NTU Deep Learning for Computer Vision (DLCV) Fall 2019 Final Project
Dunhuang Image Restoration

Alberto Herrera R08922163 * Cheng Huan An T08902105 * Zeynep Cankara T08902113

Group 10 : Dropouts

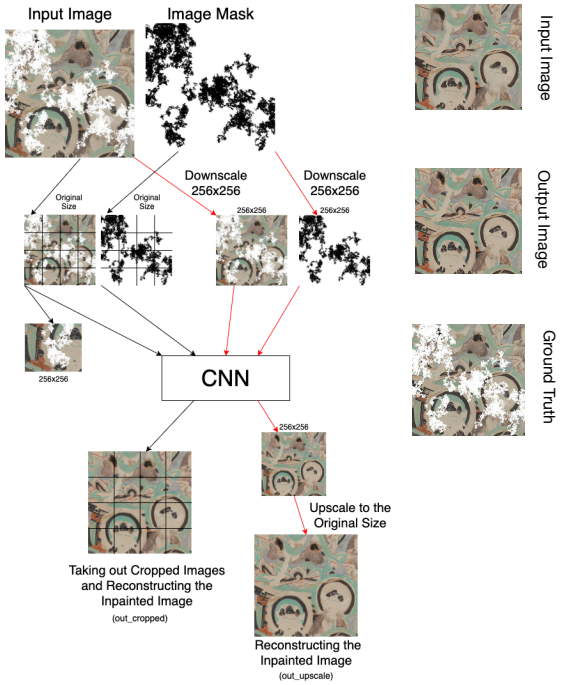
Introduction

- The mural paintings from Dunhuang caves suffering from corrosion and aging.
- We used image inpainting, a task of synthesizing contents in the missing regions to generate images which are as close as possible to the original image.
- Ground truth images are encoded in Adobe RGB, while the input masked images are in sRGB. The encoding difference creates some ripples that decrease the validation accuracy. We propose a weighted average methodology to smooth out the final result.

Training Data Augmentation

- Random crop a 256x256 patch from the input image.
- Select a random mask and rescale it to 256x256
- Apply morphological closing to remove holes in the mask
- Randomly flip and rotate (0°, 90°, 180°, 270°) the image and mask.

Image Inpainting Pipeline

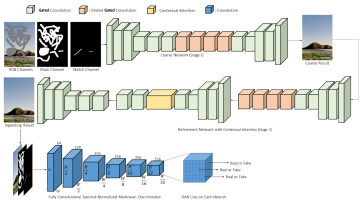


Equation for the Final Inpainting Result

Output = 0.5 * (Input * inverted mask) + 0.5 * (out_cropped) + 0.5 * (out_upscale * Mask)

Network Architecture

- Encoder-Decoder ConvNet being used to generate inpainted images.
- We used Gated Convolution Layers since they can avoid carrying convolution information from the masked regions of the images.



Experiment Results

Changes to the Model	MSE	SSIM	Final Score
Baseline	92	0.79	**1-MSE/100+SSIM**
- Coherent Semantic Attention	37.00	0.7953	1.4252
+ Pretrained Weights	37	0.80	1.429
+ Random Crops on Input Images	36.1	0.80	1.443
+ Different Kinds of Mask	36.5	0.80	1.437
+ Dilated Masks	36.6	0.80	1.438
+ Colour Space Correction	34.84	0.8105	1.462
+ Random Crops with Sliding Window Approach and Weighting the Crops together with the Original Image	33.35	0.8213	1.4877

- '-' using Coherent Semantic Attention for Image Inpainting (Hongyu Liu, Et al.) for CNN backbone
- '+' denotes to using Free-Form Image Inpainting with Gated Convolution (Jiahui Yu, Et al.) for CNN backbone.

Conclusion

- We propose Encoder-Decoder CNNs suitable for image restoration problems with low amount of data.
- Gated Convolutional Layers can use the mask information to process only valid data and improve the inpainting results.
- Taking weighted average of the generated and original images can smooth the output and improve the model performance.