Self-supervised deep-learning segmentation of corneal endothelium specular microscopy images

Sergio Sanchez[1*], Kevin Mendoza[1], Fernando Quintero[1], Angelica M. Prada[2,3,4], Alejandro Tello[2,3,4,5], Virgilio Galvis[2,3,4], Lenny A. Romero[6], Andres G. Marrugo[1]

[1*]Facultad de Ingeniería, Universidad Tecnológica de Bolívar, Cartagena, Colombia.
[2]Centro Oftalmológico Virgilio Galvis, Floridablanca, Colombia.
[3]Fundación Oftalmológica de Santander FOSCAL, Floridablanca, Colombia.
[4]Facultad de Salud, Universidad Autónoma de Bucaramanga UNAB, Bucaramanga, Colombia.
[5]Facultad de Salud, Universidad Industrial de Santander UIS, Bucaramanga, Colombia.
[6]Facultad de Ciencias Básicas, Universidad Tecnológica de Bolívar, Cartagena, Colombia.

*Corresponding author(s). E-mail(s): ssanchez@utb.edu.co;

Abstract

Computerized medical evaluation of the corneal endothelium is challenging because it requires costly equipment and specialized personnel, not to mention that conventional techniques require manual annotations that are difficult to acquire. This study aims to obtain reliable segmentations without requiring large data sets labeled by expert personnel. To address this problem, we use the Barlow Twins approach to pre-train the encoder of a UNet model in an unsupervised manner. Then, with few labeled data, we train the segmentation. Encouraging results show that it is possible to address the challenge of limited data availability using self-supervised learning. This model achieved a precision of 86%, obtaining a satisfactory performance. Using many images to learn good representations and a few labeled images to learn the semantic segmentation task is feasible.

Keywords: Self-supervised, deep learning, segmentation, corneal endothelium

1 Introduction

The corneal endothelium is a set of hexagonal cells of vital importance for maintaining the transparency of the cornea. Which progressively deteriorates as age increases. This loss can be aggravated by surgical trauma or certain diseases such as Fuchs' corneal dystrophy [1, 2]

Eye problems are highly prevalent and remain untreated in many cases. Among these diseases, Fuchs' corneal endothelial dystrophy affects the corneal endothelium, which is responsible for corneal transparency [3]. According to WHO reports, greater vigilance is needed to guarantee attention to the ophthalmological needs of communities to improve prevention, early detection, treatment, and rehabilitation [4].

Among the challenges of studying the cornea is Fuchs' dystrophy, which is characterized by the accumulation of fluid in the cornea, located in the front part of the eye. This causes the cornea to swell and become thicker. This condition manifests itself with blurred vision and eye discomfort. Causing loss of visual acuity. This pathology of the endothelium is produced in the deepest layer of the cornea, where the cells responsible for maintaining corneal transparency are located. This disease can be caused by drug use, aging, surgeries and inflammation.

Today, there are different technological tools for assessing the state of the corneal endothelium through morphometric parameters, such as cell density. The most used device is the specular microscope. However, several studies have shown that conventional approaches for estimating endothelial morphometric parameters fail in the presence of endotheliopathies such as Fuchs' CE distrophy [5–8].

Recent deep learning-based approaches have significantly improved image segmentation of corneal endothelium and estimation of morphometric parameters [9, 10]. However, these methods rely heavily on large, manually annotated data sets [11–13]. Unsupervised learning methods have emerged to avoid costly manual labeling of images, some authors such as Caron et al., Zbontar et al., Chen et al., and Punn et al., [14–16] have used these tools in their research. These methods are based on coding, clustering, transfer learning, self-monitoring and other strategies.

The problem raised above generates great challenges to be solved. However, handling unlabeled images using unsupervised learning is not straightforward, and models trained with these techniques typically perform much less efficiently than supervised ones. But in recent years, unsupervised models have significantly narrowed the gap with supervised training, particularly with the recent achievements of contrastive and non-contrastive learning methods. Giving rise to self-supervised learning; which is a strategy that combines labeled data with unlabeled data during the training of a neural network. Initially, it learns unlabeled data features, then the weights are frozen and finally used in a tuning step to learn a specific task [17–20]. These architectures are based on generative approaches [21], predictive tasks [22], contrastive and non-contrastive learning [23] and bootstrap approaches [24, 25].

In this paper, we develop a self-supervised artificial intelligence model to address the corneal endothelium segmentation problem. We use a large dataset of unlabeled images to train the encoder of a UNet network using the Barlow Twins approach to learn relevant data representations. Then with few labeled data, we train the UNet decoding path. In the following sections, we briefly review other related work, the proposed method, preliminary results, and conclusions.

2 Related work

In recent years, due to the development of deep learning technologies, researchers have developed a great interest in computer-aided diagnostic systems to support health-care services in different applications for classification, detection, and segmentation tasks [26]. This research will focus specifically on the task of segmentation of escular microscopy images of the corneal endothelium, due to the challenges present in this type of medical images when there is the presence of diseased cells.

Nowadays, many researches use models focused on supervised learning. Although it is a valuable technique in the segmentation of medical images of the corneal endothelium, it presents notable disadvantages compared to self-supervised and semi-supervised approaches. One of the main limitations of supervised learning is the large number of annotated images that the model requires for it to generalize. Not to mention that the acquisition of medical images is complex, they require expert personnel in the area and the labeled databases are limited. In contrast, self-supervised and semi-supervised approaches have the ability to memorize information from unlabeled data and learn a specific task with little annotation, thus addressing some of these limitations. Due to the above, these architectures become promising alternatives in image segmentation of the corneal endothelium.

In this context, Vigueras et al. [27] presented a fully automated method for estimating corneal endothelial parameters from specular microscopy images containing guttae. The proposed model was based on a DenseUNet neural network with non-local feedback attention to perform the semantic segmentation task. In general, the estimates agreed well with the reference values. The parameters were significantly better than those provided by commercial software, demonstrating the ability of this AI architecture to accurately estimate endothelial parameters even in the presence of edotheliopathies, like guttae in Fuchs' dystrophy.

Sierra et al. [3] proposed a UNet-based segmentation approach that requires minimal post-processing and achieves reliable CE morphometric assessment and guttae identification in all grades of Fuchs' dystrophy. They cast the segmentation problem as a regression task, using distance maps rather than a pixel-level classification task, as is typically done with the UNet architecture. These fully supervised approaches achieve decent performance but still require large annotated datasets.

There have been recent efforts in unsupervised or self-supervised methods have been developed to address the problem of the required annotation volume for data sets. In the area of biomedical image segmentation, self-supervised learning strategies can be grouped according to their approach as generative models [28], predictive tasks [22], contrastive learning [29], bootstrapping [24] and regularization [25].

Amodio et al. [30] presented the first fully unsupervised deep learning framework for medical image segmentation, which facilitated the use of the vast majority of image data that is not labeled or annotated. This unsupervised approach is based on a training objective with contrastive learning and self-coding aspects. Previous contrastive learning approaches for medical image segmentation have focused on training at the image level. This approach is proposed at the patch level within the image (pixel-centric). This model achieves improved results in several critical medical imaging tasks, as verified by expert annotations on segmenting geographic atrophy regions from multi-subject retinal images.

Felfeliyan et al. [31] proposed an alternative self-supervised deep learning training strategy on unlabeled magnetic resonance imaging. In this research, they randomly applied different distortions to unlabeled image areas and then predicted the type of distortion and information loss. To do this, they used an improved version of the Mask-RCNN architecture to locate the location of the distortion and retrieve the pixels from the original image. This self-monitored pre-training improved the Dice Index score by 20% compared to training from scratch. The proposed self-supervised learning was simple, effective, and suitable for different ranges of medical image analysis tasks, including abnormality detection, segmentation and classification according to their complexity.

Therefore, the objective of this project is to investigate how unlabeled data can be used to pre-train a network and learn important data representations, then perform fine tuning for the segmentation task with few annotated corneal endothelial miscroscopy images. Unsupervised, semi-supervised, and self-supervised learning are becoming effective substitutes for transfer learning from large data sets.

3 Methods

3.1 Dataset description

We utilized a dataset comprising 1300 in vivo images of corneal endothelial cells obtained from individuals with both healthy and pathological corneas. These images were captured at a resolution of 224x448 pixels. Among these images, we selected 230 patches measuring 96x96 pixels, which were annotated by domain experts. Additionally, we included 1719 patches of the same dimensions that lacked annotations. The acquisition of these images was performed using a Topcon SP3000P specular microscope equipped with Cell Count software. It's worth noting that the image collection process involved capturing images from either both eyes or just one eye, depending on the case. The study received approval from the ethics committee at Universidad Tecnologica de Bolıvar, Colombia. Furthermore, due to the retrospective nature of the study, the requirement for informed consent was waived, in accordance with the principles outlined in the Declaration of Helsinki.

3.2 UNet model

It is a convolutional neural network architecture that was designed for medical image segmentation. This model was originally developed Ronnenberger et al., in 2015 [32]. This architecture consists of two tracks. The first is that of contraction, also called encoder. It is used to capture the context of an image. The second way is symmetric expansion, called decoder. It also allows precise localization through transposed convolution. The mathematical model of the U-Net architecture is described by the following mathematical expressions:

$$w(\mathbf{x}) = w_c(\mathbf{x}) + w_0 \cdot \exp\left(-\frac{(d_1(\mathbf{x})) + d_2(\mathbf{x}))^2}{2\sigma^2}\right), \tag{1}$$

$$E = \sum_{x \,\epsilon\, \Omega} w(\mathbf{x}) \cdot \log(P_{l(\mathbf{x})}(\mathbf{x})), \tag{2}$$

where $P_{l(\mathbf{x})}$ is the output of the softmax function, $d_1(\mathbf{x})$ and $d_2(\mathbf{x})$ indicate distances to the nearest boundary points, w_c represents the weight maps, w_0 and σ are constants.

3.3 Barlow Twins

It is a self-supervised learning method that applies redundancy reduction. It's to learn representations that are invariant to image distortions. It does not require large batches, gradient stops, momentum encoders and predictor networks. To overcome this problem, the Barlow Twins approach was proposed to pre-train the encoder in an unsupervised way, to then perform fine tuning, taking the weights of the pre-trained network and using them for the segmentation task with a limited number of samples annotated [33]. The mathematical model that describes the Barlow Twins method is as follows:

$$LBT = \sum_i (1 - C_{ij})^2 + \lambda \sum_i \sum_{j \neq i} C_{ij}^2, \tag{3}$$

$$C_{ij} = \frac{\sum_b * Z_{bi}^A * Z_{bj}^B}{\sqrt{\sum_b *(Z_{bi}^A)^2} * \sqrt{\sum_b *(Z_{bj}^B)^2}}, \tag{4}$$

where $\sum_i (1 - C_{ij})^2$ is an invariance term (diagonal or identity term) to drive neurons to produce the same output under different magnifications, and $\lambda \sum_i \sum_{j \neq i} C_{ij}^2$ is a redundancy reduction term (off-diagonal term) making each neuron produce a different output. The term λ is used to balance the contribution of the redundancy and invariance reduction terms.

3.4 Self-supervised learning

Self-supervised learning is a machine learning paradigm that allows unlabeled data to be processed in order to obtain useful representations that can assist in subsequent learning tasks.

Due to the major challenges of medical databases, this strategy can help overcome the limitations of the availability of labeled data, allowing artificial intelligence models to capture important features and patterns more effectively. This may lead to an improvement in the accuracy of medical image processing tasks, such as the segmentation task, which in turn may be beneficial for the early detection and diagnosis of ophthalmological diseases.

For this reason, this research proposes the implementation of an architecture based on self-supervised learning focused on the task of segmentation in images of the corneal endothelium, with the aim of improving the performance of CNN networks when there are no labels. In figure 1 we can see the block diagram of the implemented model.

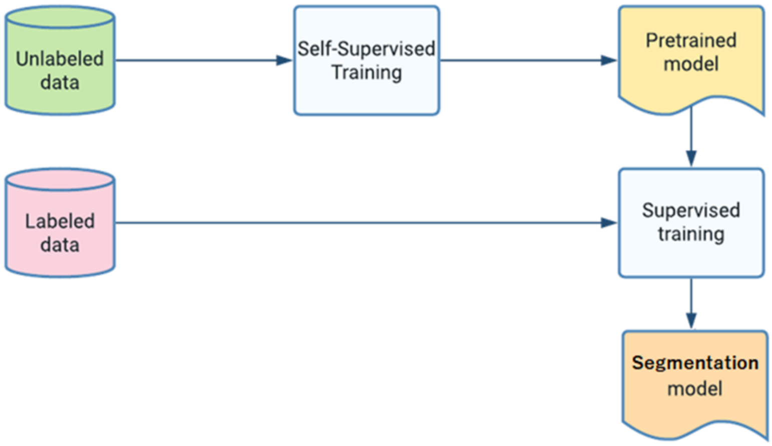

Fig. 1 Block diagram of the proposed SSL models.

In figure 1 we can detail the advantages that weakly supervised learning brings, because you can learn from many unlabeled images and do fine tuning with few annotated images.

3.5 Data augmentation

The proposed model uses the geometric transformations strategy to mitigate the drawback of the absence and imbalance of annotations in the corneal endothelium image databases and is also part of the Siamese network strategy so that the model is more resistant to variations in the input data and generalize better. Particularly, this architecture applies a distortion to the input images randomly before the training stage, using the horizontal flip, vertical flip and rotations operations. In the following figure we can see the details.

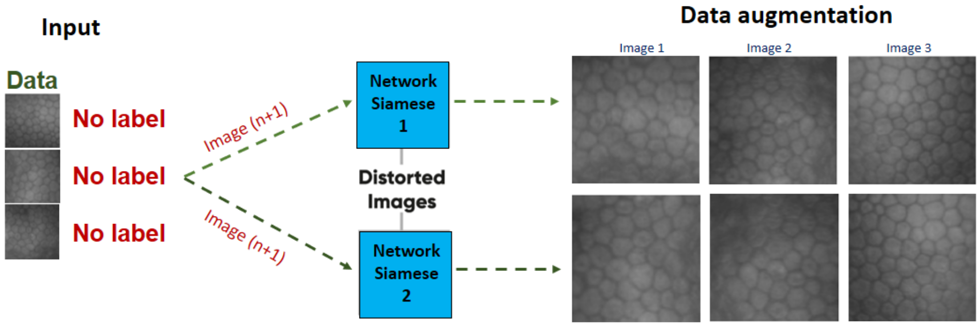

Fig. 2 Data augmentation with geometric transformations.

In figure 2 , we can see how the images enter the Siamese network, then geometric transformations are applied so that both networks are capable of learning intrinsic characteristics and representations despite the input distortions. Which will help to better generalize the proposed model.

3.6 Experiment configuration

For the training and testing of the model, the cross-validation technique was used, where 70% of the images were used for training and 30% for validation. Of the 1300 images supplied with resolution of 224x448, With the Python patchify function, 1719 patches were generated with a resolution of 96x96 without labels and 230 patches with the same resolution annotated by expert personnel. In order to increase and strengthen the feature maps of the unsupervised learning stage, without being required to annotate a very large volume of data, a higher proportion of unlabeled patches (1719 $\approx$ 88%) is provided. On the other hand, we worked with the Adam optimizer with a learning rate initialized at 1×10^{-3} for all the experiments, which had a decay of a factor of 0.1 once the learning stagnated to obtain better results of segmentation.

4 Proposed approach

Despite the good results of CNN-based approaches, they generally exhibit limitations for modeling an explicit long-range relationship, due to the intrinsic locality of the convolution operations. Therefore, these architectures generally produce weak performance, especially for target structures that show large variations between images in terms of texture, shape, and size. To overcome this limitation, it is proposed to establish attention modules, blocks which have skip connections and transfer learning in the encoder of the proposed architecture. Which would allow us to model global and specific contexts. In the following figure we can see the architecture of our proposed model.

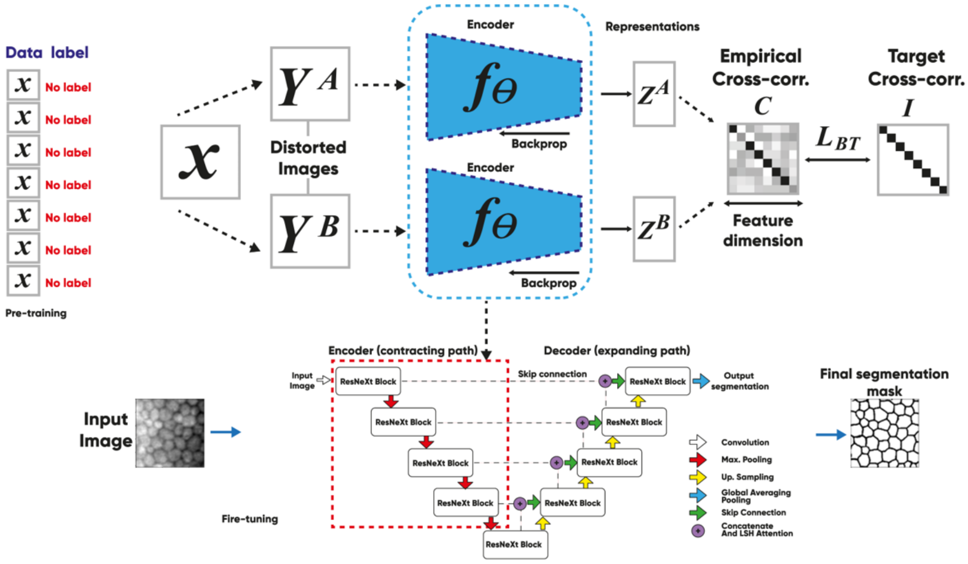

Fig. 3 Architecture of the proposed models for self-supervised learning training of the segmenter.

We aim to improve the predictions of the traditional UNet model. The proposed model addresses the challenge of limited data availability in two phases, as shown in figure 3. In phase one, a pre-training of the UNet encoder is performed with the Barlow Twins method, using the redundancy reduction principle to learn feature representations in an unsupervised way (without data annotations). Finally, for phase 2, fine-tuning is performed by taking the weights of the pre-trained network and then using them for the semantic segmentation task with a limited number of annotated samples.

This proposal combines the advantages of the ResNet (Residual Network) which contains skip connections and subsequent residual blocks to extract semantic features that reduce the number of model parameters and improve the inference speed. Moreover, the benefits of the vision transformer (attention module) of the decoder, manage to combine multi-level functions and capture the global context.

We can see in figure 3 that the model receives a set of unlabeled images which enter a Siamese network. Then, a transformation is applied (rotations, translations, color changes, among others) and an encoder is introduced to generate the most representative feature maps using an empirical cross-correlation function. In the encoder stage, five filters [16, 32, 64, 128, 256], a global average pooling block, a fully connected layer block, ReLU activation, and normalization (FC+ReLU+BN) were used. These representations were frozen and became the encoder of a UNet network. Subsequently, these feature maps are concatenated with the corresponding decoder block, using the skip connections for the feature upsampling operation. Finally, a 1×1 convolution is performed on the output layer to generate a segmentation mask and categorize each pixel of the input images.

5 Results and discussion

We have implemented a self-supervised learning model based on a Barlow Twins approach to train the encoder of a U-Net model in an unsupervised way. This model was pre-trained with a 32-core Linux server and 94 GB of RAM, where specular microscopy images of the corneal endothelium with different resolutions were used. We worked with the TensorFlow framework, the Python programming language, and the NumPy and Pandas libraries. Below you can see in table 1 the quantitative results of the segmentation task in images of the corneal endothelium with the proposed model and the traditional UNet technique.

Table 1 Performance analysis of the UNet model and the proposed one [Accuracy (Acc), Precision (Pr), Dice Coefficient (DC) and mean Intersection-over-Union (mIoU)].

DS	**Model**	**70% Training - 30% Testing**			
		Acc	**Pr**	**DC**	**mIoU**
1	**UNet**	0.8083	0.8420	0.8212	0.2797
	Ours	**0.8619**	**0.9393**	**0.9082**	**0.2939**

In table 1 we can see that our proposed model presents better performance in the four metrics evaluated with respect to the UNet model. This can be corroborated in the segmentations obtained in figure 4. The mask predicted by our proposed method is quite similar to the reference mask, except in cases where images present problems of non-uniform illumination (limitations that will be covered in future approaches) and poor sharpness. These preliminary results demonstrate the benefit of using the pre-training strategy to improve the encoding stage, especially in cases where the availability to collect annotations is limited. In the following figure you can see the results obtained in the segmentation task in images of the corneal endothelium with healthy and diseased cells.

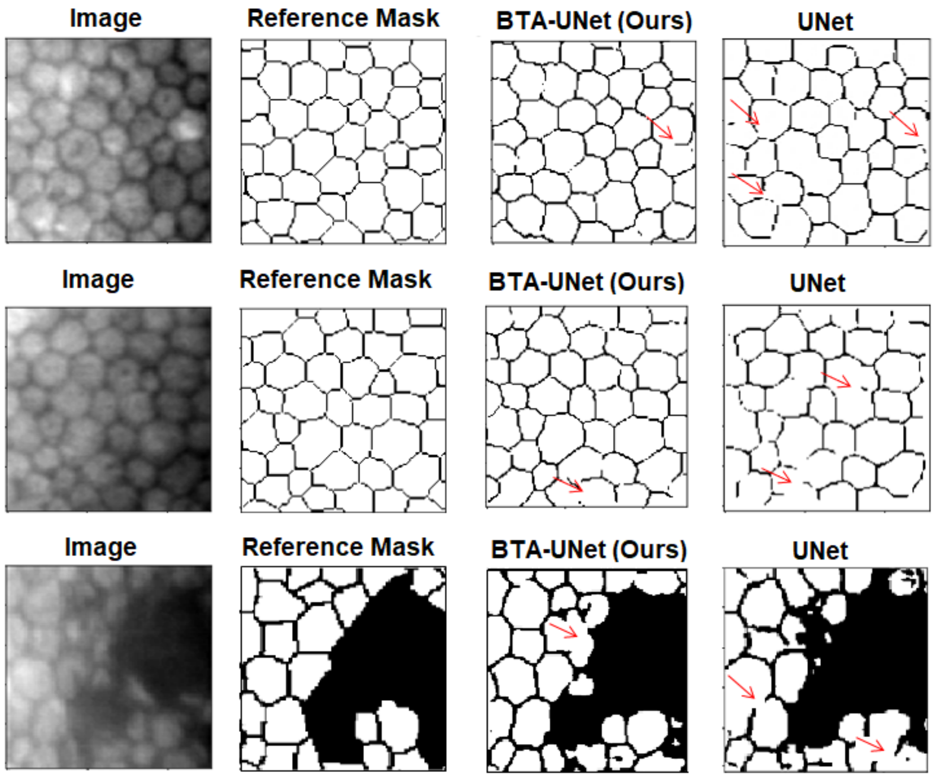

Fig. 4 Segmentation results. The proposed method defines better the intercellular boundary of the segmentation, despite some problems due to non-uniform illuminated areas.

In figure 4, you can see the qualitative performance of the self-supervised model compared to the traditional UNet model. Where it is evident that the proposed model improves its performance by freezing the weights of unsupervised training with unlabeled images, and then performing fine tuning with few images of the corneal endothelium with healthy cells and with Fuchs dystrophy. The images used present great challenges, due to their variations in scale, lighting, shadows, brightness, poor sharpness, among other aspects that make the training more complex.

6 Conclusion

We have developed a self-supervised learning model for the semantic segmentation task of images of the corneal endothelium obtained by specular microscopy to address one of the main challenges of the limited availability of annotated data. The results showed that the proposed strategy improves the segmentation performance of the UNet model. This improvement is evidenced due to encoder tuning, using jump connections, residual blocks, and attention modules. Future work involves further experiments and exploring different training strategies and settings in the encoder to generate better feature maps and ensure more accurate image segmentation. Finally, it is proposed to use pre-trained encoders with different databases that have learned low, medium and high level characteristics that help to better generalize the network.

Acknowledgments. This work has been partly funded by Ministerio de Ciencia, Tecnología e Innovacioón, Colombia, Project 124489786239 (Contract 763-2021), Universidad Tecnológica de Bolívar (UTB) Project CI2021P02. K. Mendoza and F. Quintero thank UTB for a post-graduate scholarship. S. Sanchez thanks Fondo Bicentenario for a Ph.D. Scholarship.

References

[1] Jeang, L.J., Margo, C.E., Espana, E.M.: Diseases of the corneal endothelium. Experimental eye research **205**, 108495 (2021)

[2] Catala, P., Thuret, G., Skottman, H., Mehta, J.S., Parekh, M., Dhubhghaill, S.N., Collin, R.W., Nuijts, R.M., Ferrari, S., LaPointe, V.L., *et al.*: Approaches for corneal endothelium regenerative medicine. Progress in retinal and eye research **87**, 100987 (2022)

[3] Sierra, J.S., Pineda, J., Rueda, D., Tello, A., Prada, A.M., Galvis, V., Volpe, G., Millan, M.S., Romero, L.A., Marrugo, A.G.: Corneal endothelium assessment in specular microscopy images with fuchs' dystrophy via deep regression of signed distance maps. Biomedical optics express **14**(1), 335–351 (2023)

[4] Knauer, C., Pfeiffer, N.: The value of vision. Graefe's Archive for Clinical and Experimental Ophthalmology **246**, 477–482 (2008)

[5] Huang, J., Maram, J., Tepelus, T.C., Sadda, S.R., Chopra, V., Lee, O.L.: Comparison of noncontact specular and confocal microscopy for evaluation of corneal endothelium. Eye & Contact Lens: Science & Clinical Practice **44**, 144–150 (2017)

[6] Price, M.O., Fairchild, K.M., Price, F.W.: Comparison of manual and automated endothelial cell density analysis in normal eyes and dsek eyes. Cornea **32**(5), 567–573 (2013) https://doi.org/10.1097/ico.0b013e31825de8fa

[7] Luft, N., Hirnschall, N., Schuschitz, S., Draschl, P., Findl, O.: Comparison of 4 specular microscopes in healthy eyes and eyes with cornea guttata or corneal grafts. Cornea **34**(4), 381–386 (2015) https://doi.org/10.1097/ico.0000000000000385

[8] Gasser, L., Reinhard, T., Böhringer, D.: Comparison of corneal endothelial cell measurements by two non-contact specular microscopes. BMC ophthalmology **15**, 87 (2015) https://doi.org/10.1186/s12886-015-0068-1

[9] Piórkowski, A., Gronkowska–Serafin, J.: Towards precise segmentation of corneal

endothelial cells. In: Ortuño, F., Rojas, I. (eds.) Bioinformatics and Biomedical Engineering, pp. 240–249. Springer, Cham (2015)

[10] Selig, B., Vermeer, K., Rieger, B., Hillenaar, T., Luengo Hendriks, C.: Fully automatic evaluation of the corneal endothelium from in vivo confocal microscopy. BMC medical imaging **15**, 13 (2015) https://doi.org/10.1186/s12880-015-0054-3

[11] Shilpashree, P., Kaggere, S., Sudhir, R., Srinivas, S.: Automated image segmentation of the corneal endothelium in patients with fuchs dystrophy. Translational Vision Science & Technology **10**, 27 (2021) https://doi.org/10.1167/tvst.10.13.27

[12] Daniel, M., Atzrodt, L., Bucher, F., Wacker, K., Böhringer, S., Reinhard, T., Böhringer, D.: Automated segmentation of the corneal endothelium in a large set of 'real-world' specular microscopy images using the u-net architecture. Scientific Reports **9** (2019) https://doi.org/10.1038/s41598-019-41034-2

[13] Vigueras-Guillén, J., Rooij, J., Dooren, B., Lemij, H., Islamaj, E., Van Vliet, L., Vermeer, K.: Denseunets with feedback non-local attention for the segmentation of specular microscopy images of the corneal endothelium with guttae. Scientific Reports **12** (2022) https://doi.org/10.1038/s41598-022-18180-1

[14] Caron, M., Touvron, H., Misra, I., Jégou, H., Mairal, J., Bojanowski, P., Joulin, A.: Emerging Properties in Self-Supervised Vision Transformers (2021)

[15] Zbontar, J., Jing, L., Misra, I., LeCun, Y., Deny, S.: Barlow twins: Self-supervised learning via redundancy reduction. In: International Conference on Machine Learning, pp. 12310–12320 (2021). PMLR

[16] Punn, N.S., Agarwal, S.: Bt-unet: A self-supervised learning framework for biomedical image segmentation using barlow twins with u-net models. Machine Learning **111**(12), 4585–4600 (2022)

[17] Jiao, R., Zhang, Y., Ding, L., Cai, R., Zhang, J.: Learning with Limited Annotations: A Survey on Deep Semi-Supervised Learning for Medical Image Segmentation (2022)

[18] Wu, Y., Ge, Z., Zhang, D., Xu, M., Zhang, L., Xia, Y., Cai, J.: Mutual Consistency Learning for Semi-supervised Medical Image Segmentation (2022)

[19] Fang, K., Li, W.-J.: DMNet: Difference Minimization Network for Semi-supervised Segmentation in Medical Images, pp. 532–541 (2020). https://doi.org/10.1007/978-3-030-59710-8_52

[20] Chen, S., Bortsova, G., Juarez, A.G.-U., Tulder, G., Bruijne, M.: Multi-Task Attention-Based Semi-Supervised Learning for Medical Image Segmentation (2019)

[21] Devlin, J., Chang, M.-W., Lee, K., Toutanova, K.: Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805 (2018)

[22] Pathak, D., Krähenbühl, P., Donahue, J., Darrell, T., Efros, A.A.: Context encoders: Feature learning by inpainting. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2536–2544 (2016). https://doi.org/10.1109/CVPR.2016.278

[23] Balestriero, R., LeCun, Y.: Contrastive and Non-Contrastive Self-Supervised Learning Recover Global and Local Spectral Embedding Methods (2022)

[24] Grill, J.-B., Strub, F., Altché, F., Tallec, C., Richemond, P.H., Buchatskaya, E., Doersch, C., Pires, B.A., Guo, Z.D., Azar, M.G., Piot, B., Kavukcuoglu, K., Munos, R., Valko, M.: Bootstrap your own latent: A new approach to self-supervised Learning (2020)

[25] Ghosh, S., Seth, A., , Mittal, D., Singh, M., Umesh, S.: DeLoRes: Decorrelating Latent Spaces for Low-Resource Audio Representation Learning (2022)

[26] Wang, R., Lei, T., Cui, R., Zhang, B., Meng, H., Nandi, A.K.: Medical image segmentation using deep learning: A survey. IET Image Processing **16**(5), 1243–1267 (2022) https://doi.org/10.1049/ipr2.12419

[27] Vigueras-Guillén, J.P., Sari, B., Goes, S.F., Lemij, H.G., Rooij, J., Vermeer, K.A., Vliet, L.J.: Fully convolutional architecture vs sliding-window cnn for corneal endothelium cell segmentation. BMC biomedical engineering **1**, 4 (2019) https://doi.org/10.1186/s42490-019-0003-2

[28] Devlin, J., Chang, M.-W., Lee, K., Toutanova, K.: BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding (2019)

[29] Noroozi, M., Favaro, P.: Unsupervised Learning of Visual Representations by Solving Jigsaw Puzzles (2017)

[30] Liu, C., Amodio, M., Shen, L.L., Gao, F., Avesta, A., Aneja, S., Wang, J.C., Priore, L.V.D., Krishnaswamy, S.: CUTS: A Fully Unsupervised Framework for Medical Image Segmentation (2023)

[31] Felfeliyan, B., Hareendranathan, A., Kuntze, G., Cornell, D., Forkert, N.D., Jaremko, J.L., Ronsky, J.L.: Self-Supervised-RCNN for Medical Image Segmentation with Limited Data Annotation (2022)

[32] Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18, pp. 234–241 (2015). Springer

[33] Marsocci, V., Scardapane, S.: Continual barlow twins: Continual self-supervised learning for remote sensing semantic segmentation. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing **16**, 5049–5060 (2023) https://doi.org/10.1109/JSTARS.2023.3280029