Old Photo Restoration: Pix2Pix vs Partial Convolutions

Nivedita Rahurkar and Daria Reshetova
CS230: Deep Learning

Problem

Damaged B&W photo $\overset{?}{\Longrightarrow}$ new-looking photo

There are two ways to solve this problem-

- Using image processing software, ex. Photoshop. This takes a considerable amount of time and money.
- Using Machine learning, solving it as an **image-to-image translation**[1] problem.

We decided to go with the second approach, using a cGan and a modified VGG16 architecture to restore the damaged images.

Data

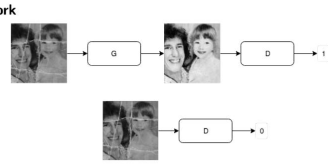
- ≈ 200 manually acquired paired images of old photos and their Photoshopped versions
- varying image sizes and quality
- 100 training set, 20 test and 20 validation

Preprocessing

- resizing to $(256, 256)$
- aligning damaged and photoshopped images
- applying Fourier transform and a high-pass filter to it as tears and rips usually have high frequency and thresholding the image so that 10% of the pixels are noisy to get a damage mask
- applying the mask to a photoshopped image

Method 1: Conditional GAN

Framework



- UNet-based generator
- Patch-discriminator

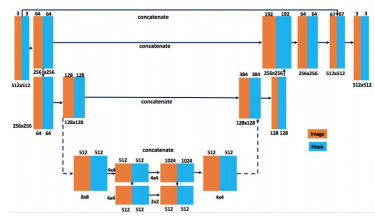
Loss We used a cGAN loss combined with L1(MAE) loss on the generator to encourage less blurring: if x is the damaged image, y is the restored image and z is the noise (damage mask)

$$\mathcal{L}(G, D) = \mathbb{E}_y[\log(D(y))] + \mathbb{E}_{(x,z)}[\log(1 - D(G(x, z))] + \lambda\mathbb{E}[\|y - G(x, z)\|_1]$$

The objective is a minimax problem:

$$G^* = \arg\min_G \max_D \mathcal{L}(G, D)$$

Model Architecture



Method 2: Partial Convolutions

- Treat damaged image restoration as an inpainting problem
- Assumption: Most damaged pixels are white(more than 0.9 * max intensity)
- UNet-based architecture, replacing conv layers with partial conv layers [2]

Pconv layer: For the conv. filter W and the corresponding bias b

$$x' = \begin{cases} \mathrm{W}^T(\mathrm{X} \odot \mathrm{M})\frac{1}{\mathrm{sum(M)}} + b, & \text{if } \mathrm{sum(M)} > 0 \\ 0, & \text{otherwise} \end{cases} \qquad m' = \begin{cases} 1, & \text{if } \mathrm{sum(M)} > 0 \\ 0, & \text{otherwise} \end{cases}$$

M is a binary mask of 0s and 1s.

- After every Pconv if the convolution was able to condition its output on at least one valid input, then the mask is removed at that location.

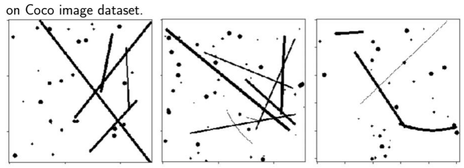
Loss We borrowed the loss function from [2]. It includes - Per-pixel losses, Perceptual loss based on ImageNet pre-trained VGG-16, Style loss on VGG-16 features, Total variation loss for a 1-pixel dilation of the hole region

$$\mathcal{L}_{total} = \mathcal{L}_{valid} + 6\mathcal{L}_{hole} + 0.05\mathcal{L}_{perceptual} + 120(\mathcal{L}_{style_{out}} + \mathcal{L}_{style_{comp}}) + 0.1\mathcal{L}_{tv}$$

Training

We initialized our network with VGG16 weights.

Mask Generation We simulated damaged images by applying randomly generating masks containing circles, lines and ellipses of varying sizes and number on Coco image dataset.



Midway through our training we had to increase the amount of damage in simulated images as the restoration was poor for heavily damaged images.

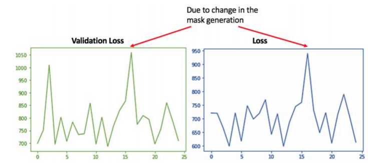
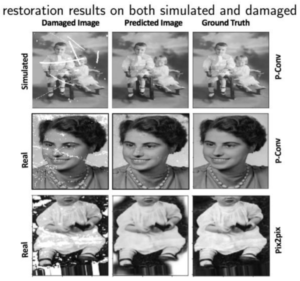


Figure 1: Losses

Results

We show the restoration results on both simulated and damaged images.



Discussion and Future Work

- The restoration results obtained on simulated damaged images are better than the real damaged images since both the images have different distributions. However, for training we only used simulated ones as the corresponding masks for real damaged images were not available.
- We observed that restoration fails near corners because: 1) Less information from neighborhood patches around corners as compared to interior patches. 2) Limited images with damage around corners in the training set
- In future, we plan to resolve the issue of poor results at damaged corners by generating more masks with non-zero pixels around corners for training. We would also like to estimate the damaged regions in an image better by training it for the same. Finally we would like to color the image.

References

[1] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros. Image-to-image translation with conditional adversarial networks. *arXiv preprint*, 2017.

[2] Guilin Liu, Fitsum A Reda, Kevin J Shih, Ting-Chun Wang, Andrew Tao, and Bryan Catanzaro. Image inpainting for irregular holes using partial convolutions. *arXiv preprint arXiv:1804.07723*, 2018.