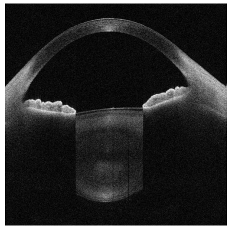
CSGAN: a consistent structural GAN for AS-OCT image despeckling by image translation

Sanqian Li, Muxing Xiong, Risa Higashita, Jiang Liu

#1 Motivations and background

What is AS-OCT?

- AS-OCT: Anterior segment optical coherence tomography, a non-invasive imaging technique, which is widely utilized in ophthalmology to diagnose anterior segment disorders.



AS-OCT image with speckle noise

Why despeckling?

- Speckle noise inherited in ASOCT images degrades the visual quality and hampers the subsequent medical analysis.

Why preserve image structure?

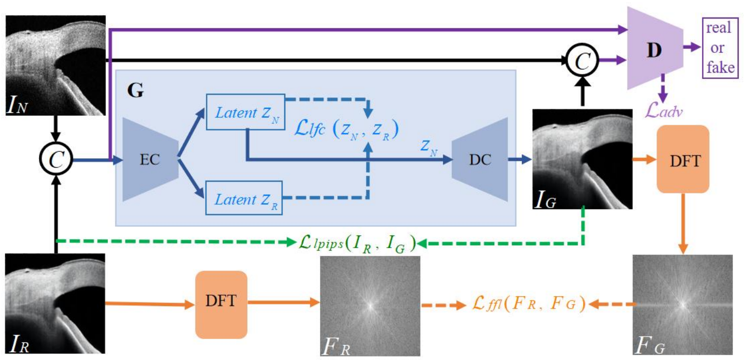
- Previous work was devoted to removing the speckles and acquiring satisfying images. But according to the clinical requirements, it is desirable to maintain locally higher data fidelity rather than enforcing visually appealing but obtaining rather wrong image structural features.

How to achieve that?

- Use our proposed Consistent Structural Generative Adversarial Network (CSGAN) to learn the clean style of low-speckle in repeated AS-OCT images and simultaneously preserve the tiny but vital structural knowledge among the latent feature, spatial and frequency domains.

#2 Methods

Model Structure



- Basic structure of GAN:
  Resnet Generator
  Patch GAN discriminator
- Latent constraint module
- Frequency constraint module
- Perceptual loss module

Objective functions

The generative network demonstrate the powerful representation ability in low-level image processing task while it always neglects some tiny but vital details for clinical practice. To balance the speckle suppression and structure consistency, we exploit objective function L_{lfc}, and introduce L_{lpips}, L_{ffl}.

$$\mathcal{L}_{lfc}(z_N, z_R) = \frac{1}{MN}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}\left|z_N(x,y) - z_R(x,y)\right|^2$$

$$\mathcal{L}_{lpips}(I_R, I_G) = \frac{1}{W_{i,j}H_{i,j}}\sum_{x=1}^{W_{i,j}}\sum_{y=1}^{H_{i,j}}\left|\phi_{i,j}(I_R)_{x,y} - \phi_{i,j}(I_G)_{x,y}\right|^2$$

$$\mathcal{L}_{ffl}(F_R, F_G) = \frac{1}{MN}\sum_{u=0}^{M-1}\sum_{v=0}^{N-1}\left|F_R(u,v) - F_G(u,v)\right|^2$$

$$\mathcal{L}_{adv} = \log D(I_R, I_N) + \log(1 - D(I_G, I_N))$$

- Latent feature constraint L_{lfc}: capture structural semantics in the latent feature maps and allow some slight variation to improve the generalization ability of network.
- Perceptual loss L_{lpips}: explicitly adopted to learn perceptual image patch similarity and incorporate structural texture and minimize the structural gap the speckled and generated image in the spatial domain.
- Focal frequency loss L_{ffl}: adopted to adaptively focus the model on the frequency components that are hard to deal with but can be pivotal for ameliorating quality in the frequency space .

#3 Experimental Results and analysis

Dataset: AS-CASIA

The visual comparison among the noisy images, repeated images, and despeckling results of the proposed CSGAN is shown in figure below. We can observe that the repeated images easily get trapped in edge artifacts at the border of the iris (see the orange arrow in the enlarged green region) and loss of detail (see the blue arrow) caused by involuntary movements during the acquisition. At the same time, the despeckling result can capture the structure knowledge in noisy images and learn the style of repeated images owing to the designing of structure consistency in CSGAN and the style understanding ability in PatchGAN-discriminator.