King's Research Portal

DOI:

[10.1007/978-3-030-87000-3_3](https://doi.org/10.1007/978-3-030-87000-3_3)

Document Version

Peer reviewed version

[Link to publication record in King's Research Portal](#)

Citation for published version (APA):

Komninos, C., Pissas, T., Flores, B., Bloch, E., Vercauteren, T., Ourselin, S., Da Cruz, L., & Bergeles, C. (2021). Intra-operative OCT (iOCT) Image Quality Enhancement: A Super-Resolution Approach Using High Quality iOCT 3D Scans. In H. Fu, M. K. Garvin, T. MacGillivray, Y. Xu, & Y. Zheng (Eds.), *Ophthalmic Medical Image Analysis - 8th International Workshop, OMIA 2021, Held in Conjunction with MICCAI 2021, Proceedings* (pp. 21-31). (Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics); Vol. 12970 LNCS). Springer Science and Business Media Deutschland GmbH. https://doi.org/10.1007/978-3-030-87000-3_3

Citing this paper

Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

General rights

Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Research Portal

Take down policy

If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Intra-operative OCT (iOCT) Image Quality Enhancement: A Super-Resolution Approach using High Quality iOCT 3D Scans ^{*}

Charalampos Komninos¹[0000-0001-5011-0521], Theodoros Pissas^{1,2}[0000-0001-9500-7085], Blanca Flores³[0000-0001-6730-5413], Edward Bloch^{2,3}[0000-0002-0799-3258], Tom Vercauteren¹[0000-0003-1794-0456], Sébastien Ourselin³[0000-0002-5694-5340], Lyndon Da Cruz^{3,4,*}[0000-0002-7695-6354], and Christos Bergeles^{1,*}[0000-0002-9152-3194]

¹ School of Biomedical Engineering & Imaging Sciences, King's College London, SE1 7EU, London, UK

`charalampos.komninos@kcl.ac.uk`

² Wellcome/EPSRC Centre for Interventional and Surgical Sciences, University College London, W1W 7TS, London, UK

³ Moorfields Eye Hospital, EC1V 2PD, London, UK

⁴ Institute of Ophthalmology, University College London, EC1V 9EL, London, UK

Abstract. Effective treatment of degenerative retinal diseases will require robot-assisted intraretinal therapy delivery supported by excellent retinal layer visualisation capabilities. Intra-operative Optical Coherence Tomography (iOCT) is an imaging modality which provides real-time, cross-sectional retinal images partially allowing visualisation of the layers where the sight restoring treatments should be delivered. Unfortunately, iOCT systems sacrifice image quality for high frame rates, making the identification of pertinent layers challenging. This paper proposes a Super-Resolution pipeline to enhance the quality of iOCT images leveraging information from iOCT 3D cube scans. We first explore whether 3D iOCT cube scans can indeed be used as high-resolution images by performing Image Quality Assessment. Then, we apply non-rigid image registration to generate partially aligned pairs, and we carry out data augmentation to increase the available training data. Finally, we use CycleGAN to transfer the quality between low-resolution (LR) and high-resolution (HR) domain. Quantitative analysis demonstrates that iOCT quality increases with statistical significance, but a qualitative study with expert clinicians is inconclusive with regards to their preferences.

Keywords: iOCT · Super-Resolution · Image Quality Assessment.

^{*} Supported by King's Centre for Doctoral Studies - Centre for Doctoral Training in Surgical & Interventional Engineering, and an ERC Starting Grant [714562].
L. Da Cruz and C. Bergeles are equally contributing senior authors.

1 Introduction

Regenerative therapies are considered as promising treatments for retinal diseases such as Age-Related Macular Degeneration (AMD) [2] that cause blindness. These novel therapies, however, must be precisely delivered subretinally or intraretinally over prolonged periods of time. Retinal layers cannot be visualised under conventional vitreoretinal surgical protocols, i.e. with en face binocular biomicroscopy, which calls for the use of Intra-operative Optical Coherence Tomography (iOCT) to capture cross-sectional retinal images (Fig. 1a).

OCT is a non-invasive imaging modality, which uses light in the near-infrared spectral range to visualise tissue layer information. Pre-operatively acquired OCT images are of excellent quality taking advantage of spatiotemporal signal averaging albeit requiring a prohibitively long acquisition time for real-time visualisation of interventions. iOCT, on the other hand, allows real-time scanning but at the cost of inferior image quality making iOCT interpretation and layer discrimination challenging. Therefore, the interventional utility of iOCT images generated from current commercial systems is debatable.

In parallel with development of advanced iOCT hardware carried out internationally, we are pursuing computational enhancement of iOCT image quality to support the microprecise delivery of sight-restoring therapies within the retina. Previous studies on iOCT mainly focused on needle localisation [23], tool tracking [16] and anterior segment anatomies detection [17] but the enhancement of posterior eye segment visualisation in iOCT is still unaddressed.

Several classical software-based techniques have been proposed to reduce speckle noise in OCT images [1] and thus improve image quality. Averaging multiple registered B-scans of uncorrelated speckle noise, acquired at the same position on the retina [9, 19] and single B-scan filtering [1, 14] have positive denoising results. However, the need for prolonged scans and perfect alignment lessen the effectiveness of solutions for intra-operative applications and iOCT.

Deep learning approaches have been developed for image quality enhancement and super-resolution in natural images [7, 8, 11, 24]. In the medical imaging domain [4] proposed autoencoders built using convolutional layers for denoising, while in [22] a generative adversarial network (GAN) accompanied by Wasserstein distance and perceptual similarity was used to successfully denoise CT images. In [15] and [21], the authors focused on unsupervised Super-Resolution (SR) techniques while several works developed SR models for OCT image denoising [3, 10].

This paper reports underpinning work towards real-time iOCT Super Resolution. First, we perform No-Reference Image Quality Assessment (IQA) to investigate whether the available 3D iOCT scans are of better quality than low resolution (LR) 2D iOCT frames and thus if they can be used as high resolution (HR) images. Second, in order to provide weak supervision to the network, we apply non-rigid registration to create partially aligned pairs of LR and HR images. Then, we augment our initial paired dataset by exploiting the temporal continuity in 2D iOCT frames and use CycleGAN for quality transfer between LR and HR domain. We evaluate our pipeline through quantitative analysis us-

Fig. 1: (a): Left: Surgical microscope view. Right: iOCT B-scans. (b) From top to bottom: **V-iOCT**, **C-iOCT** and **C-pOCT**.

ing four different IQA metrics calculated on the super-resolved iOCT image. We also report qualitative data by capturing the individual preferences of 11 vitreoretinal surgeons with regards to original versus SR iOCT images.

2 Methods

In this section we present the Image Quality Assessment between the two iOCT image types, the partial alignment between LR and HR images and the generation of Super-resolved images using adversarial training with cycle consistency.

2.1 Datasets

The data used in this study originate from an internal dataset of intra-operative retinal surgery videos and OCT/iOCT scans from 66 patients acquired at Moorfields Eye Hospital, London, UK. The data was acquired in accordance with the Declaration of Helsinki (1983 Revision) and its ethical principles. To qualify for inclusion in our dataset, a “patient” should have all of the following three types of OCT data acquired by ZEISS devices (Carl Zeiss Meditec Inc., Dublin, CA):

1. Intra-operative OCT video frames (**V-iOCT**), recorded from RESCAN 700 integrated into the OPMI LUMERA 700 acquired during surgery.
2. Intra-operative OCT cube frames (**C-iOCT**), extracted from 512x1024x128 OCT cubes, acquired at the onset of surgery using RESCAN 700.
3. Preoperative OCT cube frames (**C-pOCT**), extracted from 512x1024x128 OCT cubes, acquired prior to the intervention using Cirrus 5000.

This led to a subset 18 data tuples/patients to be considered in this paper.

As a preprocessing step, we manually identified the frames of the three OCT sources that contain corresponding underlying anatomy given that the videos and the scans of the same patient visualise alike retinal structures. We ended up having 983 images per type (**V-iOCT**, **C-iOCT**, **C-pOCT**).

2.2 Image Quality Assessment

Even though arguably **C-pOCT** scans are the OCT images with the best quality (Fig. 1b), we investigated whether **C-iOCT** can be used as the HR domain to enhance the **V-iOCT** (LR). As both **C-iOCT** and **V-iOCT** were obtained under the same conditions (device, date, patient position), they represent a smaller HR/LR domain gap. Therefore, we explored methods in the context of No-Reference Image Quality Assessment to compare the quality of the two iOCT sources as in our case true HR reference images are not available.

First, we used a perceptual loss function [8] which measures the high level perceptual difference between input and target images through their distance in the feature space of Imagenet-pretrained Deep Convolutional Network [18]. Specifically, for each patient and for each OCT type, e.g., **V-iOCT**, we extracted the VGG-16 [20] **conv**₄₃ feature representations for every input **V-iOCT** image and we averaged them generating one averaged feature map for the **V-iOCT** images. We applied the same process for **C-iOCT** and **C-pOCT** images obtaining two more averaged feature maps per patient. Then, we calculated pairwise Euclidean distances between the averaged feature maps (Φ) of **V-iOCT** and **C-pOCT** as well as between **C-iOCT** and **C-pOCT**:

$$\ell_{feat} = \|\Phi_i - \Phi_p\|_2 \quad (1)$$

The per patient ℓ_{feat} values indicate the perceptual similarity of **V-iOCT** and **C-iOCT** relative to **C-pOCT**.

Furthermore, we used Frechet Inception Distance (FID)[6], a metric used to evaluate the performance of GANs [5]. For image generation tasks using GANs, FID can capture how similar two sets of images are by comparing the statistics of the distributions of feature representations extracted from the ImageNet-pretrained Inception-v3. In this context, we use FID to calculate the similarity of **V-iOCT-C-pOCT** and **C-iOCT-C-pOCT** datasets across the 18 patients:

$$FID = \|\mu_i - \mu_p\|^2 + Tr(\Sigma_i + \Sigma_p - 2(\Sigma_i \Sigma_p)^{1/2}) \quad (2)$$

where $X_i \sim N(\mu_i, \Sigma_i)$ and $X_p \sim N(\mu_p, \Sigma_p)$ are the activations of the first max pooling layer for intra-operative (**V-iOCT**, **C-iOCT**) and preoperative (**C-pOCT**) samples respectively. We chose the first max pooling layer 64-dimensional features of Inception-v3 as every patient contains at least 64 images per OCT type which are sufficient to calculate the statistics.

Finally, we considered Natural Image Quality Evaluator (NIQE) [13], one of the state-of-the-art No-Reference IQA metrics. We trained a custom NIQE model from the database of **C-pOCT** images, and we calculated the NIQE score for each **V-iOCT** and **C-iOCT** image.

Lower NIQE scores indicate better perceptual quality of the examined iOCT image with respect to **C-pOCT**.

The intuition behind the above methods is that if the values of ℓ_{feat} , FID and NIQE for **C-iOCT** images across the 18 patients are lower than the corresponding values of **V-iOCT**, then **C-iOCT** would be perceptually more similar to the

Fig. 2: Overview of the proposed augmentation approach.

C-pOCT which are the highest quality images of our dataset. Therefore, considering **C-iOCT**'s as the HR domain would be meaningful. IQA results, reported in 3.1, demonstrate that **C-iOCT** are perceptually closer to the **C-pOCT** and thus they will be used as HR domain throughout the next sections.

2.3 Registration

Despite the fact that **V-iOCT** and **C-iOCT** frames were both acquired intra-operatively and under the same conditions, there were misalignments between underlying structures due to patient motion. CycleGAN which is the SR model of our pipeline, has been used mainly for unpaired datasets. However, in our first experiments we noticed that unpaired **V-iOCT** and **C-iOCT** data led to inconsistent SR results, so we decided to partially register them aiming to provide weak supervision to the network. In order to create a valid set of paired **V-iOCT-C-iOCT** images, we applied an affine transformation as initialization and performed multi-resolution non-rigid registration using B-splines as parameterisation between the **V-iOCT** (fixed) and **C-iOCT** frames (moving).

Following registration, we applied Retinal Pigment Epithelium (RPE) layer delineation using a heuristic method in the resulting moved image. The method considers the RPE as the most hyper-reflective retinal layer and estimates RPE points by calculating the brightest pixels across the A-scans (columns) of the iOCT image. Then, based on the calculated RPE points, we defined the horizontal and vertical cropping points considering as useful part of the moved image the region between RPE and ILM. The cropping points defined the crop region for both moved (**C-iOCT**) and fixed (**V-iOCT**) frames leading to images of size 380x150.

2.4 Data Augmentation

We exploited the temporal continuity in **V-iOCT** frames to augment the training pairs (so far 537). Knowing the position of one of the already paired **V-iOCT** (reference), we examined which of the ± 5 neighbour frames capture relevant retinal tissue. We delineated the RPE layers in neighbour frames and calculated their Hausdorff Distance with respect to the reference's RPE. Registration with the

reference’s **C-iOCT** image was performed only for the neighbour frames where Hausdorff Distance was less than 15 (pixels). We then rigidly registered the selected frames to the reference’s **C-iOCT**. This data augmentation methodology (Fig. 2) led to a total of 1022 paired images. The source code of our methods will be available online at <https://github.com/RViMLab/OMIA2021-iOCT-Super-Resolution>.

2.5 Super Resolution with Cycle Consistency

To perform Super Resolution (SR), we adopted CycleGAN’s architecture [24], which is widely used in image-to-image translation tasks. Given two datasets of LR (**V-iOCT**) and HR (**C-iOCT**) images, CycleGAN can learn a mapping $G : LR \rightarrow HR$ such that the distribution of the generated images $G(LR)$ is indistinguishable from the distribution of HR images. Unlike to typical GANs, it also learns an inverse mapping from output to input $G : HR \rightarrow LR$ and uses cycle consistency to enforce mappings in both directions.

3 Results

In this section, we present the results obtained from IQA between **C-iOCT** and **V-iOCT** as well as the quantitative and the qualitative analysis that we conducted to validate our Super-Resolution pipeline.

3.1 Image Quality Assessment

Table 1 summarises IQA results. The reported values are the averaged results across 18 patients. The values of the three metrics ($FID, \ell_{feat}, NIQE$) are lower for **C-iOCT** which indicates better perceptual quality with respect to **C-pOCT**. Therefore, **C-iOCT** are closer to the preoperative data (**C-pOCT**), which implies that they contain more information than **V-iOCT** and thus they can be used as HR images. The statistical significances of the pairwise comparisons were assessed using paired t-test and all the p-values were $p < 0.001$.

Table 1: IQA results of **V-iOCT** and **C-iOCT** with respect to **C-pOCT**. Arrows indicate that lower means better.

Metric	V-iOCT	C-iOCT
FID ($\downarrow$)	3.77 ± 1.87	0.65 ± 0.22
$\ell_{feat}(\downarrow)$	597.57 ± 80.84	503.50 ± 66.72
NIQE ($\downarrow$)	26.29 ± 1.52	16.32 ± 2.06

Fig. 3: LR **V-iOCT** images and their Super-resolved **SR-iOCT**.

Fig. 4: From left to right: **V-iOCT**(LR), SR by CycleGAN, SR by Pix2Pix, **C-iOCT**(HR).

3.2 Quantitative Analysis

We performed quantitative analysis to validate the quality enhancement of the SR images (Fig. 3) with respect to **C-iOCT**. The analysis was based on the: **V-iOCT** (LR), **SR-iOCT**(CGAN) (SR) and **C-iOCT** (HR) test sets (204 images each). Three already analysed in Sec. 2.2 metrics (FID , ℓ_{feat} , $NIQE$) and the no-reference Global Contrast Factor (GCF) [12] metric were used.

First, we divided the test set to 3 subsets of 64 images each and we calculated the FID between **SR-iOCT** and **C-iOCT** as well as the FID between **V-iOCT** and **C-iOCT** and we averaged the results across the 3 subsets. ℓ_{feat} was computed between each **V-iOCT**, **SR-iOCT** and its corresponding **C-iOCT** frame while NIQE score was obtained per **V-iOCT**, **SR-iOCT** image after training a model on **C-iOCT** images. GCF, a reference-free metric which calculates the image contrast, was also computed per **V-iOCT**, **SR-iOCT**. The reported values (Table 2) are the averaged results across the 204 **V-iOCT** and **SR-iOCT** im-

Table 2: Quantitative analysis results. Arrows show if higher/lower means better.

Metric	V-iOCT	SR-iOCT(CGAN)	SR-iOCT(Pix)
FID ($\downarrow$)	3.18 $\pm$ 0.93	3.22 $\pm$ 0.39	3.26 $\pm$ 0.64
$\ell_{feat}(\downarrow)$	632.31 $\pm$ 74.29	602.55 $\pm$ 77.24	621.94 $\pm$ 72.99
NIQE ($\downarrow$)	32.01 $\pm$ 5.81	28.22 $\pm$ 4.74	39.07 $\pm$ 3.60
GCF ($\uparrow$)	7.85 $\pm$ 0.60	8.21 $\pm$ 0.58	6.49 $\pm$ 0.46

ages. To further validate our work, we compared our approach against Pix2Pix, **SR-iOCT(Pix)**, which is one of the state-of-the-art methodologies for image-to-image translation tasks. Pix2Pix is an implementation of the Conditional GAN where the generation of an image is conditional on a given image. It needs aligned pairs of LR and HR images, a requirement which is fulfilled to a great extent in our case because of the registration step. The quantitative and visual results of the method are presented in Table 2 and Fig. 4 respectively.

Our analysis aimed at investigating whether our pipeline can improve the quality of the LR OCT frames. From these results, **SR-iOCT(CGAN)** images, super-resolved images using CycleGAN, are perceptually more similar to the HR than the corresponding LR according to ℓ_{feat} . Furthermore, lower NIQE score and higher GCF values show that **SR-iOCT(CGAN)** has better perceptual quality and higher contrast respectively compared to **V-iOCT**. However, FID demonstrates that **V-iOCT** statistics are more similar to **C-iOCT**, which is probably because the metric uses lower layer features trying to address the problem of having limited samples for statistics calculation. On the other hand, **SR-iOCT(Pix)** images are worse than the LR images according to three out of four metrics (Table 2). Pix2Pix’s architecture does not include a Cycle Consistency Loss term and as a result structures from original sample may not appear in the reconstructed sample. We assessed the statistical significance of the results using paired t-test and the p-values for all the metrics except FID were $p < 0.001$. It is also worth mentioning that CycleGAN needs 0.029 (s) time to generate one super-resolved image of size 350x350 which is acceptable for the real-time requirement of our application.

3.3 Qualitative Analysis

Our survey included 30 pairs of **V-iOCT** and **SR-iOCT(CGAN)** frames (Fig. 3), randomly arranged to avoid order bias during the evaluation. We asked 11 retinal doctors/surgeons to select the image that they would prefer to observe in a clinical setting. The results indicate that clinicians marginally prefer the Super-resolved OCT obtained by our methodology, as 51% of the total experts’ answers are in favour of **SR-iOCT** instead of **V-iOCT**. Furthermore, it is worth noting that 2 out of the 11 clinicians strongly preferred the **SR-iOCT** (in 83.33% and 73.33% of the occasions), 3 were against **SR-iOCT** (less than 35% cases) and the rest fluctuated between 46% and 63%. In addition, in 5 out of 30 pairs, more than 73% of the experts tend to prefer **SR-iOCT** than the original **V-iOCT**.

Therefore, despite the fact that our quantitative metrics demonstrate that our pipeline can produce images of higher quality than the original **V-iOCT** frames, the qualitative analysis shows that most participants, in the majority of the cases were unable to choose between **V-iOCT** and **SR-iOCT**, showcasing the need for further work to deliver qualitatively superior iOCT images.

4 Discussion and Conclusions

This paper presents initial work on the Super-resolution of intra-operative Optical Coherence Tomography images obtained in real-time during various vitreoretinal surgical procedures.

First, we assessed the quality of **C-iOCT** and **V-iOCT** using three different No Reference Image Quality Assessment methods demonstrating that **C-iOCT** can be used as HR domain. Furthermore, we performed Non-rigid Registration between LR and HR images and we deployed an automatic way to augment the available training dataset. Finally, we used CycleGAN to apply Super-Resolution by quality transfer between LR and HR domains.

We evaluated our pipeline by performing both quantitative and qualitative analysis. In the former case, the obtained results showed that the proposed method can generate enhanced iOCT images, while the later analysis demonstrated that there is not a clear preference between original and Super-resolved images in the majority of the occasions.

To the best of our knowledge, this paper is the first which addresses the problem of Super-Resolution in intra-operative OCT images leveraging the available HR iOCT 3D scans. Further research into temporal alignment of iOCT frames must be carried out to maximise the potential of creating iOCT frames of truly superior image quality than what currently available.

References

1. Adler, D.C., Ko, T.H., Fujimoto, J.G.: Speckle reduction in optical coherence tomography images by use of a spatially adaptive wavelet filter. *Optics letters* **29**(24), 2878–2880 (2004)
2. da Cruz, L., Fynes, K., Georgiadis, O., Kerby, J., Luo, Y.H., Ahmado, A., Vernon, A., Daniels, J.T., Nommiste, B., Hasan, S.M., et al.: Phase 1 clinical study of an embryonic stem cell-derived retinal pigment epithelium patch in age-related macular degeneration. *Nature biotechnology* **36**(4), 328 (2018)
3. Devalla, S.K., Subramanian, G., Pham, T.H., Wang, X., Perera, S., Tun, T.A., Aung, T., Schmetterer, L., Thiéry, A.H., Girard, M.J.: A deep learning approach to denoise optical coherence tomography images of the optic nerve head. *Scientific reports* **9**(1), 1–13 (2019)
4. Gondara, L.: Medical image denoising using convolutional denoising autoencoders. In: 2016 IEEE 16th International Conference on Data Mining Workshops (ICDMW). pp. 241–246. IEEE (2016)
5. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y.: Generative adversarial nets. *Advances in neural information processing systems* **27** (2014)

6. Heusel, M., Ramsauer, H., Unterthiner, T., Nessler, B., Hochreiter, S.: Gans trained by a two time-scale update rule converge to a local nash equilibrium. *arXiv preprint arXiv:1706.08500* (2017)
7. Isola, P., Zhu, J.Y., Zhou, T., Efros, A.A.: Image-to-image translation with conditional adversarial networks. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. pp. 1125–1134 (2017)
8. Johnson, J., Alahi, A., Fei-Fei, L.: Perceptual losses for real-time style transfer and super-resolution. In: *European conference on computer vision*. pp. 694–711. Springer (2016)
9. Jørgensen, T.M., Thomadsen, J., Christensen, U., Soliman, W., Sander, B.A.: Enhancing the signal-to-noise ratio in ophthalmic optical coherence tomography by image registration—method and clinical examples. *Journal of biomedical optics* **12**(4), 041208 (2007)
10. Lazaridis, G., Lorenzi, M., Ourselin, S., Garway-Heath, D.: Improving statistical power of glaucoma clinical trials using an ensemble of cyclical generative adversarial networks. *Medical Image Analysis* **68**, 101906 (2021)
11. Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., et al.: Photo-realistic single image super-resolution using a generative adversarial network. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. pp. 4681–4690 (2017)
12. Matkovic, K., Neumann, L., Neumann, A., Psik, T., Purgathofer, W.: Global contrast factor—a new approach to image contrast. *Computational Aesthetics* **2005**(159-168), 1 (2005)
13. Mittal, A., Soundararajan, R., Bovik, A.C.: Making a “completely blind” image quality analyzer. *IEEE Signal processing letters* **20**(3), 209–212 (2012)
14. Ozcan, A., Bilenca, A., Desjardins, A.E., Bouma, B.E., Tearney, G.J.: Speckle reduction in optical coherence tomography images using digital filtering. *JOSA A* **24**(7), 1901–1910 (2007)
15. Ravi, D., Szczotka, A.B., Pereira, S.P., Vercauteren, T.: Adversarial training with cycle consistency for unsupervised super-resolution in endomicroscopy. *Medical image analysis* **53**, 123–131 (2019)
16. Rieke, N., Tan, D.J., Alsheakhali, M., Tombari, F., di San Filippo, C.A., Belagiannis, V., Eslami, A., Navab, N.: Surgical tool tracking and pose estimation in retinal microsurgery. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. pp. 266–273. Springer (2015)
17. Roodaki, H., Grimm, M., Navab, N., Eslami, A.: Real-time scene understanding in ophthalmic anterior segment oct images. *Investigative Ophthalmology & Visual Science* **60**(11), PB095–PB095 (2019)
18. Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., et al.: Imagenet large scale visual recognition challenge. *International journal of computer vision* **115**(3), 211–252 (2015)
19. Sander, B., Larsen, M., Thrane, L., Hougaard, J.L., Jørgensen, T.M.: Enhanced optical coherence tomography imaging by multiple scan averaging. *British Journal of Ophthalmology* **89**(2), 207–212 (2005)
20. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556* (2014)
21. Wolterink, J.M., Dinkla, A.M., Savenije, M.H., Seevinck, P.R., van den Berg, C.A., Išgum, I.: Deep mr to ct synthesis using unpaired data. In: *International workshop on simulation and synthesis in medical imaging*. pp. 14–23. Springer (2017)

22. Yang, Q., Yan, P., Zhang, Y., Yu, H., Shi, Y., Mou, X., Kalra, M.K., Zhang, Y., Sun, L., Wang, G.: Low-dose ct image denoising using a generative adversarial network with wasserstein distance and perceptual loss. *IEEE transactions on medical imaging* **37**(6), 1348–1357 (2018)
23. Zhou, M., Huang, K., Eslami, A., Roodaki, H., Zapp, D., Maier, M., Lohmann, C.P., Knoll, A., Nasser, M.A.: Precision needle tip localization using optical coherence tomography images for subretinal injection. In: 2018 IEEE International Conference on Robotics and Automation (ICRA). pp. 4033–4040. IEEE (2018)
24. Zhu, J.Y., Park, T., Isola, P., Efros, A.A.: Unpaired image-to-image translation using cycle-consistent adversarial networks. In: Proceedings of the IEEE international conference on computer vision. pp. 2223–2232 (2017)