DSC 180B Project Checkpoint 3

Final Report Draft

Artistic-Style-3 (Emmanuel Diaz): RestoreNet

Quantifying Image Restoration of World War II Photos

Abstract

In this project, we look at an approach towards restoring degraded images from the World War II era. These target photos have certain artifacts like "photo creasing" that develop due to natural factors in handling photos and our goal is to devise a method that attempts to restore these images back to their original form. We collected a dataset of World War II images and selected degraded images to analyse features of the degradation process, like Signal-to-Noise Ratio and Structural Similarity Index, in order to further understand reversing this process. Using a Generative Adversarial Network called "Deep Generative Prior", we show how an inpainting restoration is sensitive to the training data and observe the difficulties of proving a 'good' restoration subjectively and quantitatively.

Introduction

The field of digital image restoration has been developed for more than 30 years with its foundations in signal processing techniques. These techniques helped remove noise from 'degraded' images which are images that have some form of reduced quality from their original version. However, these image processing techniques have their bounds in how much lost information from degradation may be recovered. The recent advancements of neural networks and collections of huge image datasets have led many researchers to believe that neural networks hold some potential in proper image restoration.

Previous Work

Researchers began looking at different types of neural network architectures to find the best in restoring images. It began with the use of Recurrent Neural Networks (RNNs)[1] which takes the degraded image as input and performs a feedback loop, further restoring the image in each network pass. This method has shown perceptually good results on small dimensional test images, however RNNs suffered from inherent biases of the initial conditions, causing this method not to generalize well beyond a few examples. The next network challenger up to the task was the Convolutional Neural Network (CNN) Autoencoder[2] which use CNNs as feature extractors and encode the convolutions to a small dimension latent variables then deconvolutional layers learns the upscale process for these variables, giving a 'cleaned' image. CNN Autoencoder proved exceptional perceptive results for denoising and inpainting but its complexity and optimization left researchers to keep exploring new methods.

A breakthrough in image restoration came through with the application of Generative Adversarial Networks (GANs)[3] which gave rise to many facets in image processing and computer vision like DeepFakes[4]. GANs are constructed using two networks, the Generator and Discriminator, where the Generator creates constructed samples from the distribution of the training set and the Discriminator classifies these generated samples as being 'real (could be from the training set)' or 'fake' (not likely in training set). These two networks have contrasting objectives as the Generator wants to create 'realistic' samples to fool the Discriminator, and the Discriminator wants to detect as many 'fake' samples that the Generator creates. This adversarial battle leads each network to achieve better loss each iteration and once certain epochs are reached, the training is terminated and the Generator is kept to create similar examples from the training set.

Problem

It is widely known that old photos converted from film to digital might have experienced some degradation in the process. World War II photos are no exception and properly restoring these images will benefit us in preserving historical documents. We will investigate if using GANs to approach the 'image inpainting' problem, where sections of the image are obstructed and we have to 'fill in the blanks', will give us not only a perceptual improvement in quality but also a quantification of how well the restoration occurred.

Data Description

To approach this problem, we start by investigating the population of World War II images. The data collected includes ~3,000 digital images of events from World War II

scraped from ww2db.com[5]. The images were separated between three categories: Pre-War, Mid-War, and Late-War.
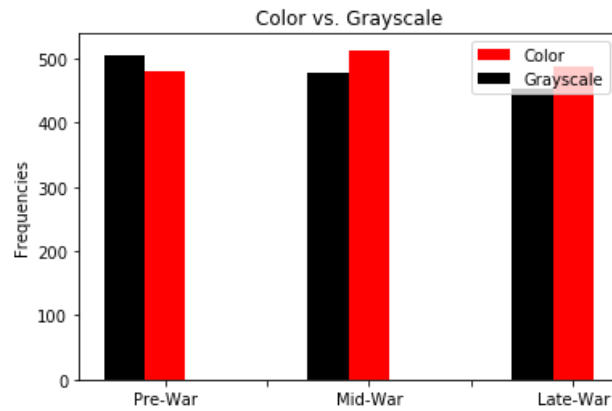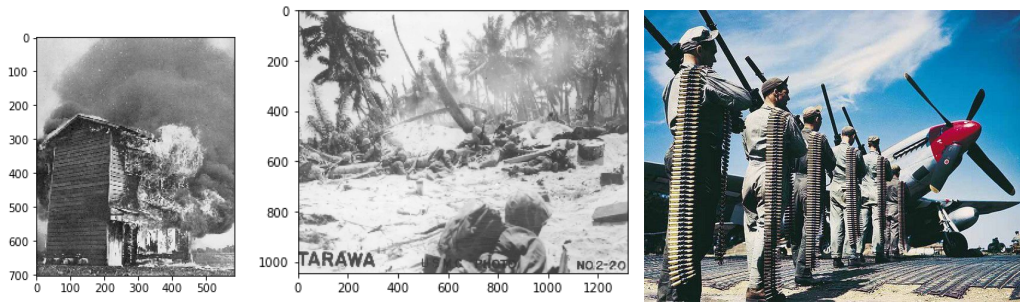

Fig (1)

 Fig (2)-(4)

Similar amounts of color and grayscale images were collected for each time period in order to have enough data to represent both subsets of the population in further analysis/training. The population we are seeking to refine is the collection of World War II photos and by analysing a sample of these photos, we hope to gain insight behind the degradation process.

Better Chart Visualizations coming here soon

Methods

Quantifiable Metric

Describe my combination of PSNR and SSIM for quantitative metric

Pre-Processing

Given a target image that is of size $h \times w$, it is loaded into the model where we apply a degradation transformation. To perform this, we use the previously defined binary mask $Mask(x, y)$ and receive the output $Target'(x, y) = Target(x, y) \cdot M(x, y)$. For the block inpainting, we crop a square in the center of the image of sides $min(h, w)$. This block will be what the network will attempt to predict. For the selective mask inpainting, the mask applied is chosen based on the region to be restored (the creases).

Choosing the model

        In approaching a GAN architecture, we chose "Deep Generative Prior"[6] which is based on the BigGAN[7] layers/data-loaders as well as "Deep Image Prior"[8], an approach for image restoration learned from parameterizing the network based on the image rather than learning the parameters of the image on the network. *more detail*

Training the model

        Detailed description of model training process

Parameter Selection

        Block inpainting, we have resolution=256x256, 2000 iterations *more detail*
        Update parameters for selective mask

Model Evaluation

        Difference between ground truth and inpainted estimates
        Creating table of mse_loss, PSNR-SSIM metric
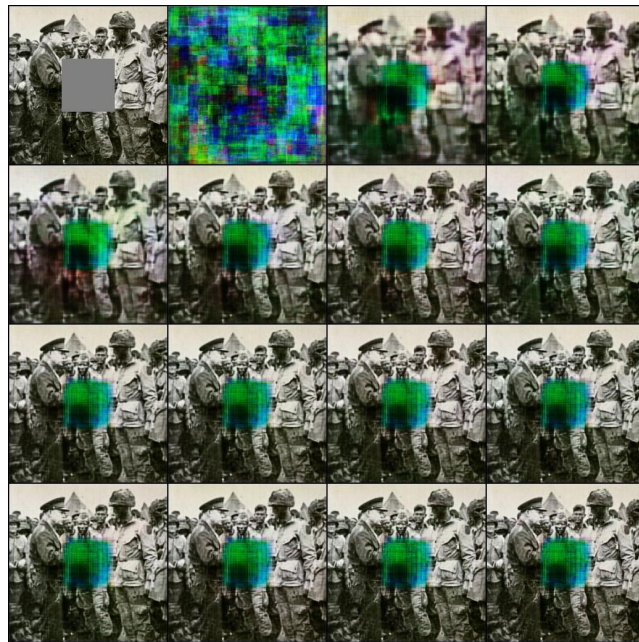
Results

Figure (5): Using block inpainting w/ Random Generator
Final PSNR: +6.76
Final SSIM: 0.1427

Figure (6): Using block inpainting w/ trained weights
Final PSNR: +11.46
Final SSIM: 0.1394

In Progress
Figure (7): Using selective inpainting mask w/ trained weights

Discussion

The results seen so far show some promise in using the weights over the 'random generator' method. As we see in the first iterations between figure (5) and (6), figure (5) starts with a random image prior and does not converge to a semantically improved restoration as it had no prior knowledge of how to generate the occluded region. In figure (6), we see that the first iteration has aspects of an image learned before and this converges towards a visually better restoration *discuss metrics*

- Describes how the results compare with prior work on the problem

- Discusses the impact and applicability of the results

- Discusses limitations of the approach and why the results are stronger (or more broadly applicable).

In the future, we would like to extend this investigation into different restoration objectives like super-resolution, denoising and colorization. We would also like to create an automated selective mask creator that can detect the creases and segment them to make a mask for the selective mask GAN process. *more detail*

References

[1] *On learning optimized reaction diffusion processes for effective image restoration*, Y Chen, W Yu, T Pock, 2016 https://arxiv.org/pdf/1503.05768v2.pdf

[2] *Image Restoration Using Convolutional Auto-encoders with Symmetric Skip Connections,* Xiao-Jiao Mao and Chunhua Shen and Yu-Bin Yang, 2016, https://arxiv.org/abs/1606.08921

[3](*GANs PAPER*)

[4](*DeepFake PAPER*)

[5]Chen, C. P. (n.d.). *World War II Database*. Lava Development LLC, Retrieved from https://ww2db.com/photo.php

[6]*Exploiting Deep Generative Prior for Versatile Image Restoration and Manipulation*, Pan, Xingang and Zhan, Xiaohang and Dai, Bo and Lin, Dahua and Loy, Chen Change and Luo, Ping, 2020 https://arxiv.org/pdf/2003.13659.pdf

[7] BigGAN - https://github.com/ajbrock/BigGAN-PyTorch

[8]Ulyanov, D., Vedaldi, A., & Lempitsky, V.S. (2018). *Deep Image Prior*. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 9446-9454.

Appendix

In this section, I will describe thoroughly

- Peak Signal-to-Noise Ratio
    - Definition
    - Interpretation
    - Significance to Image Quality
    - Example
- Structural Similarity Index
    - Definition
    - Achieving rigid motion invariance for similarity comparison
- *more terms from Methods section*