Convolutional neural network-based regression for biomarker estimation in corneal endothelium microscopy images

Vigueras Guillén, Juan Pedro; van Rooij, Jeroen G.J.; Lemij, Hans G.; Vermeer, Koen; van Vliet, Lucas

DOI
10.1109/EMBC.2019.8857201

Publication date
2019

Document Version
Final published version

Published in
2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2019

Convolutional neural network-based regression for biomarker estimation in corneal endothelium microscopy images

Juan P. Vigueras-Guillén*[1,3], Jeroen van Rooij[2], Hans G. Lemij[2], Koenraad A. Vermeer[1], and Lucas J. van Vliet[3]

***Abstract*— The morphometric parameters of the corneal endothelium – cell density (ECD), cell size variation (CV), and hexagonality (HEX) – provide clinically relevant information about the cornea. To estimate these parameters, the endothelium is commonly imaged with a non-contact specular microscope and cell segmentation is performed to these images. In previous work, we have developed several methods that, combined, can perform an automated estimation of the parameters: the inference of the cell edges, the detection of the region of interest (ROI), a post-processing method that combines both images (edges and ROI), and a refinement method that removes false edges. In this work, we first explore the possibility of using a CNN-based regressor to directly infer the parameters from the edge images, simplifying the framework. We use a dataset of 738 images coming from a study related to the implantation of a Baerveldt glaucoma device and a standard clinical care regarding DSAEK corneal transplantation, both from the Rotterdam Eye Hospital and both containing images of unhealthy endotheliums. This large dataset allows us to build a large training set that makes this approach feasible. We achieved a mean absolute percentage error (MAPE) of 4.32% for ECD, 7.07% for CV, and 11.74% for HEX. These results, while promising, do not outperform our previous work. In a second experiment, we explore the use of the CNN-based regressor to improve the post-processing method of our previous approach in order to adapt it to the specifics of each image. Our results showed no clear benefit and proved that our previous post-processing is already highly reliable and robust.**

I. INTRODUCTION

In the last decade, Convolutional Neural Networks (CNNs) have become a very popular approach for image classification [1], [2], [3] and image segmentation, either via pixel classification or semantic segmentation [4]. Regression problems in images have also been solved with CNNs by adding fully connected layers with linear activations functions at the end of the network [5]. Overall, CNNs have led to a series of breakthroughs in the recent years, being employed with high success in many disciplines.

In corneal endothelium images, it is clinically relevant to determine the morphometric parameters of the cells in order to assess the quality and health status of the cornea. A quantitative analysis can provide information on endothelial cell density (ECD, reported as the number of cells per mm^2), cell size variation (CV, expressed by the coefficient of variation in cell size), and hexagonality (HEX, defined as the percentage of 6-sided cells). These image-based biomarkers can be easily estimated if the cell boundaries are identified in the image, therefore image segmentation has been the approach most commonly employed for solving this task [6], [7], [8]. In previous work, we achieved state-of-the-art results by employing a CNN U-net to segment the cell boundaries [9], but in order to make the process completely automated we developed a dense U-net to infer the area in the image (region of interest, ROI) in which the biomarkers could be reliably estimated [10]. Furthermore, it is not a trivial task to combine both CNN output images (the edge probability image and the ROI image) and transform them into the final binary edge image to perform the biomarker estimation. Indeed, any small error or discontinuity in an edge could cause a significant error in the resulting estimates. This step was called *post-processing* in our previous work [9].

In this work, we explore the possibility of using a CNN-based regressor, named *DenseNet-R35*, to directly estimate the biomarkers from the edge probability images, substituting the inference of the ROI, the post-processing, and refinement. In addition, we explore the use of DenseNet-R35 to improve the post-processing itself, and we compare it against a simpler approach based on Fourier analysis.

II. MATERIALS AND PREVIOUS WORK

The images used in this work are part of one ongoing study regarding the implantation of a Baerveldt glaucoma drainage device and one retrospective standard clinical care regarding the visual function after corneal transplantation (Descemet Stripping Automated Endothelial Keratoplasty, DSAEK), both from the Rotterdam Eye Hospital. For the Baerveldt study, images were retrieved before the surgical implantation and 3, 6, 12, and 24 months after surgery, in both the central and the temporal superior (TS) cornea. Here, we used 400 images from 100 patients. For the DSAEK group, 338 images from 41 patients were acquired 1, 3, 6, and 12 months after surgery, only from the central cornea. In both cases, images were obtained with a non-contact specular microscope Topcon SP-1P (Topcon Co., Japan). Data was collected in accordance with the tenets of the Declaration of Helsinki. Signed informed consent was obtained from all

*This work was supported by the Dutch Organization for Health Research and Health Care Innovation (ZonMw) under grants 842005004 and 842005007, and by the Combined Ophthalmic Research Rotterdam (CORR, project number 2.1.0). *Asterisk indicates corresponding author.*

[1]J.P. Vigueras-Guillén and K.A. Vermeer are with the Rotterdam Ophthalmic Institute, 3011BH Rotterdam, The Netherlands. J.ViguerasGuillen@oogziekenhuis.nl; K.Vermeer@oogziekenhuis.nl

[2]J. van Rooij and H.G. Lemij are with the Rotterdam Eye Hospital, 3011BH Rotterdam, The Netherlands. J.vanRooij@oogziekenhuis.nl; H.Lemij@oogziekenhuis.nl.

[3]J.P. Vigueras-Guillén and L.J. van Vliet are with Delft University of Technology, Department of Imaging Physics, 2611CN Delft, The Netherlands. J.P.ViguerasGuillen@tudelft.nl; L.J.vanVliet@tudelft.nl.

978-1-5386-1311-5/19/$31.00 ©2019 IEEE

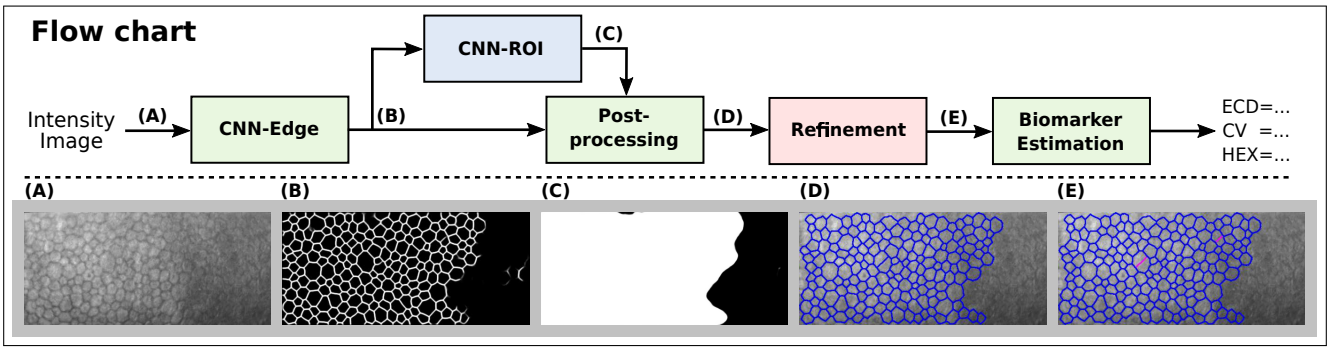

Fig. 1. Flow chart of our previous framework for biomarker estimation. CNN-Edge infers the edge probability images, CNN-ROI estimates the region of interest, the post-processing combines both images to provide a binary edge image, and the refinement method removes potential false edges (magenta lines in E) before estimating the biomarkers. The different images at each step (A-E) are displayed below the dashed line. For (D) and (E), the results are binary images, depicted in blue (superimposed to the intensity images for display purposes).

participants. For the Baerveldt study, approval was obtained from the Medical Ethical Committee of the Erasmus Medical Center (Rotterdam, The Netherlands), with trial registration NTR4946.

The images were saved as grayscale images of 240×528 pixels, covering an area of 0.25 mm × 0.55 mm. Images from the Baerveldt study generally show a good contrast between cell edges and cell bodies. The main challenge is that large fractions of the TS images are out of focus due to the curvature of the cornea at that position (as in Figure 1-A). In contrast, images from the DSAEK group are more prone to appear blurred, with low contrast, and slightly out of focus due to the optical distortions produced by the graft-recipient interface and the rough surface of the graft. An expert created the gold standard for each image by manually delineating the cell boundaries.

Previously, we presented a fully automated framework to estimate the endothelial biomarkers from specular microscopy images (Figure 1). This framework is subdivided into four different methods: a CNN based on a Dense U-net to segment endothelial cells using an intensity image as input and providing an edge probability image as output (CNN-Edge) [9], [10]; a CNN based on a Dense U-net to segment the region of interest where cells are correctly detected using the edge probability image as input (CNN-ROI) [10]; a post-processing method based on Fourier analysis and watershed that combines both output images and yields the binary edge image [9], [11]; and a machine learning approach based on Support Vector Machines (SVM) that removes potential false edges from the binary edge image [8]. The biomarkers could then be estimated from the final binary image.

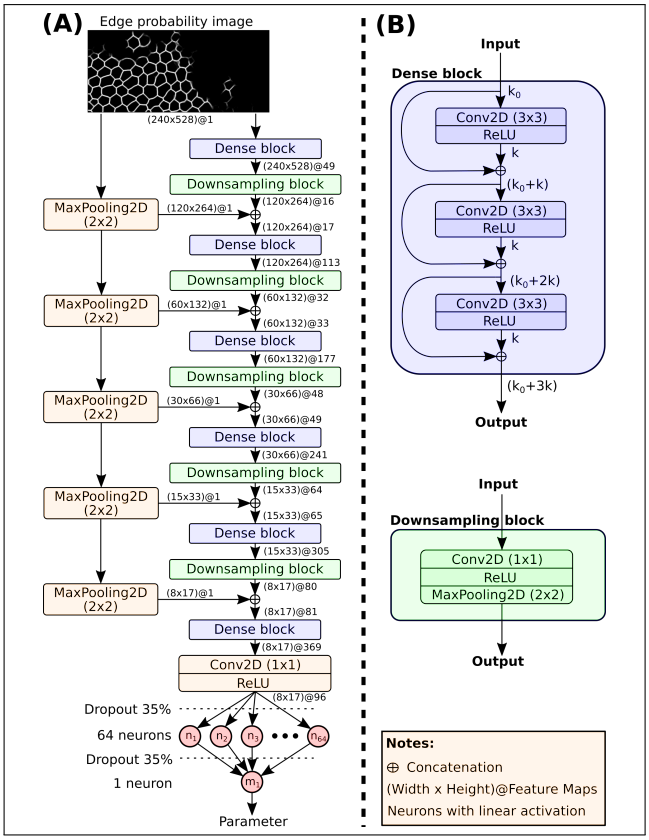

Fig. 2. (A) A schematic overview of the proposed DenseNet-R35 ; and (B) a schematic overview of one dense block and one downsampling block.

III. METHODS

A. Biomarker estimation from edge images

In the first approach, we aim to use the DenseNet-R35 (Figure 2) after the CNN-Edge (Figure 1) in order to directly infer the biomarkers from the edge probability image. Ideally, it would be preferred to use DenseNet-R35 as a substitution of the whole previous framework, but that would require a much larger annotated dataset. This possibility was tested, but failed to obtain reliable results. As described below, we can perform more data augmentation on the edge probability images, thereby making this approach feasible.

The proposed DenseNet-R35 (Figure 2) borrows the concept of densely connected convolutional layers from Huang et al. [3]. In dense blocks, a layer receives the feature maps of all preceding layers, which improves the direct information flow, helping back-propagating gradients and thereby facilitating better network learning during training. Huang et al. [3] used dozens of convolutional layers in each dense block, which creates bottlenecks due to the excessive number of piled up layers. To solve this, they added one 1×1

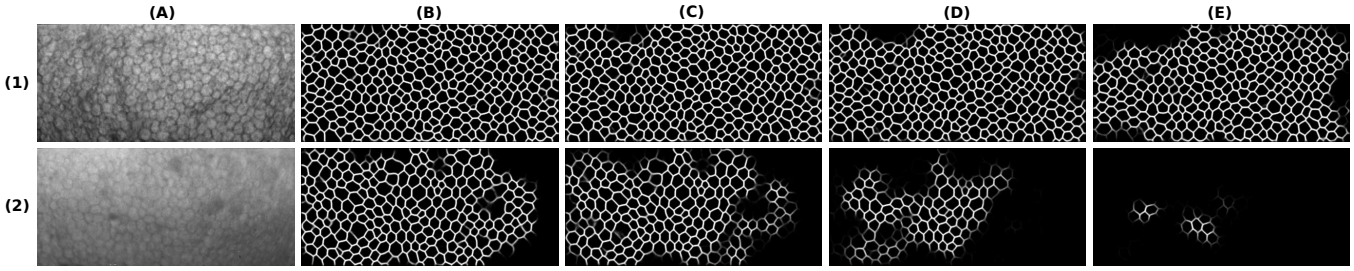

Fig. 3. Representative output images of our previous Dense U-net (CNN-Edge) for four configurations based on the dropout rate of the dropout layers: 5% (B), 20% (C), 35% (D), and 50% (E), given (A) as input. Image (1) belongs to the Baerveldt study, whereas image (2) comes from the DSAEK study.

convolutional layer before each 3×3 convolutional layer in order to reduce the number feature maps. Here, we use only three 3×3 convolutional layers in each dense block, hence the 1×1 convolutional layers are not needed. Each layer in a dense block provides k feature maps, where k is the growth rate parameter. We set $k = 16$ in the first dense block, adding 16 more features at each new resolution step. After each dense block, the feature maps are reduced to k by a 1×1 convolutional layer in the downsampling block. Furthermore, we introduce the edge probability image, properly scaled, at the beginning of each resolution step. Finally, the dense layer of neurons and the output neuron employ a linear activation function. The chosen loss function is the mean squared error (MSE), and the mean absolute percentage error (MAPE) is computed for analysis.

Another key aspect is the absence of batch normalization or dropout layers in the convolutional part, which either increased the MSE loss or prevented the network to converge, respectively. In contrast, the use of dropout in the layer of densely connected neurons was highly beneficial, obtaining the best performance for a dropout rate of 35%. This architecture was computationally efficient, avoiding bottlenecks during network training. It has a receptive field of 379×379 (approx. $(\text{height} + \text{width})/2$), which showed to be optimal as either adding or removing a convolutional layer in each block was detrimental.

Data augmentation was a critical aspect in our training. Horizontal and vertical flipping multiplied the available data by four. Rotation had the inconvenience of losing the corners of the image once rotated, and thus it was evaluated. We obtained that uniform sampling of the rotation angle between $\pm 35^\circ$ yielded the best performance. Elastic deformations and translation were detrimental once rotation was employed. Finally, we further increased the training data by employing, for the same intensity image, slightly different versions of the edge probability images, obtained as discussed in our previous work [10]. Briefly explained, the Dense U-net (CNN-Edge, Figure 1) could be tuned to provide more conservative edge images by simply increasing the dropout rate, thereby not inferring potentially spurious cell boundaries in highly blurred and/or noisy areas (Figure 3). For images with high contrast (Figure 3-1A), the resulting edge images at different dropout rates did not show major differences unless the dropout was largely increased. However, for images with low contrast (Figure 3-2A), larger areas of the images were rapidly not inferred after just small increments of the dropout rate. We argued that the biomarkers estimated from those images should, in principle, be very similar if enough cells were detected. The DenseNet-R35 performance improved drastically as we were including more images built with larger dropout rates, reaching the minimum test loss when images up to a dropout rate of 35% were used. Therefore, they were added to the training with the same labels and were named *twin sets*. Considering all augmentation, the original 738 images could be converted to 1.6 million images.

We employed a 10-fold cross-validation, and the same fold order was applied to all twin sets. The validation set contained images from all twin sets, but the test set only contained the original dataset so that it could be compared to our previous framework. The original dataset (here and in our previous work) was obtained with a dropout rate of 5%.

B. Improvement of the post-processing method

In a second approach, the DenseNet-R35 was used to improve the post-processing method from our previous framework. The post-processing involved estimating the most common cell size in the edge probability images by using Fourier analysis ($l = 1/f^*$, being f^* the characteristic frequency), then smoothing the edge probability images with a Gaussian filter whose standard deviation (SD) is $\sigma = k_\sigma/(\alpha f^*)$, and finally applying the classic watershed [12], which does not require any parameter. The scaling factor of the sigma, $k_\sigma = 0.20$, was estimated in previous work [11], which assumed $\alpha = 1$.

The idea of estimating the cell size by Fourier analysis was proposed after observing that the 2D Fourier Transform (FT) of an endothelium image shows a distinctive ring related to the regular hexagonal patterns of the cells [13]. If the 2D FT was applied to the edge probability image instead, we got rid of the noisy patterns of the intensity image that could obscure the ring in Fourier domain. By simply computing the radial mean of the magnitude of the 2D FT [11], we could estimate the peak of the ring, f^* (Figure 4). In previous work [11], we proposed a model based on an exponential and a Gaussian to fit the radial mean, thereby estimating the ring size with the fitted Gaussian mean. This method is highly reliable to estimate the ECD of the image (Figure 5-A), obtaining a high linear correlation between the estimated f^* and the true ECD (Pearson correlation coefficient $r = 0.965$, $p < 0.001$).

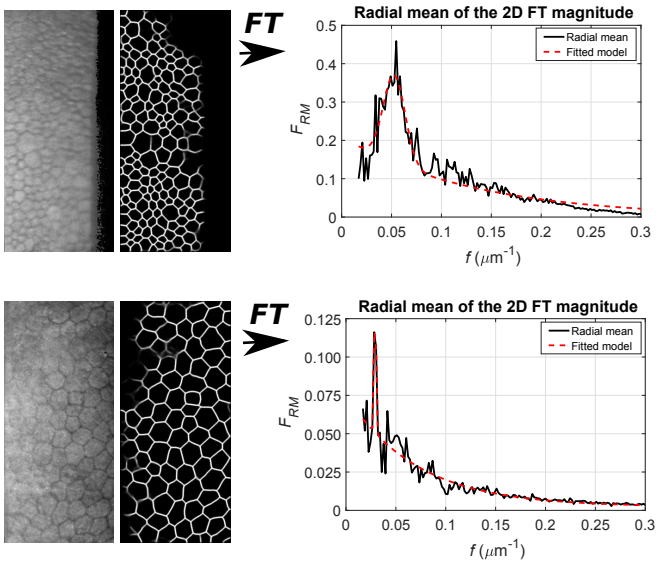

Fig. 4. The radial mean of the magnitude of the 2D FT (in black) of the probability edge image. A fitted model based on a Gaussian and an exponential (red) is used to estimate the peak (Gaussian mean, referred as characteristic frequency f^*) and Gaussian standard deviation.

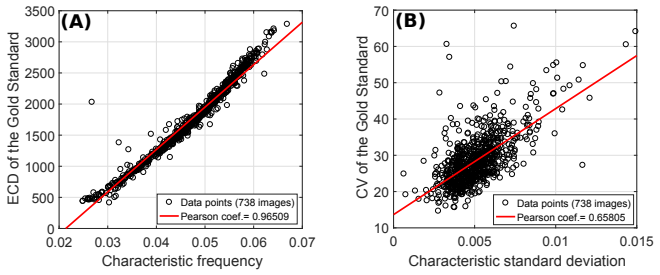

Fig. 5. (A) Relation between the estimated characteristic frequency in Fourier domain and the true ECD; and (B) the relation between the estimated characteristic SD and the true CV. Each circle is an image and the red line is the first degree polynomial that best fits the data (in a least-squares sense).

The main benefit of our post-processing step is the detection of faded, true edges in the edge probability images. Since watershed requires a single local minimum per cell in order to yield a good segmentation, the former smoothing of the edge image is crucial. Given the robustness of the estimation of f^*, our post-processing is simple yet very reliable. However, two flaws are observed in this approach: it does not take into account the (1) variation in cell size and (2) the blurriness of the edge probability image. Indeed, an image with high variation in cell size requires to reduce the filter size in order to avoid merging nearby edges and thereby missing small cells (Figure 4-top), whereas edge images with some blurred areas and regular cell size would benefit from a larger smoothing filter (Figure 4-bottom). To correct this, we define a scaling factor, α, to f^*.

In summary, our goal is to employ DenseNet-R35 to determine the best α for each image. We hypothesize that our network would be able to learn the characteristics of each image and therefore able to adapt the post-processing. For this, we estimated the best α for each image (twin sets included). Specifically, we computed two values from the binary edge output images: the total number of cells detected (n_{total}) and the number of cells correctly segmented (n_{corr}); as well as one value from the gold standard: the number of real cells (n_{real}). Precision, $p = n_{corr}/n_{total}$, and recall, $r = n_{corr}/n_{real}$, were computed and combined into the F-measure, $F = 2pr/(p + r)$. We performed this evaluation for values of α between 0.7 and 1.3 in steps of 0.05, being the best α the one that provides the largest F-measure. Therefore, DenseNet-R35 would have an edge probability image as input and its best α as target. The same 10-fold cross-validation was used.

C. Improvement of the post-processing by Fourier Analysis

The Fourier Analysis also showed a weak correlation between the fitted Gaussian SD (named 'characteristic SD') and the biomarker CV (Figure 5-B, Pearson correlation coefficient $r = 0.658$, $p < 0.0001$). Indeed, if the endothelium exhibits a large variation in cell size, many frequency components appear in the FT (Figure 4-top), whereas if the cells are rather uniform in size, only a very narrow range of frequencies are enhanced (Figure 4-bottom). Therefore, we consider the use of the characteristic SD, which seems to roughly encode information about the cell size variation, to estimate the best scaling factor α in our post-processing. For this, we employ a sigmoid model, defined as

$$\alpha = \frac{a}{1 + \exp(-b \cdot \mathrm{SD} + c)} + d, \qquad (1)$$

where a, b, c, and d are the parameters of the model. We set $d = 1 - a/2$ to center the sigmoid at $\alpha = 1$. The remaining parameters are estimated by employing a grid search and the F-measure defined above.

IV. RESULTS

A. Experiment for biomarker estimation

The training did not show over-fitting and converged rapidly. The error produced in each biomarker was significantly different, being HEX the most complicated parameter to predict (Figure 6). In the test set, the MAPE for the DenseNet-R35 was 4.32% for ECD, 7.07% for CV, and 11.74% for HEX, whereas our previous framework, with post-processing and solving for oversegmentation, achieved a MAPE of 2.34% for ECD, 5.56% for CV, and 6.08% for HEX (Figure 7). The non-parametric Wilcoxon signed rank test on the MAPE distributions revealed a statistically significant difference ($p < 0.0001$, all biomarkers), thereby suggesting that the error in our previous framework was significantly smaller.

The endothelium of a healthy adult usually shows an ECD around 2500 cells/mm^2 (or higher), a CV around 30% (or lower), and a HEX around 60% (or higher) [14]. It is clinically more relevant to assess an accurate estimation in images with low ECD or HEX, or high CV. In this respect, Figure 7-A shows an ECD error with similar magnitude for images with low and high ECD in both methods (within ± 100 cells/mm^2, slightly higher for the DenseNet-R35), with the

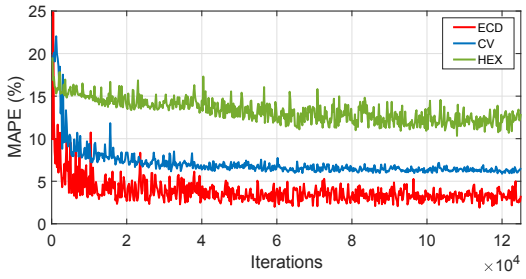

Fig. 6. MAPE in the validation set (fold 1) during training.

exception of some outliers (around 5% of images). While that error can be considered high for cases with very low ECD, both methods present a robust, stable inference of the ECD along different types of images, which means they are not affected by either a high or a low number of detectable cells. In contrast, the CV error is large with respect to the reference CV (Figure 7-B), and there is a small CV overestimation in images with low CV and an underestimation in images with high CV (same with HEX). Regression analysis determined that the errors were significantly correlated with the reference parameter ($p < 0.001$ for all biomarkers in both methods), yielding a slope of -0.03, 0.04, and 0.29 for ECD, CV, and HEX (respectively) in DenseNet-R35, and -0.01, 0.07, and 0.15 for ECD, CV, and HEX in our old framework. Thus, the correlation was very small for ECD and CV, but not for HEX. In previous work, these problems were the results of just a few minor errors in the segmentation since CV and HEX are biomarkers highly sensitive to segmentation errors. In this sense, DenseNet-R35 cannot overcome this problem.

B. Experiment for post-processing (α estimation)

The histogram distribution of the true α showed a Gaussian shape (Figure 8-B), which suggested that, for the majority of images, an α close to 1 yields the best segmentation. However, DenseNet-R35 was unable to detect the pattern in the image that would determine the best α and instead estimated all values around $\alpha \approx 1$ (Figure 8).

C. Experiment for post-processing (α estimation) by Fourier Analysis

The grid search provided the largest F-measure $= 0.947$ for the parameters $a = 0.4$, $b = 300$, and $c = 1.8$. The sigmoid resembled a linear response for an SD of 0–0.01 with an α range 0.85–1.10. In contrast, an F-measure $= 0.949$ was obtained if $\alpha = 1$ was employed in the post-processing method. Therefore, it was counterproductive to employ the characteristic SD for the inference of α.

V. DISCUSSION

The use of a CNN regressor to directly infer the biomarkers from edge probability images is a simple and fast method, yielding good results considering the complexity of the dataset under study, but it does not outperform our previous method [10]. In the latter, the selection of the ROI, the post-processing, and the subsequent refinement (Figure 1) were designed to detect, discard, and/or minimize any mistake in the edge probability images yielded by the CNN-Edge. The substitution of those three methods by a single CNN-based regressor would simplify this task, and the main challenge was to build a large annotated dataset (in the order of millions of images). In this respect, the creation of twin sets was a simple way to increase the data. Different from other types of data augmentation, twin sets did not provide unrealistic cases, and they had enough differences to make a profitable contribution to the training. However, due to the time cost of obtaining the true biomarkers from each image in the twin sets, we assumed they had the same parameter estimates as their respective original edge images. We believed this might be the major cause of our modest results, which also can explain why HEX had a considerably larger error. Indeed, since the twin sets show less detectable cells (Figure3-C,D), the real HEX in those images could be significantly different, whereas ECD would be barely affected.

Once we concluded that DenseNet-R35 could not outperform our previous method, we considered its use to improve it instead. As discussed in Section III-B, our post-processing method does not consider the blurriness of the image or its variation in cell size. Although the latter could be roughly inferred by Fourier analysis (Figure 5-B), it had a significant estimation error. The experiment in Section IV-C indicated that the inference of α based on the characteristic SD from Fourier analysis degraded the performance, therefore suggesting that blurriness might be the major factor to consider. Subsequently, we tested the use of DenseNet-R35 to infer the scaling factor α, thereby allowing the network to find the relevant features directly from the edge images. Different from the first experiment, we computed the target α for each independent image in the twin sets, thus avoiding any bias. However, DenseNet-R35 could not converge towards a reasonable solution. It is indubitable that this problem implies a larger complexity. For instance, blurriness might appear only in a fraction of the image while having high intensity contrast in the rest, and a CNN should be able to find the appropriate α for that small blurred portion while overlooking the characteristics of the rest of the image. Probably, a dataset of dozens of millions of images would be necessary to solve this problem.

The lack of convergence and similar behavior as in Figure 8-A was also observed if we attempted to estimate the biomarkers from the intensity images directly. Even with the use of more data augmentation (adding blurriness or noise), the network could not perform regression from the intensity images.

VI. CONCLUSIONS

This study has shown that our CNN-based regressor, which directly infers corneal endothelium biomarkers from the edge probability images (created by our previous work), does not provide a better solution than aiming for the binary segmentation instead. Based on our experiments, the amount of annotated data that would be required to provide similar accuracy seems unreasonable. Indeed, the augmentation of

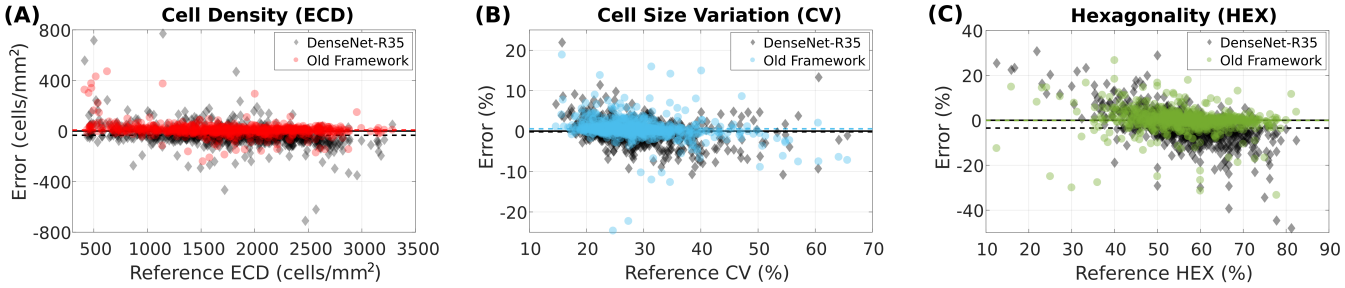

Fig. 7. Error of the biomarkers estimates – ECD (A), CV (B), and HEX (C) –, for the DenseNet-R35 (black) and our old framework (colored). The x-axis indicates the value for the gold standard, and the y-axis indicates the error computed as the difference between the proposed estimates and the gold standard. Each point corresponds to one image in the dataset (738 images in total). Dashed lines indicate the average error.

the data up to 1.6 million of annotated images were not enough to outperform our previous work. While it is possible that the limitation of the proposed data augmentation might have jeopardized the success, it is not clear whether a flawless augmentation would have made a large impact. Further work needs to be done to this respect. Nonetheless, considering that these biomarkers are defined based on the shape of the cells, it is reasonable to conclude that cell segmentation is still the preferable approach.

This study has also shown that the post-processing method of our previous work, which uses the average cell size estimated by Fourier analysis to smooth the edge images such that a subsequent watershed provides the proper segmentation, is simple and yet very robust. Our attempts to improve the post-processing by using more information from the Fourier analysis (the characteristic SD, which encodes the variation in cell size) have been unfruitful. We believe this is because the correlation between the characteristic SD and the cell size variation is not strong enough and the blurriness of the edge images should also be considered. This conclusion made us hypothesize that our CNN-based regressor could be satisfactorily employed for this task, as it could encode both factors (blurriness and cell size variation) to adapt the smoothing filter of the post-processing. However, our experiments were unsuccessful, suggesting that these features might be too subtle to identify. A possible alternative could be to adapt α to a local window instead of inferring a global α for each image.

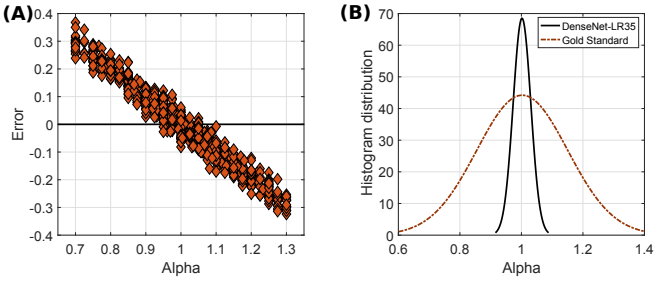

Fig. 8. (A) Error of the estimation of alpha (each point correspond to one image); and (B) distributions of the histograms of both sets, the estimated alphas by DenseNet-R35 and the true alphas (Gold Standard).

References

[1] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems 25*, 2012, pp. 1097–1105.

[2] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778.

[3] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *30th IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 2261–2269.

[4] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, ser. LNCS, 2015, vol. 9351, pp. 234–241.

[5] T. Shimobaba, T. Kakue, and T. Ito, "Convolutional neural network-based regression for depth prediction in digital holography," in *27th IEEE International Symposium on Industrial Electronics (ISIE)*, 2018, pp. 1323–1326.

[6] F. Scarpa and A. Ruggeri, "Development of a reliable automated algorithm for the morphometric analysis of human corneal endothelium," *Cornea*, vol. 35, no. 9, pp. 1222–1228, 2016.

[7] B. Selig, K. A. Vermeer, B. Rieger, T. Hillenaar, and C. L. Luengo Hendriks, "Fully automatic evaluation of the corneal endothelium from in vivo confocal microscopy," *BMC Medical Imaging*, vol. 15:13, 2015.

[8] J. P. Vigueras-Guillén, E.-R. Andrinopoulou, A. Engel, H. G. Lemij, J. van Rooij, K. A. Vermeer, and L. J. van Vliet, "Corneal endothelial cell segmentation by classifier-driven merging of oversegmented images," *IEEE Transactions on Medical Imaging*, vol. 37, no. 10, pp. 2278–2289, 2018.

[9] J. P. Vigueras-Guillén, B. Sari, S. F. Goes, H. G. Lemij, J. van Rooij, K. A. Vermeer, and L. J. van Vliet, "Fully convolutional architecture vs sliding-window CNN for corneal endothelium cell segmentation," *BMC Biomedical Engineering*, vol. 1:4, 2019.

[10] J. P. Vigueras-Guillén, H. G. Lemij, J. van Rooij, K. A. Vermeer, and L. J. van Vliet, "Automatic detection of the region of interest in corneal endothelium images using dense convolutional neural networks," in *Proc. SPIE, Medical Imaging 2019: Image Processing*, vol. 10949, no. 1094931, 2019.

[11] J. P. Vigueras-Guillén, A. Engel, H. G. Lemij, J. van Rooij, K. A. Vermeer, and L. J. van Vliet, "Improved accuracy and robustness of a corneal endothelial cell segmentation method based on merging superpixels," in *15th International Conference Image Analysis and Recognition (ICIAR)*, ser. LNCS, vol. 10882, 2018, pp. 631–638.

[12] S. Beucher and F. Meyer, "The morphological approach to segmentation: the watershed transformation. Mathematical morphology in image processing," *Optical Engineering*, vol. 34, pp. 433–481, 1993.

[13] M. Foracchia and A. Ruggeri, "Automatic estimation of endothelium cell density in donor corneas by means of Fourier analysis," *Medical and biological engineering and computing*, vol. 42, no. 5, pp. 725–731, 2004.

[14] B. E. McCarey, H. F. Edelhauser, and M. J. Lynn, "Review of corneal endothelial specular microscopy for FDA clinical trials of refractive procedures, surgical devices and new intraocular drugs and solutions," *Cornea*, vol. 27, no. 1, pp. 1–16, 2008.