity to learn. Further work will be done to optimize the model, such as adding additional layers to the model and tuning hyperparameters. Additional experiments will be done to stress test the model by feeding in faces that are more obfuscated and thus have less data to see how well it can be reconstructed.