A Modified U-net for Dispersion Compensation of OCT Simulated Layered Images

Yueming Zhuo, and David Lindell

Abstract—Chromatic dispersion, a common problem that exists in optical coherence tomography (OCT) imaging systems, degrades the axial resolution of retinal layer structures in an OCT image. This study is to develop a simple OCT simulation with multiple back-scattering planes in the sample arm and use a modified U-net (mUnet) to enable dispersion compensation of dispersed simulated data using a set of corresponding ground truth. The effects of dispersion and noises are carefully studied. Axial depth corrections on training data is needed to enable robust network performance.

Index Terms—Optical Coherence Tomography, Dispersion Compensation, Fully Convolutional Neural Network, U-net

1 Introduction

Optical coherence tomography (OCT) is a non-invasive interferometric imaging technique which uses low-coherence light to capture micrometer-resolution images of optical scattering media, e.g., human retina. The core of OCT can be a simple Michelson interferometer, shown in Fig.1. Light from low-coherence (wide-band) light source is split into two separate arms: reference arm towards a reflective mirror, and sample arm towards a sample, which contains multiple scattering planes. The lights upon reflections interferes in the detection path and pass through a diffraction grating which decomposes the interference spatially in wavelength(λ)-space. To reconstruct OCT images in the spatial domain (z), the measured interferogram in λ-space is re-sampled to wavenumber(k)-space linearly, which are referred as the spectral interference. Taking the Fourier transform of the spectral interferogram, the spatial layer structures are reconstructed. Samples are usually dispersive mediums, therefore, lights of different frequencies will travel at different velocity within the sample, which causes peak broadening. Detailed derivations of basic OCT theory are shown in Section 3. Section 2 discusses related works on dispersion compensation of OCT images. Section 4 describes data preparation procedure. Section 5 shows the proposed method and section 6 discusses the experimental results.

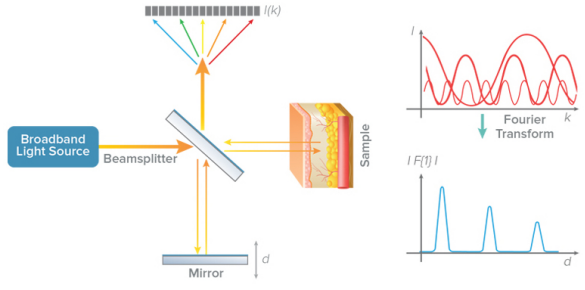

Fig. 1. Simplest OCT setup [1].

- *Y.Zhuo, D.Lindell are with the Department of Electrical Engineering, Stanford University, California, CA, 94305.*

2 Related Works

Dispersion compensation has been extensively studied in the literature. There are hardware-based approaches and software-based approaches. Hardware-based approach basically utilizes glasses (e.g., SF11) and place them in the reference arm to compensate the sample dispersion as well as the dispersion introduced by the different optics used within two arms. Hardware-based approaches are not very flexible as glasses or other compensating optics are made with fixed dimensions (especially thickness), therefore, it is hard to compensate the system accurately. Numerical corrections are extensively explored in the past twenty years. Fercher et al. [2] used a depth-dependent kernel to correlate the spectral interference to compensate dispersion. However, this method relies on the prior knowledge of the dispersive characteristics of the sample, which could be very different among different biological tissues. Wojtkowski et al. [3] applies a phase correction term (up to third order) on the spectral interference and updates this correction term to optimize a custom sharpness metric of the measured OCT image. However, different depths of the biological tissues require different amount of compensation, therefore, a single phase mask for all depths may not be very effective. Hofer et al. [4] used entropy information of the spatial signal as the sharpness metric. However, sharpness metrics in general are very susceptible to the prevalent speckle noise in OCT B-scans. There are a few other techniques: Lippok et al. [5] introduced fractional Fourier transform; Pan et al. [6] derived an analytical formula to estimate the second-order dispersion compensation coefficients in different depths based on a linear fitting approach.

3 Theory

This section provides a short summary on the optical theory of spectral-domain optical coherence tomography (SD-OCT) [7]. Adopting the Michelson interferometer, assume no reflective or scattering surfaces/planes have frequency-invariant reflectivities, 50-50 splitting at the beamsplitter, and denoting the electric field in the reference arm as $E_R(\omega)$, the mirror reflectivity as r_R, reference arm length as l_R, the

field in the reference arm can be expressed in the angular-frequency(ω) domain as:

$$E_R(\omega) = E_o(\omega) r_R e^{i(2k_R(\omega) l_R - \omega t)}. \quad (1)$$

Let the sample reflectivity profile be $r_S(z)$ and sample arm length be l_S, the field in the sample arm can be written as:

$$E_S(\omega) = E_o(\omega) \int_{-\infty}^{\infty} r_S(z) e^{i(2k_S(\omega) z - \omega t)} dz. \quad (2)$$

Note the factor of 2 in the exponents comes from the double-pass from the beamsplitter to mirror/sample. Common path length scaling factor is assumed to be factored into $E_o(\omega)$. In addition, k_R and k_S are by default functions of ω and this is the reason that dispersive effects embedded in the interference. The interference can be found as:

$$\begin{aligned} I(\omega) &= I(kc) && (3)\\ &= |E_R(kc) + E_S(kc)|^2 && (4)\\ &= I_R(kc) + I_S(kc) + I_{\text{int}}(kc). && (5) \end{aligned}$$

The constant c can be dropped out of the functional dependency. The first term is referred as the reference background:

$$I_R(k) = r_R^2 S(k), \quad (6)$$

where $S(k) = |E_o(kc)|^2$ denotes the source power spectral density distribution. The second term is referred as the sample self-interference:

$$I_S(k) = S(k) \left| \int_{-\infty}^{\infty} r_S(z) e^{i2kn_S z} dz \right|^2. \quad (7)$$

The last term, which is the most important one, is the interference term:

$$I_{\text{init}}(k) = 2S(k) r_R \int_{-\infty}^{\infty} r_S(z) \cos\big(2k(n_S z - n_R l_R)\big) dz. \quad (8)$$

For brevity in notation, set $l_R = 0$ and assume $r_S(z)$ is an even function that is symmetrical around the imaged reference mirror along the sample arm, and the interference term can be rewritten as:

$$I_{\text{init}}(k) = r_R S(k) \int_{-\infty}^{\infty} r_S(z) e^{i2kn_S z} dz. \quad (9)$$

In practice, reference background and sample self-interference term are easily removed in k-space by measuring interference while blocking one of the arms. Therefore, after background subtraction, there will only be the interference left. This term already looks like a Fourier transform itself. With a simple variable transform, i.e., $z' = 2n_S z$, the interference can be directly written as a Fourier transform:

$$\begin{aligned} I_{\text{init}}(k) &= S(k) \frac{r_R}{2n_S} \int_{-\infty}^{\infty} r_S\left(\frac{z'}{2n_S}\right) e^{ikz'} dz' && (10)\\ &= S(k) \frac{r_R}{2n_S} \mathfrak{F}\left\{ r_S\left(\frac{z'}{2n_S}\right) \right\}(k). && (11) \end{aligned}$$

Above equation shows that, we can reconstruct $r_S(z)$ by just taking the inverse Fourier transform of the spectral interference. Note that the multiplicative $S(k)$ factor behaves as a convolution in spatial domain, i.e., the spatial distribution of the source convolves with the sample reflectivity profile and thus will blur the peaks at layer boundaries. For simulated data, where the spectral shape of the source is assumed, deconvolution is possible to be implemented. However, in practice, the source spectral density is not easily measured or at least not perfectly, therefore, deconvolution is not usually done for real data.

4 Data Preparation

To simulate OCT, a continuous sample reflectivity profile is not practical. In fact, in this study, only 2,3,4 discrete backscattering planes are considered. To eliminate any non-idealities from linear k-resampling of λ-space, the simulation is based on direct sampling in k-space. The source central wavelength (λ_0) and bandwidth ($\Delta\lambda$) are, 840 and 100 nm, respectively. These values can be converted into the following k-related quantities:

$$\begin{aligned} k_0 &= \frac{2\pi}{\lambda_0}, && (12)\\ \Delta k &= \frac{2\pi \Delta\lambda}{\lambda_0^2}, && (13)\\ \sigma_k &= \frac{\Delta k}{2\ln 2}. && (14) \end{aligned}$$

Above quantities fully characterize a Gaussian spectral profile. To initialize the k-vector, the range of k which depends on the number of σ_k chosen (or detected by the spectrometer in a real setup) has to be determined such that the FWHM in λ-space is about 100 nm. In this study, the number of σ_k used is 3. The k-vector and the OCT source spectrum are initialized as:

```
lambda0 = 840e-9    # center wavelength: 840 nm
dlambda = 100e-9     # FWHM: 100 nm

# Conversion from lambda-space to k-space
k0 = 2*np.pi/lambda0
dk = 2*np.pi*dlambda / lambda0**2
sigma_k = dk/np.sqrt(2*np.log(2))
nsigma = 3
k = k0 + sigma_k*np.linspace(-nsigma, nsigma, N)

# OCT source spectrum
S_k = np.exp(-(k-k0)**2 / 2 / sigma_k**2)
```

Listing 1. OCT source initialization

Since the final simulated OCT images will be normalized, the absolute magnitudes of electric fields in the reference arm and the sample arm are not important. The relative magnitudes however are important and are determined by the reflectivity of each backscattering planes. The simulation program detects how many planes are in the sample, and sums up all the fields according to equation (4). Since the reflectivity profiles in this study are simply finite sums of shifted δ functions, the detector basically detects the magnitude of a finite sum of complex exponentials. The interference can be obtained by first setting r_S to zero and run the simulation to obtain the reference background; then setting r_R to zero and run the simulation to obtain the sample self-interference. The final interference term is obtained by background subtraction in k-space as described in the previous section. The dispersion comes into the picture when n_S in the electric fields computation is a function of k. In this study, it is assumed that the sample itself has only one

unique dispersion profile along k. In real retina, different layers have different dispersion relations, which adds even more dispersive and complicated effects into imaging.

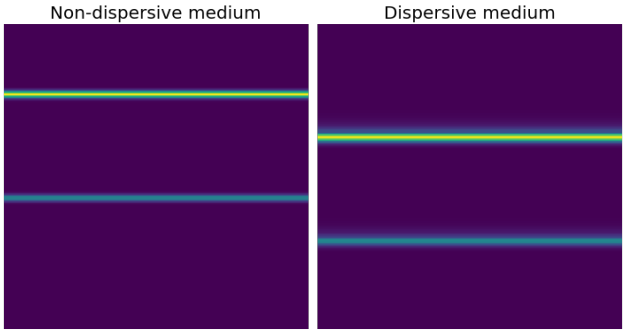

Fig. 2. Dispersive effects

Fig.2 shows an OCT image of a 2-layer structure. Dispersion not only widened the imaged layers but also shifted the whole structure. If directly using the above pair of images as the ground truth and measurement data pair for neural network training, it imposes a super challenging task for the neural network, even though the network is very sophisticated and very dedicated loss functions are used. In practice, the reference mirror can be moved to relocate the layered structures within the images. Therefore, it is not a problem by realigning the shifted measurements accordingly to match with the depths of the layers in the ground truth images. This is done using a simple intensity-based image registration technique: **cross-correlation**. The depth-mismatch or the shift between two images is determined by the lag of the maximum peak in the computed cross-correlation series. The measurement images are then circularly shifted in the vertical direction to align with the ground truth images.

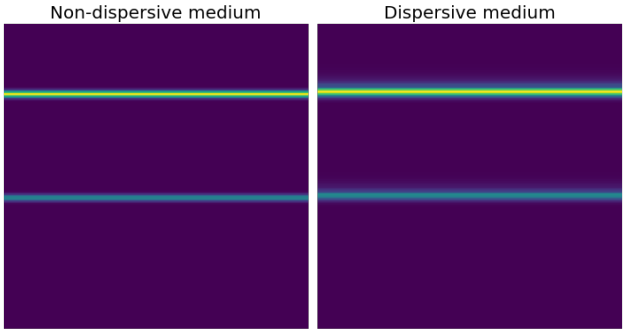

Fig. 3. Shifted corrected measurements

The resolution of the detector is chosen to be 256×256. The inverse Fourier transform of a real-valued interferogram is complex and symmetrical, the unique information is in either half of the reconstructed spatial structure, which only extends 128 pixels. The detector resolution was initially chosen to be 1024×1024, however, the resulting OCT image, size of 512×512, was impossible to be used for training. The GPU easily ran out of memory or the training time was over a day or more. The size is also chosen to be a factor 2 for mainly 2 reasons: faster Fourier transform; skip-connections in the modified U-net.

5 Proposed Methods

In this study, a deep learning approach is used to do dispersion compensation. The mUnet model is adopted from Shaiban et al. [8].

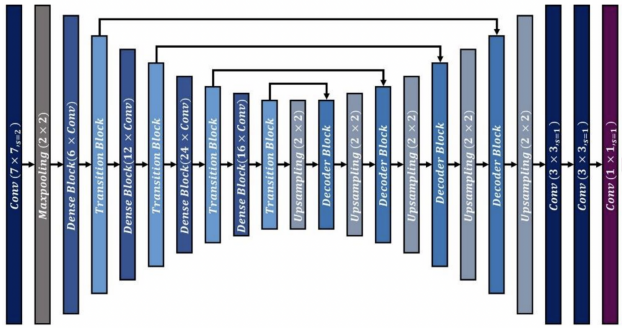

Fig. 4. Modified U-net architecture

The network is implemented in Pytorch, and optimized using Adam SGD with simple $L2$ loss. The dispersion is added using the following simple model:

```
A = 0.005
x = np.linspace(0,np.pi/2,N)
n_s = A*np.sin(x) + 1
```

Listing 2. Dispersion model

The multiplicative factor A which scales the sin function simply controls the degree of dispersive effects. The training data and test data contain different numbers of layers (2-4), different amount of dispersion (0.001-0.005), different thicknesses (1, 1.5 mm) of samples. Note that the finite thickness also controls the amount of dispersion. Furthermore, noise level of 0.05 and 0.1 are added to observe network's performance. In addition to uniform illumination across the horizontal (lateral) dimension, a Gaussian profiled illumination is also used for comparative study (see Fig. 10 for example). The ordinary Unet architecture was also trained on the same dataset and compared qualitatively with mUnet (unfortunately, the quantitative model parameters was lost; retraining was not possible within the limited time frame).

Note this method of dispersion compensation kills all phase information of the original complex OCT data, as the input to this network is the magnitude of the complex OCT image. The network will modify real pixel values, and these new pixel values cannot be converted back to complex values with meaningful phase information. Therefore, for quantitative phase applications, this method of dispersion compensation is not appropriate. Using depth-dependent spectral phase correction term will be better.

6 Results and Discussions

The total training time of the modified U-net, which contains about 1.9 millions of parameters, is about 14 minutes. The PSNR summaries of individual test images for both uniform illumination and Gaussian illumination are shown in Fig. 6. The baseline PSNRs are computed between ground truth images and the axial-corrected measurement images with no dispersion compensation.

There are a few important observations: when adding noise to the axial-corrected measurements, the baseline

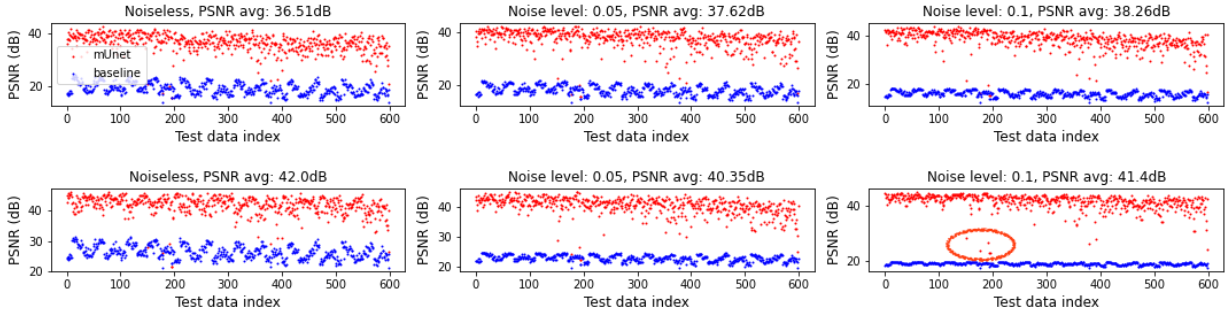

Fig. 5. PSNR summary. First row: uniform illumination; second row: Gaussian profiled illumination. Circled region: example network worse performances; this is evident in all cases.

TABLE 1
PSNR summary

-	Noiseless	$\sigma = 0.05$	$\sigma = 0.1$
Uniform (baseline)	19.02	18.15	16.33
Uniform (mUnet)	36.51	37.62	38.26
Gaussian (baseline)	26.27	23.00	18.99
Gaussian (mUnet)	42.00	40.35	41.40

PSNRs exhibit much better consistency. This makes sense as there are different amount of dispersion and different number of layers present in different OCT images. More importantly, the network seems to perform more consisent (lower standard deviation in PSNRs) when the baseline PSNRs are more consistent, for both cases. In addition, for the case of uniform illumination, adding more noises into the training data seems to gain better performance out of the neural network. The potential reason could be that mUnet sees the broadened blurry edges as noises based on the loss function used. Adding higher noises to measurements introduces more uniform distribution and thus appearing as an easier job for mUnet.

Furthermore, the red circled region shown in Fig. 6 shows some outlier performances. For these few exceptions, the measurements exhibit very high dispersion such that the distance between the planes shortened or became at least ambiguous for the cross-correlation to work out properly. An example is shown below:

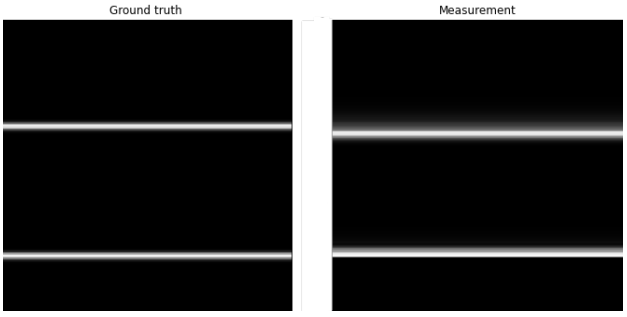

Fig. 6. Low PSNR example

As one can see that as severe dispersion occurs, the distance between layers became shorter in this case. However, the naive axial depth registration can not possibly detects this, which imposes a great challenge to the neural network, as it also has to figure out how to shift layers.

Fig. 7 shows an example of a sharpened mUnet result of a 4-layer structure with uniform illumination. It achieves a PSNR of 36.53 dB even though the dispersion was quite severe. Fig.8 shows an example of sharpening a noisy 4-layer OCT image ($\sigma = 0.05$). Even though the last layer after adding dispersion and additive noise was totally at a different position, mUnet was able to correct it and show a high PSNR of 32.61 dB. Fig. 9 shows an example of noiseless sharpening result in the case of Gaussian profiled illumination. For Gaussian illumination OCT images, mUnet achieves higher PSNRs overall as compared to uniform illumination. The primary reason is that mUnet only needed to be trained on a considerably smaller region. Fig. 10 shows an noisy case for Gaussian illumination. In the case of Gaussian illumination, adding noise does not seem to improve the network performance which is true for the case of uniform illumination. A potential reason could be that noises actually blurred the tails on both sides of the OCT signals to a significant extent (see Fig. 10 for example).

To compare with other models of neural network, the original Unet architecture was also trained using the simulated data. The mUnet performs only slightly better than the ordinary Unet even though it has about twice amount of training parameters. This makes sense as the simulated OCT data exhibits perfect symmetrical structure and very clear distinctive layers. In real OCT data, the images do not appear to be this nice at all. For example, depending on the magnification of the optical system, there could easily be some curved surfaces; there could also be some blood vessel shadows. For those images, a more sophisticated deep net may perform better.

7 Conclusion and Future Works

This study shows the performance of a modified U-net (mUnet) deep learning model on simulated OCT dispersed images. The deep neural network shows a strong recoverability for dispersed OCT images. The network can also denoise if there are additive Gaussian noises in the OCT image. There is on average of $15 \sim 20$ dB improvements in using the mUnet for dispersion compensation.

To address the issue of severe dispersion, a second-order numeric dispersion correction can be first applied to the measurements. This should reduce the distance changing

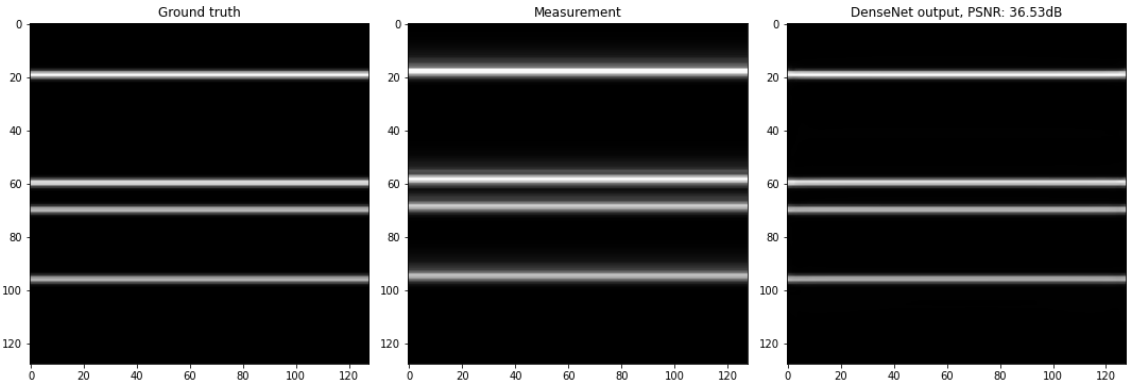

Fig. 7. 4-layer, uniform, noiseless example

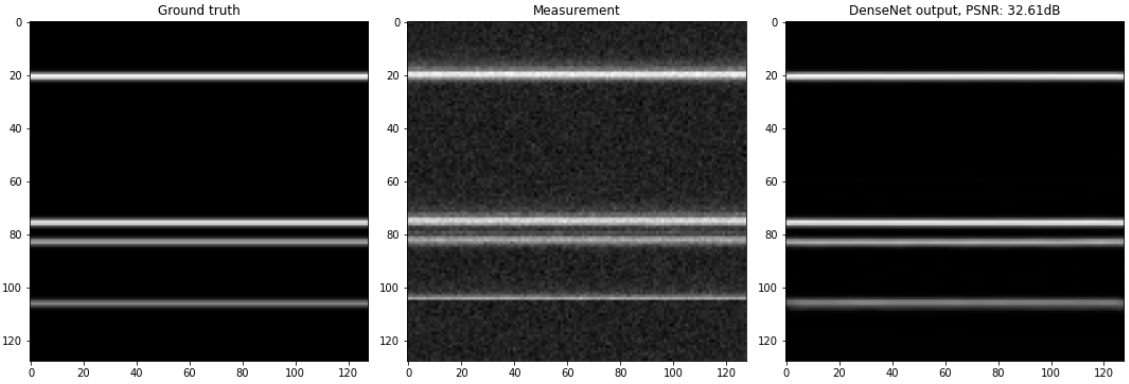

Fig. 8. 4-layer, uniform, noise level 0.05 example

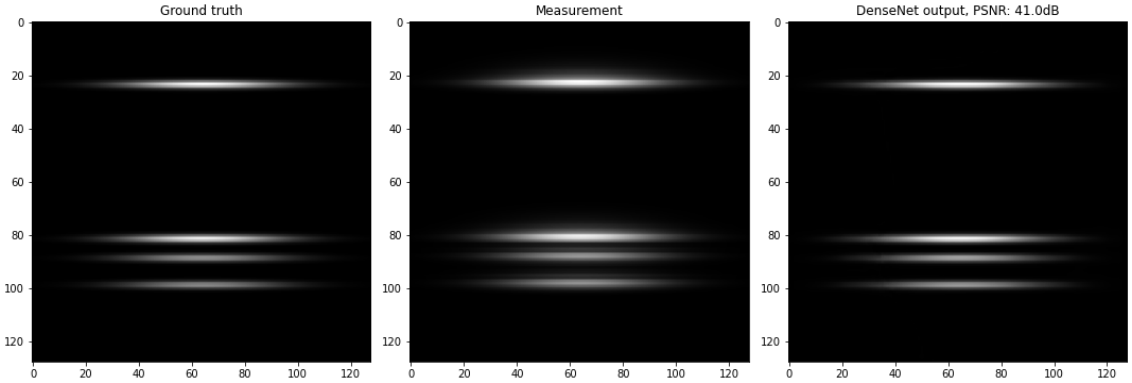

Fig. 9. 4-layer, Gaussian, noise level 0.05 example

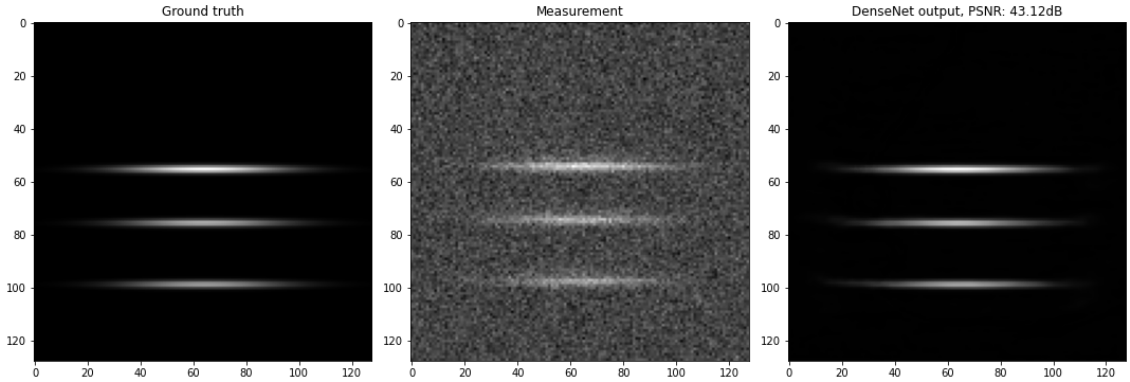

Fig. 10. 3-layer, Gaussian, noise level 0.1 example

effect as shown in Fig. 6. Then after axial depth registration, the network could be trained to output more consistent results. This could be implemented in the future as an intermediate step between data preparation and network training. Some other aspects of the deep net could also be explored for layered structures in the training data. For example, rectangular filters might be more suitable for extracting horizontal layered features.

ACKNOWLEDGMENTS

Yueming Zhuo would like to thank Mark Nishimura for valuable discussions on loss functions and Python debugging.

REFERENCES

[1] "Oct tutorial," Nov 2021. [Online]. Available: https://wasatchphotonics.com/oct-tutorial/

[2] A. F. Fercher, C. K. Hitzenberger, M. Sticker, R. Zawadzki, B. Karamata, and T. Lasser, "Numerical dispersion compensation for partial coherence interferometry and optical coherence tomography," *Opt. Express*, vol. 9, no. 12, pp. 610–615, Dec 2001. [Online]. Available: http://opg.optica.org/oe/abstract.cfm?URI=oe-9-12-610

[3] M. Wojtkowski, V. J. Srinivasan, T. H. Ko, J. G. Fujimoto, A. Kowalczyk, and J. S. Duker, "Ultrahigh-resolution, high-speed, fourier domain optical coherence tomography and methods for dispersion compensation," *Opt. Express*, vol. 12, no. 11, pp. 2404–2422, May 2004. [Online]. Available: http://opg.optica.org/oe/abstract.cfm?URI=oe-12-11-2404

[4] B. Hofer, B. Považay, B. Hermann, A. Unterhuber, G. Matz, and W. Drexler, "Dispersion encoded full range frequency domain optical coherence tomography," *Opt. Express*, vol. 17, no. 1, pp. 7–24, Jan 2009. [Online]. Available: http://opg.optica.org/oe/abstract.cfm?URI=oe-17-1-7

[5] N. Lippok, S. Coen, P. Nielsen, and F. Vanholsbeeck, "Dispersion compensation in fourier domain optical coherence tomography using the fractional fourier transform," *Opt. Express*, vol. 20, no. 21, pp. 23 398–23 413, Oct 2012. [Online]. Available: http://opg.optica.org/oe/abstract.cfm?URI=oe-20-21-23398

[6] L. Pan, X. Wang, Z. Li, X. Zhang, Y. Bu, N. Nan, Y. Chen, X. Wang, and F. Dai, "Depth-dependent dispersion compensation for full-depth oct image," *Opt. Express*, vol. 25, no. 9, pp. 10 345–10 354, May 2017. [Online]. Available: http://opg.optica.org/oe/abstract.cfm?URI=oe-25-9-10345

[7] L. V. Wang and H. Wu, *Biomedical Optics: Principles and Imaging*, 1st ed. Wiley Interscience, 2007.

[8] S. Ahmed, D. Le, T. Son, T. Adejumo, and X. Yao, "Adc-net: An open-source deep learning network for automated dispersion compensation in optical coherence tomography," *Opt. Express*, Jan 2022. [Online]. Available: arXiv:2201.12625