

Development of Cumulative Order-Preserving Image Transformation Based Variational Autoencoder for Anterior Segment Optical Coherence Tomography Images

Kilhwan Shon^{1,4}, Kyung Rim Sung², Jiehoon Kwak², Joo Yeon Lee^{3,4}, and Joong Won Shin²

¹ Department of Ophthalmology, Gangneung Asan Hospital, Gangneung, Korea

² Department of Ophthalmology, College of Medicine, University of Ulsan, Asan Medical Center, Seoul, Korea

³ Camp 9 Orthopedic Clinic, Hwaseong-si, Gyeonggi-do, Korea

⁴ Asan Artificial Intelligence Institute, Hwaseong-si, Gyeonggi-do, Korea

Correspondence: Kyung Rim Sung, Department of Ophthalmology, University of Ulsan, College of Medicine, Asan Medical Center, 388-1 Pungnap-2-dong, Songpa-gu, Seoul, Korea.
e-mail: sungeye@gmail.com

Received: February 10, 2022

Accepted: July 24, 2022

Published: August 30, 2022

Keywords: variational autoencoder; artificial intelligence; angle closure glaucoma; anterior segment optical coherence tomography

Citation: Shon K, Sung KR, Kwak J, Lee JY, Shin JW. Development of cumulative order-preserving image transformation based variational autoencoder for anterior segment optical coherence tomography images. *Transl Vis Sci Technol.* 2022;11(8):30.
<https://doi.org/10.1167/tvst.11.8.30>

Purpose: To develop a variational autoencoder (VAE) suitable for analysis of the latent structure of anterior segment optical coherence tomography (AS-OCT) images and to investigate possibilities of latent structure analysis of the AS-OCT images.

Methods: We retrospectively collected clinical data and AS-OCT images from 2111 eyes of 1261 participants from the ongoing Asan Glaucoma Progression Study. A specifically modified VAE was used to extract six symmetrical and one asymmetrical latent variable. A total of 1692 eyes of 1007 patients were used for training the model. Conventional measurements and latent variables were compared between 74 primary angle closure (PAC) and 51 primary angle closure glaucoma (PACG) eyes from validation set (419 eyes of 254 patients) that were not used for training.

Results: Among the symmetrical latent variables, the first three and the last demonstrated easily recognized features, anterior chamber area in η_1 , curvature of the cornea in η_2 , the pupil size in η_3 and corneal thickness in η_6 , whereas η_4 and η_5 were more complex aggregating complex interactions of multiple structures. Compared with PAC eyes, there was no difference in any of the conventional measurements in PACG eyes. However, values of η_4 were significantly different between the two groups, being smaller in the PACG group ($P = 0.015$).

Conclusions: VAE is a useful framework for analysis of the latent structure of AS-OCT. Latent structure analysis could be useful in capturing features not readily evident with conventional measures.

Translational Relevance: This study suggested that a deep learning-based latent space model can be applied for the analysis of AS-OCT images to find latent characteristics of the anterior segment of the eye.

Introduction

Although advances have been made in imaging techniques of the anterior segment, such as anterior segment optical coherence tomography (AS-OCT), the appropriate analysis of acquired high-resolution images has been limited by the lack of proper analytical tools. Conventional methods consist of manual measurement of hand-crafted features by the physician. The commonly used parameters are as follows;

anterior chamber depth (ACD), width, and area (ACA), lens vault, angle recess area, angle opening distance, trabecular-iris-space area, and iris thickness.¹ Conventional analysis has been successfully utilized in a variety of tasks, including subclassification, monitoring intervention-induced changes, and describing dynamic and long-term processes in the anterior segment. Still, these parameters cannot inherently comprehend the whole image and have vulnerabilities when considering highly correlated parameters.²⁻⁷

Recent advances in deep learning technology provide a new approach to image data analysis. Several studies have proven that deep learning techniques are not only suitable for the analysis of AS-OCT images but can achieve accuracies comparable to human measurements.⁸⁻¹² However, because these studies usually involve large-scale automated measurements of manually assigned labels, they are affected by all the limitations of manual measurements.

Latent space modeling is one of the machine learning approaches that could be applied to analyze high-dimensional data. An example of a latent space model (or latent variable model) used in the conventional context is factor analysis, such as the ones that have long been used in psychology.¹³ In the field of computer vision, a family of latent space models called deep generative models has been intensively developed to analyze various images.¹⁴ In our previous study, we have shown that a convolutional β -variational autoencoder (VAE) can be applied to the AS-OCT images to achieve a good disentangled latent space representation.¹⁵ Despite encouraging results, we also found shortcomings of a convolutional VAE framework, which motivated us to improve the model. The previous model had limited power in separating asymmetrical variance from symmetrical variance, which hampered the disentanglement of latent variables.

To overcome problems arising from asymmetry, we have developed a new method inspired by spatial transformer networks.¹⁶ In this new model, we have preserved the overall framework of VAE, but instead of convolution, an image warping technique we have named cumulative order-preserving image transforming network (COPIT) is used to reconstruct images from the latent space. COPIT was specifically developed and tailored for the current article's latent space representation of AS-OCT images. COPIT has been designed to have several properties: (1) the order of the x and y coordinates are preserved after transformation; (2) each latent variable defines a unique transformation; (3) multiple transformations can be combined into a single new transformation; and (4) each layer can be designed differently depending on the purposes. The entire network is based on the convolutional β -VAE with two modifications: (1) the convolutional decoder part is replaced by COPIT, and (2) the loss function has been modified to include a cosine similarity.

Methods

Participants

We selected participants from the ongoing Asan Glaucoma Progression Study who have undergone an

AS-OCT examination (Visante OCT, ver. 3.0; Carl Zeiss Meditec, Jena, Germany). From 2111 eyes of 1261 patients, we randomly assigned 80% of the patients to the training set (1692 eyes of 1007 patients) and the remaining 20% to the validation set (419 eyes of 254 patients), ensuring both eyes of the patient were assigned to the same group. Comparisons between conventional measurements and latent variables were made using patients from the validation set. A more detailed description of the population, including the clinical assessment, inclusion criteria, image acquisition, and demographics, has been included in our previous study.¹⁵

All procedures conformed to the Declaration of Helsinki, and this study was approved by the institutional review board of the Asan Medical Center, University of Ulsan, Seoul, Korea.

Image Preparation and Segmentation

Raw AS-OCT images of size 1200×1500 ($H \times W$) pixels were center cropped to create a grayscale 512×1024 image, resized to 256×512 pixels. The segmentation was done in four steps: (1) resized images were manually segmented into three segments: the iris, the corneal shell, and the anterior chamber by an experienced glaucoma specialist (KS); (2) 130 manually segmented images were used to train a modified U-net; (3) a trained modified U-net was used to segment remaining images; and (4) segmented images were aligned with rotation and translation using a spatial-transformer network.¹⁶ A modified U-net is structurally identical to the original U-net but has been reduced in depth and has been adjusted to a different resolution.^{15,17} Segmented images were cropped to a size of 192×448 , and left eyes were horizontally flipped.

Model Structure

Our model follows the general structure of VAE, but unlike the typical convolutional autoencoder, a decoder part was replaced with an image warping technique that was inspired by a spatial transformer network.¹⁶ Compared to the original spatial transformer network, our model has the following major structural differences: (1) added a "reparameterization" part; (2) replaced a "grid generator" with our new COPIT; (3) replaced the sampling technique with a linearized multi-sampling technique proposed by Jiang et al.,¹⁸ and (4) have modified the loss function.^{16,18,19}

The encoder (which corresponds to a "localization net" in the spatial transformer network) is identical to the model described in our previous study but was trained de novo.¹⁵ A reparameterization part and

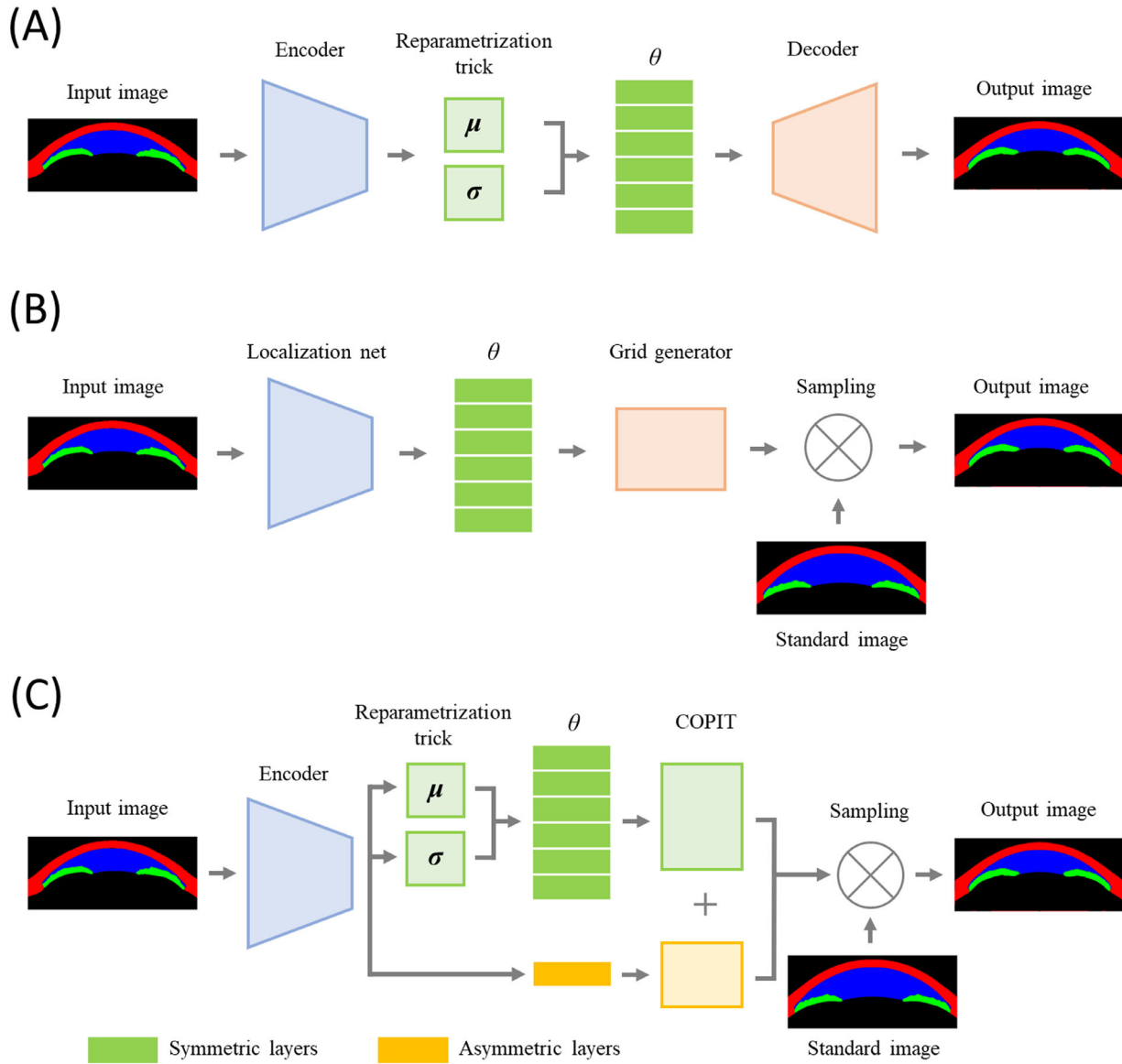


Figure 1. Schematic diagram of convolutional VAE, spatial transformer and COPIT based VAE. In a convolutional VAE, convolutional encoder is used to generate means and standard deviations μ and σ , from which latent variables are sampled, and then output image is generated from sampled latent variables through convolutional decoder (A). In a spatial transformer, convolutional localization net is used to generate latent variables which are used as transformation coefficients for generating sampling grids, and then output image is calculated using a sampling grid and the standard image (B). In our new model, latent variables are generated in the same way as in the convolutional VAE, but sampled in similar way as in the spatial transformer but grid generator has been replaced with COPIT which two versions - six symmetrical and one asymmetrical. Output image is calculated from final sampling grid and standard image using linearized multi-sampling (C).

addition of Kullback-Leibler divergence (KLD) to the loss function were taken from makes our model a VAE.¹⁹ The decoder part in our preceding work or a “grid generator” in the spatial transformer network has been replaced with a COPIT-based decoder specifically developed for the current research. In the COPIT, numbers are generated from reparametrized latent variables using fully connected layers, which are then fed into COPIT to generate a sampling grid. Then a transformed image is calculated from the standard image using a sampling grid (Fig. 1).

To separate symmetrical variability from asymmetrical variability, we used two versions of COPIT layers: an asymmetrical layer and a symmetrical layer (the left and right sides of the grid are mirror images reflected over the y axis). All variationally inferred latent variables have been matched to the symmetrical layers, while one additional variable was matched with an asymmetrical layer. Also, because generating grids with a number of points equivalent to or larger than the number of pixels, in our case $192 \times 448 = 86,016$, is not only inefficient but might also cause instability, we

have generated down-scaled grids from fully connected layers that were upsampled using bicubic interpolation. A detailed description of COPIT can be found in the supplementary material.

The Loss Function

A latent space z in the VAE is variationally inferred such that a posterior $p_\theta(z|x)$ is approximated to a prior $p_\phi(z)$, which is usually defined as a Gaussian distribution $N(0, I)$.²⁰ The distance between a posterior and a prior is measured with a KLD. However, KLD does not provide information regarding the similarity between latent variables. Hence, we have decided to add cosine similarities between all possible combinations of latent variables:

$$Sim(z) = \sum_{i=1}^{n-1} \sum_{j=i+1}^n \frac{\|z_i \cdot z_j\|}{\|z_i\| \|z_j\|} \quad (1)$$

In a special case where the data is centered at zero, cosine similarity is equivalent to Pearson's correlation coefficient.²¹

With additional hyperparameter γ and the similarity function Sim , our modified loss function VAE is given as:

$$L(\theta, \phi, \beta; x) = -\mathbb{E}_{q_\phi(z|x)} [\log p_\theta(x|z)] + \beta D_{KL}(q_\phi(z|x) || p(z)) + \gamma Sim(z) \quad (2)$$

where x is an image in our case, $q_\phi(z|x)$ is an estimated distribution of the latent space, $p_\theta(x|z)$ is the likelihood of generating a true image, and D_{KL} is a KLD.²²

Adjusting Hyperparameters and Training

The number of symmetrical latent variables was set to 6 based on our previous work with an additional asymmetrical layer, whereas the values of β and γ were both set at 5^2 , which yielded a comparable KLD to the convolutional VAE model presented in our previous work.¹⁵ The grids were scaled down by a factor of 16, 16, 16, 8, 4, 2 for symmetrical layers 1 to 6 and by a factor of 8 for the asymmetrical layer. For the calculation of the reconstruction loss, mean squared error has resulted in a shorter training time than binary cross-entropy. The layers were trained in three steps: (1) six symmetrical layers were trained sequentially (only one layer was trained at a time while the other layers were frozen) for 900 epochs, (2) the asymmetrical layer was trained for 50 epochs, and (3) all layers were simultaneously trained for 50 epochs.

Making Conventional Measurements and Calculating Latent Variables in Selected Eyes

Exclusively from patients in the validation set, we collected complete clinical information of patients diagnosed with primary angle closure (PAC) or primary angle closure glaucoma (PACG). All patients have undergone static and dynamic gonioscopy with Sussman 4-mirror gonioscope (Ocular Instruments, Bellevue, WA, USA) in a darkened room (0.5 cd/m^2) by an experienced glaucoma specialist (K.R.S). PAC was diagnosed if the eyes had an occludable angle (pigmented posterior trabecular meshwork was not visible on nonindentation gonioscopy for at least 180° in the primary position) with signs indicating trabecular obstruction (elevated intraocular pressure, distortion of the radially orientated iris fibers, "glaukomflecken" lens opacities, excessive pigment deposition on the trabecular meshwork, or presence of peripheral anterior synechiae) but without any sign suggestive of glaucoma on optic disc examination and visual field tests. PAC eyes showing glaucomatous optic disc changes (neuroretinal rim thinning, disc excavation, or optic disc hemorrhage attributable to glaucoma) or a glaucomatous visual field change were classified as PACG. Eyes with a previous history suggesting acute angle closure attack have been excluded: (1) presenting with ocular or periocular pain, nausea or vomiting, or intermittent blurred vision with haloes; (2) those with a presenting intraocular pressure of more than 30 mm Hg; and (3) those who had experienced at least three of the following: conjunctival injections, corneal epithelial edema, mid-dilated unreactive pupil. If both eyes were eligible, we selected the right eye. As a result, 125 eyes of 125 patients, including 74 PAC eyes and 51 PACG eyes, were analyzed.

A single investigator (S.K), blinded to all information, assigned the scleral spur—defined as the point showing a change in the curvature of the inner surface of the angled wall - and measured the ACD using calipers built-in in the software provided by the manufacturers.²³ Then, the software provided by the manufacturer automatically measured the scleral spur angle, angle opening distance at 500 μm and 750 μm , angle recess area at 500 μm and 750 μm , and the trabecular-iris space area at 500 μm and 750 μm . Additionally, the iris thickness at 750 μm from the scleral spur, iris curvature, ACD, anterior chamber width, ACA, and lens vault were measured using Fiji software, and pixel values were converted into real-world units by comparing pupil diameter measured with Fiji software and built-in and calipers.²⁴ More detailed descriptions of the measurement methods can be found in our previous works.^{25–27}

Clinical information, including age, gender, axial length, baseline intraocular pressure, manual measurements, and values of latent variables derived from the neural network trained in previous steps were compared between the PAC and PACG eyes with Student's t test for continuous variables and the χ^2 test for frequency variables using SAS 9.4 software (SAS Institute Inc., Cary, NC, USA). We have not collected detailed information from the training dataset due to the larger size of the dataset, but we assume the proportion of PAC/PACG eyes in the training dataset to be not statistically different because patients were randomly assigned training and validation datasets and demographics do not differ.¹⁵

Results

Exploration of the Latent Space

A satisfactory latent space disentanglement was achieved in our new method, such that the variables were discernible and appreciable in the visual analysis. Specifically, η_1 seems to represent an overall ACD and ACA, whereas η_2 seems to mainly represent the curvature of the cornea. The η_3 was associated with pupil

size changes without any noticeable changes in the corneoscleral or lens contour. The η_4 was also related to the pupil size, but in contrast to η_3 , the following differences are worth mentioning: (1) the iris is thicker and more curved with a small pupil size (negative z value); (2) the iris became relatively flat when the pupil got larger (positive z value); and (3) there was an overt increase in the lens vault with a positive z value. Little perceivable changes were seen for the η_5 variable, but by creating a subtraction image with extreme z values, we can notice some interesting characteristics: (1) the iris is flatter at negative z values and more curved at positive z values; (2) at negative z values, the iris becomes thinner, but the peripheral part remains thick; (3) the anterior chamber gets narrower at negative z values and wider at positive z values; (4) the periphery of the anterior chamber becomes shallower at negative z values and deeper at positive z values; and (5) there is a subtle change in the corneal profile such that at negative z values, the central portion of the cornea is steeper whereas the peripheral portion of the cornea is flatter and thicker. The η_6 seems to be mainly related to the corneal thickness, whereas η_A appears to represent the asymmetry as intended (Fig. 2 and Fig. 3).

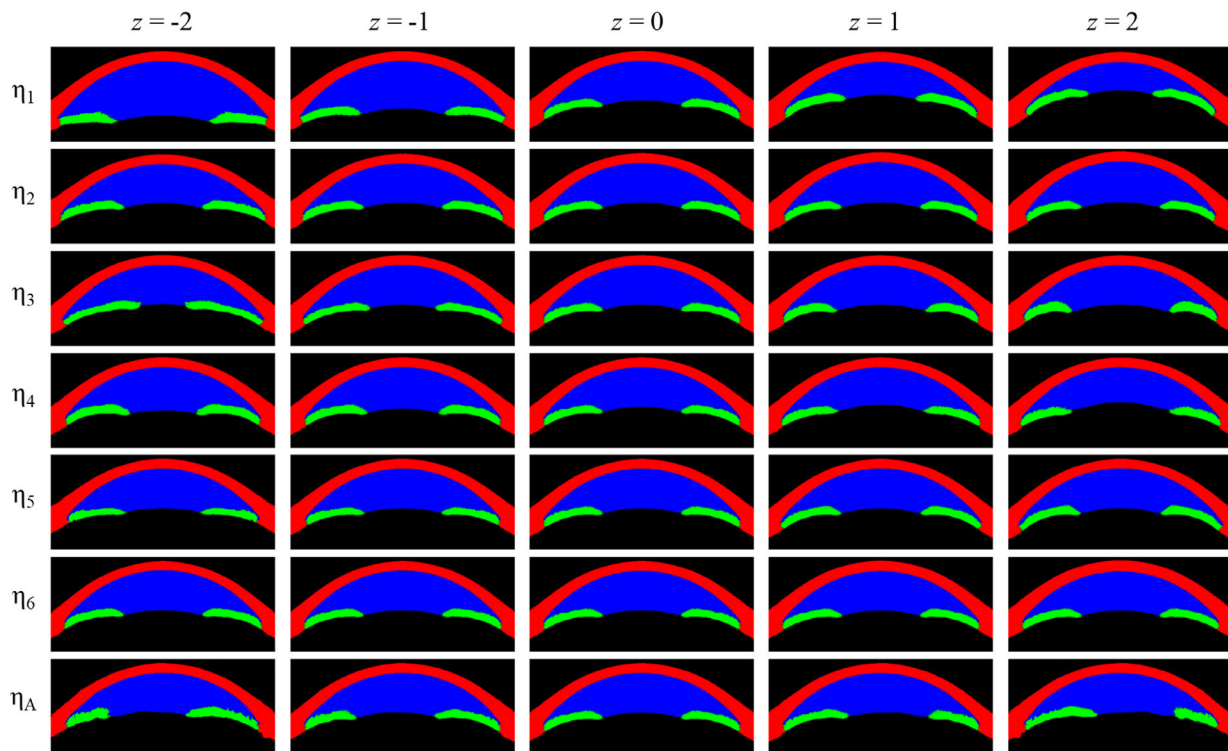


Figure 2. Visualization of the latent variables of the model. There are six symmetrical layers denoted η_1 to η_6 and one asymmetric layer denoted η_A . For asymmetric layers, z values of -2.0 , -1.0 , 0 , 1.0 , and 2.0 were used while for the asymmetric layer, z values have been reduced to -1.0 , -0.5 , 0 , 0.5 , and 1.0 for better interpretability.

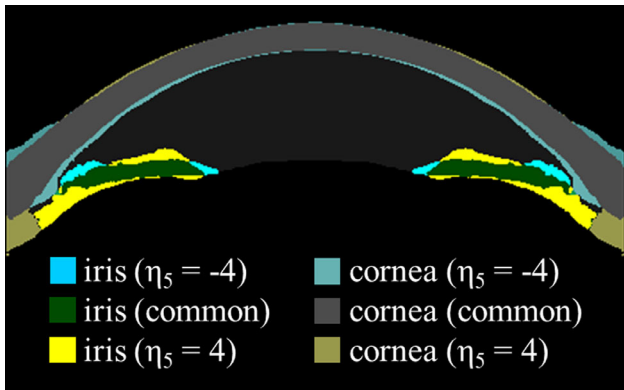


Figure 3. Subtraction image of η_5 for z values -4 and 4 . Areas that correspond to $z = 4$ are color-coded in yellow tone while areas that correspond are color-coded in blue tone. Gray color represents an area of the cornea which is common to both z values and green color represents areas of iris common to both z values.

Conventional Measurements and Latent Variables for PAC and PACG Eyes

There was no statistically significant difference between PAC and PACG eyes for any of the conventional measurements, with angle recess area having the lowest P value ($P = 0.116$). Among the latent variables, PACG eyes have a smaller value of η_4 compared to PAC eyes ($P = 0.015$; Table). Values of the latent variables can be visualized with the network in several ways. When using mean values of η_1 to η_6 for PAC and PACG eyes, there is little difference in the cornea or lens between the groups. The size of the pupil and width of the anterior chamber was slightly smaller in the PACG group compared to the PAC group, which is consistent with the conventional measurements in the Table (the top row of Fig. 4). To enhance the difference, we upscaled the mean values of the latent variables

Table. Conventional Measurements and Latent Variables for PAC and PACG Eyes

	PAC ($n = 74$)	PACG ($n = 51$)	P Value
Age, years	61.8 (± 9.3)	63 (± 8.5)	0.473
Sex, % of females	62 (83.8%)	42 (82.4%)	0.833
Laterality, right	68 (91.9%)	43 (88.2%)	0.495
Axial length, mm	22.7 (± 0.7)	22.7 (± 0.6)	0.916
Baseline IOP, mmHg			
Pupil diameter, mm	4.27 (± 0.87)	4.09 (± 1.03)	0.278
ACD, mm	2.04 (± 0.26)	2.03 (± 0.26)	0.867
ACW, mm	11.82 (± 0.54)	11.76 (± 0.43)	0.543
ACA, mm ²	13.61 (± 2.03)	13.01 (± 2.12)	0.116
LV, mm	1.14 (± 0.32)	1.09 (± 0.29)	0.378
AOD500, μm	148.3 (± 86.7)	128.8 (± 82.3)	0.210
AOD750, μm	213.2 (± 113.1)	188 (± 112.3)	0.223
ARA500, mm ²	0.077 (± 0.047)	0.070 (± 0.039)	0.353
ARA750, mm ²	0.121 (± 0.067)	0.108 (± 0.060)	0.283
TISA500, mm ²	0.058 (± 0.032)	0.053 (± 0.028)	0.428
TISA750, mm ²	0.104 (± 0.057)	0.092 (± 0.050)	0.223
SSA, degrees	16.0 (± 8.4)	14.0 (± 8.3)	0.206
IT750, μm	0.432 (± 0.094)	0.433 (± 0.079)	0.942
Iris curvature, μm	0.18 (± 0.073)	0.186 (± 0.076)	0.773
η_1	0.499 (± 0.548)	0.694 (± 0.791)	0.105
η_2	0.155 (± 0.867)	-0.089 (± 0.965)	0.142
η_3	0.034 (± 0.947)	0.114 (± 1.004)	0.650
η_4	0.412 (± 0.872)	0.012 (± 0.922)	0.015*
η_5	0.062 (± 1.083)	0.035 (± 0.856)	0.882
η_6	0.013 (± 0.845)	-0.130 (± 1.02)	0.397
η_A	0.204 (± 0.115)	0.199 (± 0.13)	0.806

IOP = intraocular pressure; ACD = anterior chamber depth; ACW = anterior chamber width; ACA = anterior chamber area; LV = lens vault; AOD500 = angle opening distance at 500 μm from the scleral spur; AOD750 = angle opening distance at 750 μm from the scleral spur; ARA500 = angle recess area at 500 μm from the scleral spur; ARA750 = angle recess area at 750 μm from the scleral spur; TISA500 = trabecular-iris space area at 500 μm from the scleral spur; TISA750 = trabecular-iris space area at 750 μm from the scleral spur; SSA = scleral spur angle; IT750 = iris thickness at 750 μm from the scleral spur.

*Statistically significant at $P = 0.05$.

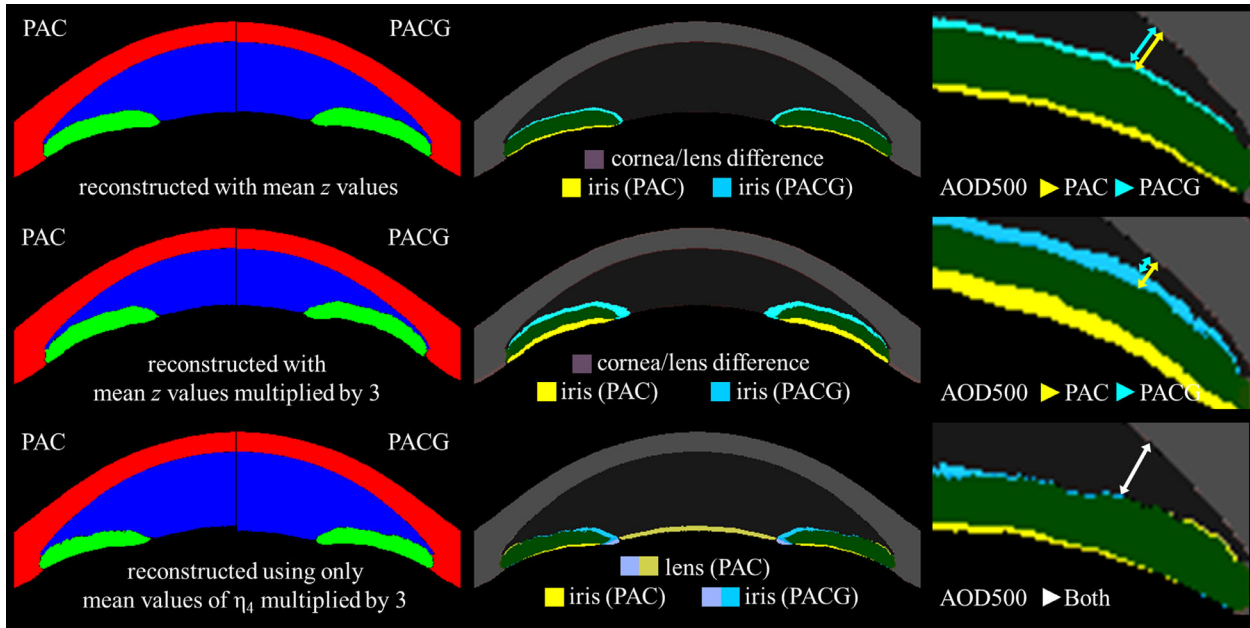


Figure 4. Reconstructed images representing PAC and PACG eyes from selected mean values of latent variables of each group. Top row: mean values of $\eta_1 \sim \eta_6$, middle row: tripled mean values of $\eta_1 \sim \eta_6$, bottom row: tripled mean values of η_4 . Left column: plain segmented images, middle column: difference map, right column: magnified difference map of the angle with approximate AOD500 marked with arrows. We can see that PACG eyes seem to have a narrower angle (top row, right column) which is more pronounced if we multiply latent variables by 3 (middle row, right column). However, there was no noticeable difference in the width of the angle regarding only η_4 (bottom row, right column). It implies that despite significant statistical difference between two disease groups in η_4 , narrower angle in PACG is not a direct result of η_4 but rather a result of combination all latent variables.

for each group by a factor of three. The angle difference was more noticeable with the narrower angle in the PACG group compared to that of the PAC group (middle row of Fig. 4), although there was little difference for the cornea or lens. Because we can select certain latent variables, we have created a subtraction image using the mean values of η_4 for the PAC and PACG eyes multiplied by 3, whereas all other latent variables are kept constant at zero. Still, there is little difference in the angle, corneal contour, or anterior chamber width, although the lens vault is slightly larger in the PAC group (bottom row of Fig. 4).

Discussion

Our model has successfully disentangled latent space with readily distinguishable main features for the first three latent variables: anterior chamber area for η_1 , the curvature of the cornea for η_2 , and pupil size for η_3 . η_4 and η_5 are more complex, with η_4 associated with at least three features: (1) pupil size, (2) curvature and thickness of the iris, and (3) the lens vault. The η_5 is the most complex, with combined changes of the iris profile, corneal profile, width of the anterior chamber, and depth of the peripheral anterior chamber. η_6 seems

to be associated with corneal thickness (Figs. 2 and 3). Given that our model is unsupervised, good interpretability of certain latent variables (η_1 , η_2 , η_3 , and η_6) is encouraging. However, some latent variables (η_4 , η_5) are difficult to interpret, which implies complex interactions between AS-OCT features but also leaves room for further improvement of the model.

Comparing conventional measurements and latent variables of PAC and PACG eyes led to interesting results: there was no statistically significant difference in any of the conventional measurements, but the PACG eyes had smaller values of η_4 compared to PAC eyes ($P = 0.015$; Table). On the visualization of latent variables of the two groups, we noticed that PACG eyes seemed to have a narrower angle, which was more pronounced if the latent variables were multiplied by a factor of 3. However, despite the statistically significant difference of η_4 between disease groups, there was no noticeable difference in the width of the angle between groups if only η_4 was visualized (Fig. 4). Hence, if there is a difference in the angle between PAC and PACG eyes, it is not a direct result of η_4 but rather due to the combined action of all latent variables. Although current research is not sufficient to draw any conclusion regarding morphological differences between PAC and PACG eyes, such results are encouraging in that

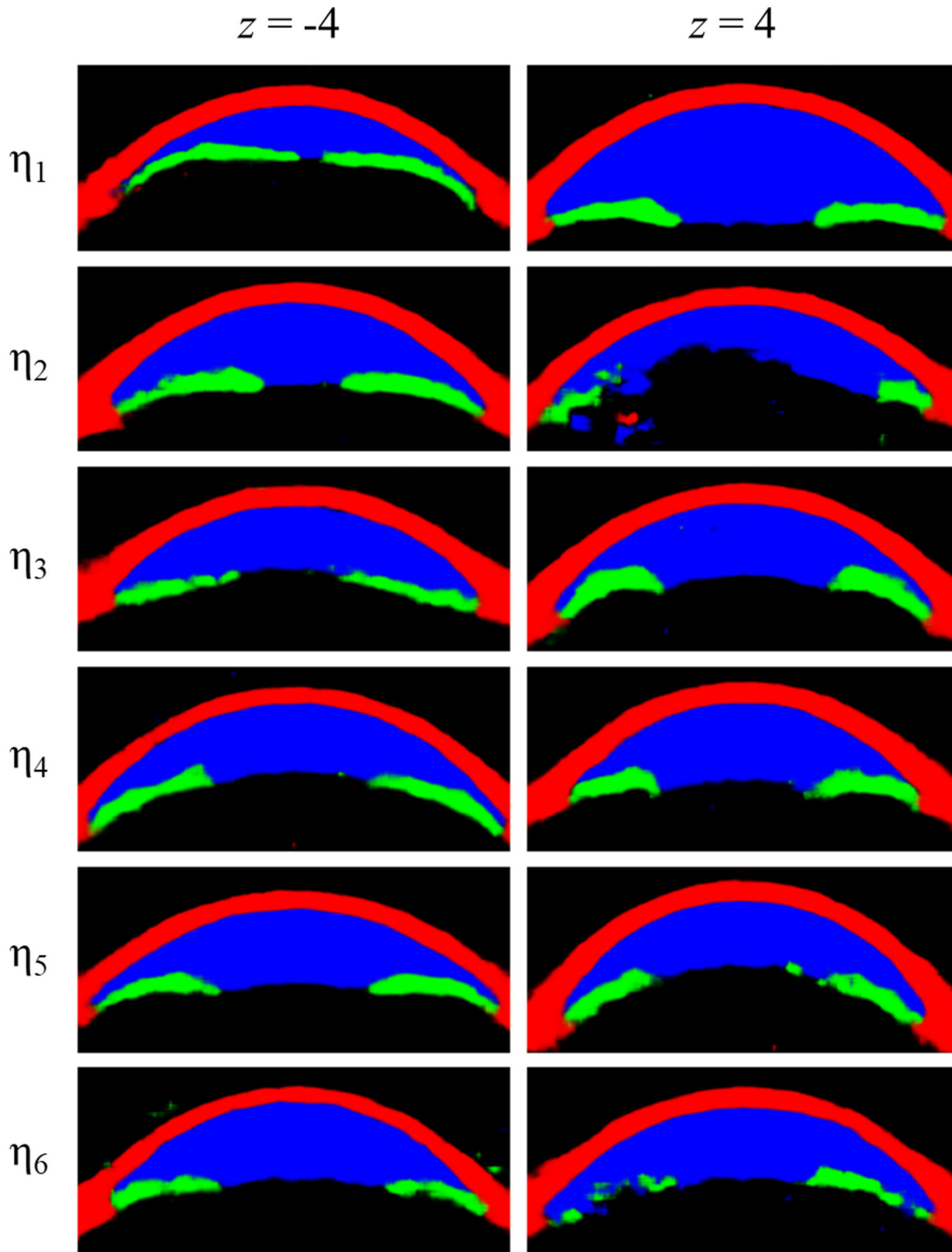


Figure 5. Latent variables of convolutional VAE presented in our previous paper at more extreme z values.

current research did not apply any disease labels during the training. Although we cannot derive any conclusion regarding the association between latent variables and disease mechanisms, we can postulate that there could be some complex interactions between conventional measurements that are not evident with conventional measurements but could be captured with deep learning techniques.

Compared to the convolutional VAE model we have presented in a previous study, our new COPIT-based model has shown similar but less blurred reconstructed images but vastly improved latent space disentanglement such that (1) every latent variable represents a unique feature, (2) features are easier to comprehend, and (3) other components of corneosclera, iris, and the lens remain relatively stable whereas the feature

of the focus changes dramatically (Figs. 2 and 3). Such encouraging results could be achieved by implementing a design specifically tailored for the application. First, we have implemented sequential training, which provides some characteristics: (1) the latent variables are ordered such that the variance explained can be expected to be the largest for the first latent variable and decrease afterward, (2) the addition of cosine similarity to the loss function encourages latent factors are dissimilar, and (3) resulting latent space is easier to interpret to human eyes. Second, in our new model, every layer could be configured to have a different design. For example, we have separated the asymmetrical layer from the symmetrical layers, which can offer a clear advantage given that physiologically, no eye is symmetrical. Hence, we expect that separating the asymmetry will help reduce confounding and enhance the extraction of clinically meaningful features in the symmetrical layers. Besides horizontal symmetry, other restrictions and transformations such as affine transformation or thin-plate spline transformation are relatively easy to incorporate and can be applied on a per-layer basis if required.

Another useful characteristic of our new model compared to the convolutional VAE is better stability at extreme z values. This can be useful because scaling up the values of the latent variables can enhance the subtle changes to improve interpretability. However, after a certain point, the generated image becomes unnatural, limiting the range of usable z values. For example, at z values of -4 and 4 , which have been used in Figure 3, latent variables in the convolutional VAE generate somewhat broken images (Fig. 5).

We believe that in the near future, deep learning techniques will be more commonly applied in the field of glaucoma, including AS-OCT image analysis. Many deep learning techniques involve the dimension reduction stage, which is related to the latent space. For example, deep clustering or longitudinal analyses can use latent space. For such strategies to be more effective, an understanding of latent structure is essential. We hope our study could promote understanding of the latent structure of AS-OCT images and provide a basis for future deep learning studies.

AS-OCT poses an important limitation to our analysis. Inadequate tissue resolution makes it difficult to delineate the exact border between the cornea and the iris and the location of the scleral spur in certain eyes with a closed angle. AS-OCT has limited penetration, limiting visualization of the posterior surface of the iris and the ciliary body, whereas in some eyes without cataracts, the lens is so transparent that some parts of the anterior capsule are not visualized. We expect those limitations to be overcome

with newer technologies. The model also has limitations inherent in generative models, including a requirement for manual tailoring of the hyperparameters. In the process of tailoring, knowledge of developers about the subject, AS-OCT images, in our case, get involved. Because there is no universally accepted feature to qualitatively assess the degree of disentanglement for AS-OCT images, the process is highly subjective. The same limitation applies when interpreting the results, especially latent variables that capture complex interactions of various features that are difficult to interpret.

Both training and validation data were derived from the same dataset, which does not include various ethnic groups, and for a machine learning study, the sample size is small. Hence, we expect the generalizability of our analysis to be limited, and the results should be assumed to be dependent on a specific dataset we have used. Comparisons between PAC and PACG eyes have all the limitations inherent in the retrospective design of the study.

Nonetheless, we have shown that an unsupervised neural network can achieve good results for the analysis of the latent structure of the AS-OCT. Also, our results suggest that latent space analysis can be useful for capturing a combination of features not readily represented with conventional measurements due to their complex interactions.

Acknowledgments

Supported by a grant (2018-0500) from the Asan Institute for Life Sciences, Asan Medical Center, Seoul, South Korea.

Disclosure: **K. Shon**, (N); **K.R. Sung**, (N); **J. Kwak**, (N); **J.Y. Lee**, (N); **J.W. Shin**, (N)

References

1. Triolo G, Barboni P, Savini G, et al. The use of anterior-segment optical-coherence tomography for the assessment of the iridocorneal angle and its alterations: update and current evidence. *J Clin Med*. 2021;10:231.
2. Baek S, Sung KR, Sun JH, et al. A hierarchical cluster analysis of primary angle closure classification using anterior segment optical coherence tomography parameters. *Invest Ophthalmol Vis Sci*. 2013;54:848–853.

3. Kwon J, Sung KR, Han S, et al. Subclassification of primary angle closure using anterior segment optical coherence tomography and ultrasound biomicroscopic parameters. *Ophthalmology*. 2017;124:1039–1047.
4. Han S, Sung KR, Lee KS, et al. Outcomes of laser peripheral iridotomy in angle closure subgroups according to anterior segment optical coherence tomography parameters. *Invest Ophthalmol Vis Sci*. 2014;55:6795–6801.
5. Lee Y, Sung KR, Na JH, et al. Dynamic changes in anterior segment (AS) parameters in eyes with primary angle closure (PAC) and PAC glaucoma and open-angle eyes assessed using as optical coherence tomography. *Invest Ophthalmol Vis Sci*. 2012;53:693–697.
6. Kwon J, Sung KR, Han S. Long-term changes in anterior segment characteristics of eyes with different primary angle-closure mechanisms. *Am J Ophthalmol*. 2018;191:54–63.
7. Moghimi S, Torkashvand A, Mohammadi M, et al. Classification of primary angle closure spectrum with hierarchical cluster analysis. *PLoS One*. 2018;13(7):e0199157.
8. Xu BY, Chiang M, Pardeshi AA, et al. Deep neural network for scleral spur detection in anterior segment OCT images: the Chinese American eye study. *Transl Vis Sci Technol*. 2020;9(2):1–10.
9. Wanichwecharungruang B, Kaothanthong N, Patanapongpaiboon W, et al. Deep learning for anterior segment optical coherence tomography to predict the presence of plateau iris. *Transl Vis Sci Technol*. 2021;10(1):1–10.
10. Fu H, Baskaran M, Xu Y, et al. A deep learning system for automated angle-closure detection in anterior segment optical coherence tomography images. *Am J Ophthalmol*. 2019;203:37–45.
11. Hao H, Zhao Y, Yan Q, et al. Angle-closure assessment in anterior segment OCT images via deep learning. *Med Image Anal*. 2021;69:101956.
12. Pham TH, Devalla SK, Ang A, et al. Deep learning algorithms to isolate and quantify the structures of the anterior segment in optical coherence tomography images. *Br J Ophthalmol*. 2021;105:1231–1237.
13. Bollen KA. Latent variable in psychology and social sciences. *Ann Rev Psychol*. 2002;53:605–634.
14. Turhan CG, Bilge HS. Recent trends in deep generative models: a review. In: *2018 3rd International Conference on Computer Science and Engineering (UBMK)*. IEEE; 2018:574–579.
15. Shon K, Sung KR, Kwak J, et al. Development of a β -variational autoencoder for disentangled latent space representation of anterior segment optical coherence tomography images. *Transl Vis Sci Technol*. 2022;11(2):11.
16. Jaderberg M, Simonyan K, Zisserman A, et al. Spatial transformer networks. *Adv Neural Inf Process Syst*. 2015;2015:2017–2025.
17. Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. *IEEE Access*. 2015;9:16591–16603.
18. Jiang W, Sun W, Tagliasacchi A, et al. Linearized multi-sampling for differentiable image transformation. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2019:2988–2997.
19. Kingma DP, Welling M. Auto-encoding variational bayes. *2nd International Conference on Learning Representations, ICLR 2014 - Conference Track Proceedings*. 2014;(MI):1–14.
20. Kingma DP, Welling M. Auto-encoding variational Bayes. *2nd International Conference on Learning Representations, ICLR 2014 - Conference Track Proceedings*. 2014:1–14.
21. van Dongen S, Enright AJ. Metric distances derived from cosine similarity and Pearson and Spearman correlations. arXiv preprint [arXiv:1208.3145](https://arxiv.org/abs/1208.3145). Accessed April 4, 2012.
22. Higgins I, Matthey L, Pal A, et al. Beta-VAE: learning basic visual concepts with a constrained variational framework. Available at: <https://openreview.net/forum?id=Sy2fzU9gl>.
23. Sakata LM, Lavanya R, Friedman DS, et al. Assessment of the scleral spur in anterior segment optical coherence tomography images. *Arch Ophthalmol*. 2008;126:181–185.
24. Schindelin J, Arganda-Carreras I, Frise E, et al. Fiji: an open-source platform for biological-image analysis. *Nat Methods*. 2012;9:676–682.
25. Lee KS, Sung KR, Kang SY, et al. Residual anterior chamber angle closure in narrow-angle eyes following laser peripheral iridotomy: anterior segment optical coherence tomography quantitative study. *Jpn J Ophthalmol*. 2011;55:213–219.
26. Lee Y, Sung KR, Na JH, et al. Dynamic changes in anterior segment (AS) parameters in eyes with primary angle closure (PAC) and PAC glaucoma and open-angle eyes assessed using as optical coherence tomography. *Invest Ophthalmol Vis Sci*. 2012;53:693–697.
27. Bookstein FL. Principal Warps: Thin-Plate Splines and the Decomposition of Deformations. *IEEE Trans Pattern Anal Mach Intell*. 1989;11:567–585.