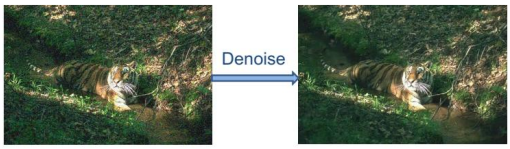
Leveraging Residual Neural Network for UNet Image Denoising with Perceptual Loss

Jingyu Liu, Xinrong Zhou, Xinyu Zhang
Department of Computer Science, University of Toronto

Motivation



Traditional model-based method

- Non-local-mean(NLM)
- Block-batching/3D filtering(BM3D)

BUT Their Limitation:

❖ Rely on prior
❖ Time-consuming optimization

Machine learning method

- Deep convolution neural network(DnCNN)
- Autoencoder

AND Their strength:

❖ End-to-end optimization

Related Work

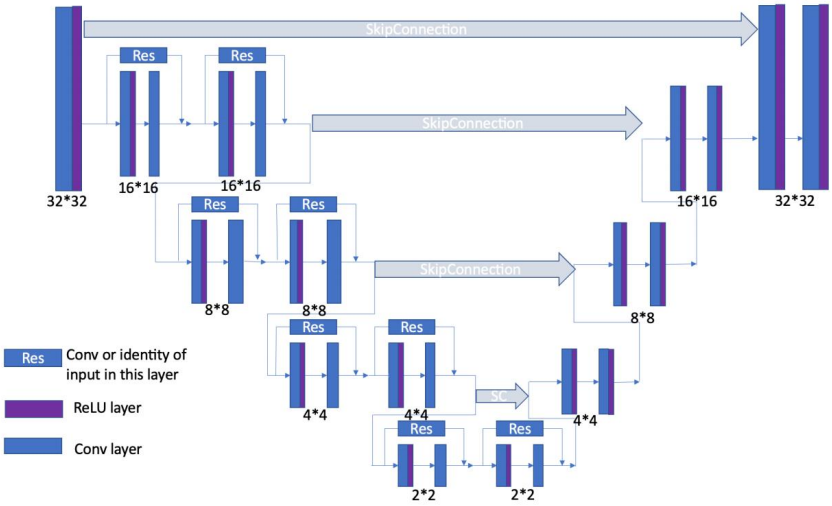
- **DnCNN [1]:** Zhang et al. proposed a deep convolutional neural network which improves the denoising performance by stacking multiple blocks of convolutional layers, batch normalization, and ReLU activations.
- **RDUnet [2]:** Javier Gurrola et al. present a residual dense neural network for image denoising based on the densely connected hierarchical network.
- **UNet [3]:** Ronneberger et al. proposed the UNet framework for Biomedical Image Segmentation. There are 4 down-sampling blocks and 4 up-sampling blocks.
- **ResNet [4]:** Zhang et al. proposed in 2016. Each block was given by adding residual learning operation to improve the performance of image recognition.

References

[1] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising," IEEE transactions on image processing, vol. 26, no. 7, pp. 3142–3155, 2017.
[2] J. Gurrola-Ramos, O. Dalmau, and T. E. Alarc´on, "A residual dense u-net neural network for image denoising," IEEE Access, vol. 9, pp.31 742–31 754, 2021
[3] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in International Conference on Medical image computing and computer-assisted intervention. Springer, 2015, pp. 234–241.
[4] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770–778

New Technique

Adapt UNet structure with residual blocks



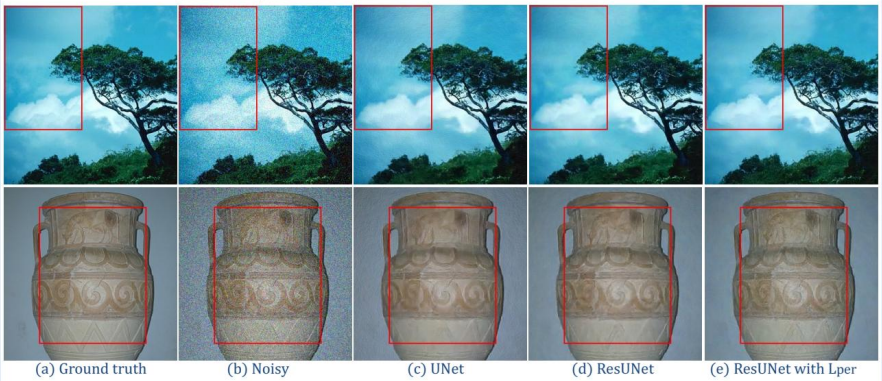
$$\mathcal{L}_{loss} = \lambda \mathcal{L}_{MSE} + \mathcal{L}_{per}$$

$$\mathcal{L}_{MSE} = \frac{\|y - x'\|^2}{W * H} \qquad \mathcal{L}_{per} = \frac{\|\phi(y) - \phi(x')\|^2}{W * H}$$

Experimental Results

- Results on denoise work:

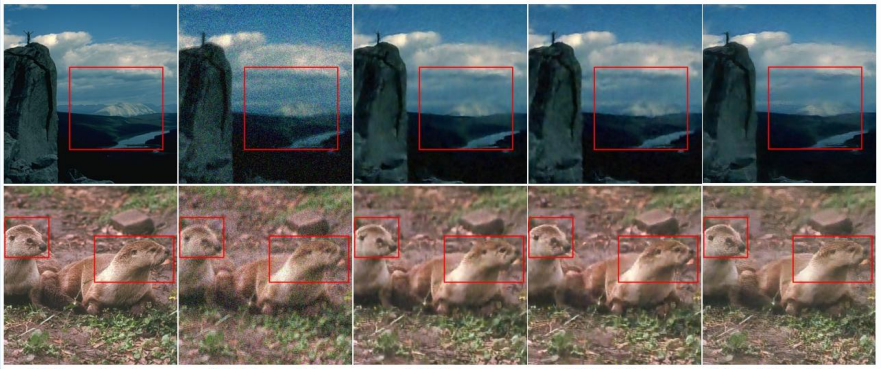
sigma	UNet	ResUNet	ResUNet with $\mathcal{L}_{per}$
0.1	29.3147	30.1055	29.5025



(a) Ground truth (b) Noisy (c) UNet (d) ResUNet (e) ResUNet with Lper

- Results on deblur-denoise work:

sigma	UNet	ResUNet	ResUNet with $\mathcal{L}_{per}$
0.1	25.6711	25.8606	25.4638



(a) Ground truth (b) Noisy and blur (c) UNet (d) ResUNet (e) ResUNet with Lper