Image Restoration of Noisy and Low-Quality Retinal Images

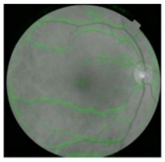
Katherine Sytwu[1], Fariah Hayee[2]
Dept. of Applied Physics[1] and Electrical Engineering[2]

High quality medical images are needed for diagnosis

Medical images contain subtle features that are crucial for reliable medical diagnosis.

Recent developments in deep-learning has enabled detection of diabetic retinopathy and even prediction of cardiovascular disease risks from retinal fundus images.

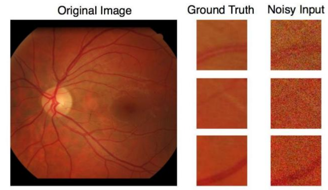
- Diagnosis depends on availability of high resolution images
- A high quality fundus camera can cost $15,000



Attention maps for hemoglobin and blood pressure detection [1] are highest around thin vessels, signifying the importance of image resolution

Aim: Use deep-learning networks to restore/improve image resolution of retinal images that retain small features like retinal blood vessels, which may enable point-of-care diagnosis (image classification) with low-quality (cell-phone) camera images.

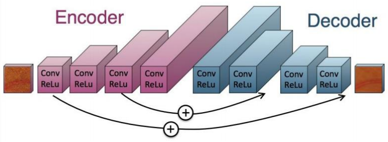
Dataset of retinal images



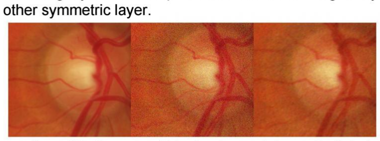
Train set: 436 high-resolution retinal images were downloaded [2] and randomly cropped into smaller images (80x80 pixels) to create a training set of >500,000 images. We then added Gaussian noise to all channels with a standard deviation of 25/255 to replicate a low-quality image. We used 148,224 patches to train the autoencoder and 501,888 images to train the modified deep-CNN, with 128 image batch size.

Test set: 26 high resolution images (1696x1696 pixels)

Autoencoder Network



Baseline network: A 9-layer convolutional autoencoder network with skip connections. It has 4 encoding layers with increasing number of filters, and 4 symmetric decoding layers, with skip connections connecting every other symmetric layer.



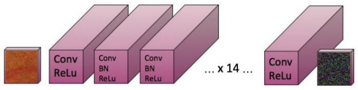
Ground Truth **Noisy Input** **Autoencoder Output**

Results: The network recovers both the thick and most of the thin veins visibly well enough, but at the expense of keeping the noise. This can be due to the skip connection being close to the output layer without any weighting factor.

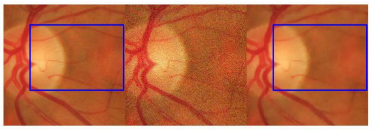
Ground Truth **Autoencoder Output**

The network can resolve thin vessels in images with higher contrast values and lighter background (first 2) while performing poorly for brighter background images.

Deep Convolutional Network



Deep CNN: To improve denoising we use a network with 17 identical convolutional layers (64 filters, 3x3 kernel) that learn the noise of the input, which can then be subtracted from the input image.
The network is pre-trained on standard color image-sets [4].



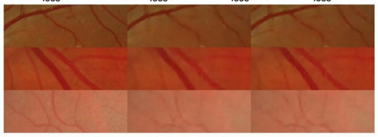
Ground Truth **Noisy Input** **CNN output**

Results:

- Performed well in removing noise and retaining large features,
- Small features like thin veins (highlighted in blue) were often blurred.

Modified loss function:

$$L2_{loss}(R,G,B) = L2_{loss}(R) + 3L2_{loss}(G) + 2L2_{loss}(B)$$

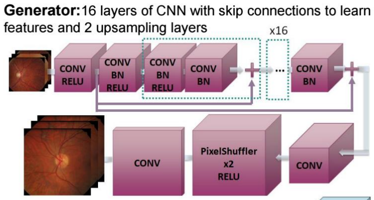


Ground Truth **Pre-trained CNN** **Modified CNN**

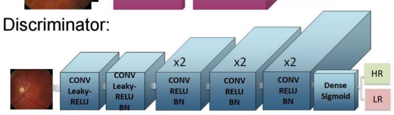
This modified network produces images with slightly better contrast than the original CNN, but still blurs out the thin veins.

Super-Resolution GAN

Generator: 16 layers of CNN with skip connections to learn features and 2 upsampling layers



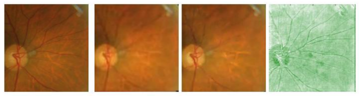
Discriminator:



Goal: Resolve finer texture details in the denoiser output by using a GAN as a super-resolution block

$Loss_{GENERATOR} = 1\times10^{-3}$ x Adversarial loss + content loss (pixel-wise mse loss) + 1×10^{-6} x V_{GG-19} loss (euclidean difference between features of ground truth and generated image)

- Modified pixel loss to highly weigh the loss in G-channel
- V_{GG}-19 and Generator pretrained, further generator training done on 128 retinal images



Ground truth **Low-res input** **SrGAN output** **$\Delta_{GEN-Truth}$(Green)**

The small blood-vessel contrast is not recovered, with the most difference pronounced in green channel.

Error Analysis

Network	PSNR
Autoencoder	33.00
Pre-trained CNN	39.96
Modified CNN	40.43
Modified SrGAN	41.96

A good denoised image will have high PSNR. The noisy input has a PSNR of ~22.

The difference between the modified CNN and the pre-trained CNN is 0.46±0.11.

- Modified DnCNN loss saturated at 0.3 with l.r rate decay.
- SrGAN needs to be trained longer over larger dataset to get to D_{errors}<0.7

Conclusions

- General trained denoiser CNNs cannot retain smaller blood vessels in retinal images
- Performance improves with modified loss function of disincentivize green-channel loss
- SrGAN with Vgg-19 loss can generate 'unreal' looking features
- Retinal feature exclusive loss parameter needs to be designed for the generator.

Future Work

- Hyperparameter search for the modified-CNN loss function weights
- Modify the loss function to include a term that promotes vessel structure and/or a term that promotes contrast.
- Also want to explore network performance on more complex noise-addition algorithms.

References

[1] Poplin, R., et al. *Nat. BioMed. Eng.* (2018)
[2] Decencière, E., et al. *IRBM* (2013)
[3] Mao, X.J., et al. *NIPS* (2016)
[4] Zhang, K., et al. *IEEE Transactions on Image Processing* (2017)
[5] Ledig, C., et al. *IEEE CVPR* (2016)