Old Photo Restoration: Pix2Pix vs Partial Convolutions
Nivedita Rahrkar and Daria Reshetova
CS230: Deep Learning

Problem
Damaged&B&W photo ⇒ new-looking photo
- Using image processing software, e.g., Photoshop. This takes a considerable amount of time and money.
- Using machine learning, solving it as an image-to-image translation problem.

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

Data
- 250 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
- resize to (256, 256)
- Aligning damaged and photoshopped images
- Applying Fourier transform and a high-pass filter to the text and rips usually have high frequency and thresholding the image so that 70% of the pixels are noisy to get a damage mask
- Applying the mask to a photoshopped image

Method 1: Conditional GAN

Framework
- U-Net based generator
- Patch discriminator

Loss We used a GAN 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)

$$L(G, D) = E_x[log D(x)] + E_{z,y}[log(1-D(G(x, z)))] + \lambda E_z[|y-G(x, z)|_1]$$

The objective is a minimax problem:

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

Method 2: Partial Convolutions
- Treat damaged image restoration as an inpainting problem
- Assumption: Most damaged pixels are white (more than 0.95 max intensity)
- U-Net based architecture, replacing convolutional kernel with partial convolution layer [2]

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

$$\hat{y} = W'X \odot M + b,$$

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

After each 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 loss, Perceptual loss based on ImageNet pre-trained VGG-18. Style loss on VGG-16 features, Total variation loss for a 1-pixel dilation of the hole region

$$L_{total} = L_{x2} + 0.01L_{perceptual} + 100L_{style} + 0.0001L_{tv}$$

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 Costa 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.

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