







OPEN ACCESS

Reproducibility of deep learning based scleral spur localisation and anterior chamber angle measurements from anterior segment optical coherence tomography images

Peng Liu ^{1,2}, Risa Higashita,^{2,3} Philip Yawen Guo,⁴ Keiichiro Okamoto,³ Fei Li,⁵ Anwell Nguyen ⁶, Rei Sakata,⁷ Lixin Duan,¹ Makoto Aihara,⁷ Shan Lin,^{6,8} Xiulan Zhang ⁵, Christopher Kai-Shun Leung ⁴, Jiang Liu^{2,9,10,11}

► Additional supplemental material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/bjophthalmol-2021-319798>).

For numbered affiliations see end of article.

Correspondence to

Professor Jiang Liu, Department of Computer Science and Engineering, Southern University of Science and Technology, Shenzhen, China; liuj@sustech.edu.cn and Dr Christopher Kai-Shun Leung, Department of Ophthalmology & Visual Sciences, The Chinese University of Hong Kong, Hong Kong, China; cksleung.cuhk@gmail.com

PL and RH contributed equally.

Received 3 June 2021
Accepted 24 December 2021
Published Online First
28 January 2022



© Author(s) (or their employer(s)) 2023. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

To cite: Liu P, Higashita R, Guo PY, et al. *Br J Ophthalmol* 2023;**107**:802–808.

ABSTRACT

Aims To apply a deep learning model for automatic localisation of the scleral spur (SS) in anterior segment optical coherence tomography (AS-OCT) images and compare the reproducibility of anterior chamber angle (ACA) width between deep learning located SS (DLLSS) and manually plotted SS (MPSS).

Methods In this multicentre, cross-sectional study, a test dataset comprising 5166 AS-OCT images from 287 eyes (116 healthy eyes with open angles and 171 eyes with primary angle-closure disease (PACD)) of 287 subjects were recruited from four ophthalmology clinics. Each eye was imaged twice by a swept-source AS-OCT (CASIA2, Tomey, Nagoya, Japan) in the same visit and one eye of each patient was randomly selected for measurements of ACA. The agreement between DLLSS and MPSS was assessed using the Euclidean distance (ED). The angle opening distance (AOD) of 750 µm (AOD750) and trabecular-iris space area (TISA) of 750 µm (TISA750) were calculated using the CASIA2 embedded software. The repeatability of ACA width was measured.

Results The mean age was 60.8±12.3 years (range: 30–85 years) for the normal group and 63.4±10.6 years (range: 40–91 years) for the PACD group. The mean difference in ED for SS localisation between DLLSS and MPSS was 66.50±20.54 µm and 84.78±28.33 µm for the normal group and the PACD group, respectively. The span of 95% limits of agreement between DLLSS and MPSS was 0.064 mm for AOD750 and 0.034 mm² for TISA750. The respective repeatability coefficients of AOD750 and TISA750 were 0.049 mm and 0.026 mm² for DLLSS, and 0.058 mm and 0.030 mm² for MPSS.

Conclusion DLLSS achieved comparable repeatability compared with MPSS for measurement of ACA.

INTRODUCTION

Primary angle-closure glaucoma (PACG) is one of the leading causes of irreversible visual impairment, accounting for 50% of bilateral glaucoma blindness worldwide. By 2040, the number of global patients with PACG will increase by 58.4% over 2013, reaching 32.04 million.¹ The assessment of the anterior chamber angle (ACA) dimensions is critical to the detection of angle closure, which is a condition conventionally evaluated with gonioscopy.

However, gonioscopy assessment is subjective with low repeatability.^{2–3} Another approach to evaluate the ACA is anterior segment optical coherence tomography (AS-OCT) imaging. AS-OCT provides quantitative assessment of the ACA such as the trabecular-iris space area (TISA) and angle opening distance (AOD). Since the ACA dimensions are typically measured from the scleral spur (SS), localisation of the SS is critical for reliable measurements of the ACA. However, the inconsistency of manual labelling of the SS in AS-OCT images would lead to low test–retest variability of ACA parameters. Therefore, a reliable means to locate the SS can significantly improve the repeatability of ACA measurements.⁴

Machine learning has been shown to take an important role in many clinical applications.^{5–7} A number of studies have applied deep learning or machine learning algorithms (eg, support vector machine) to determine whether the ACA is closed.^{8–12} These studies have demonstrated high diagnostic performance of machine learning, and deep learning models in the diagnosis of angle closure. However, deep learning has been criticised for being an end-to-end black-box model, which is difficult to meet the interpretability requirement for clinical use since these models only provide the final diagnosis results without proper interpretation. In this study, we developed a novel deep learning artificial intelligence (AI) method to locate the SS automatically and evaluated the repeatability of ACA measurements based on deep learning located SS (DLLSS) and manually plotted SS (MPSS).

METHODS

Participants and AS-OCT imaging

The participants were from four ophthalmology centres: The Chinese University of Hong Kong (CUHK), Tokyo University (TU), University of California, San Francisco (UCSF) and Zhongshan Ophthalmic Center (ZOC). Patients with a history of corneal scarring, corneal surgery or active infection of the eye were excluded. Demographic information, including patient age, gender and examination findings, were collected.

Healthy subjects have normal intraocular pressure (IOP), open angles on darkroom gonioscopy

for both eyes and no evidence of ocular abnormalities except refractive errors and visually insignificant cataract. Gonioscopic angle closure was defined as having invisible posterior trabecular meshwork for $\geq 180^\circ$ on darkroom gonioscopy; PACS had gonioscopic angle closure without a history of IOP > 21 mm Hg, peripheral anterior synechiae (PAS) or glaucoma (ie, abnormal visual field tests or glaucomatous retinal nerve fibre layer (RNFL)/optic disc changes); primary angle closure (PAC) had gonioscopic angle closure with IOP > 21 mm Hg in ≥ 2 consecutive visits or PAS; PACG had gonioscopic angle closure and glaucoma (ie, narrowed neuroretinal rim and thinned RNFL with or without visual field defects). PACD includes PACS, PAC and PACG.¹³ The inclusion criteria were phakic eyes, best corrected visual acuity better than 20/40, the spherical equivalent ≥ -6.0 D and no history of intraocular surgery or laser procedure (eg, laser iridotomy).

AS-OCT images were obtained with the AS-OCT instrument CASIA2 (Tomey Corporation, Nagoya, Japan). Eighteen B-scans evenly spaced over 360° were taken for each eye. The training and validation dataset was provided by ZOC, which contained 8658 AS-OCT images from 481 eyes (377 eyes for training and 104 eyes for validation). The test dataset contained 5166 AS-OCT images of 171 eyes with PACD and 116 healthy eyes with open angles recruited from four ophthalmology clinics (CUHK, TU, UCSF and ZOC). Each eye underwent AS-OCT examination twice in the same visit, named Dark1 and Dark2 dataset in this study. Participants were consecutively recruited between 1 April 2017 and 31 December 2019. All B-scans in train, validation and test set were examined for scan quality, but the low-quality images, for example, blurred images, artefacts of motion/blink, etc were excluded. To evaluate the 36 angles, we excluded the eyes with eyelids obscuring the ACA more than one B-scan or with indiscernible SS for ≥ 5 angle meridians. For eyes with indiscernible SS < 5 but ≥ 1 angle meridians, the SS was localised with reference to the visible SS in the adjacent B-scan that was 10° or 20° apart. We further check that after scan quality check, each eye should have at least two available captures for the following reproducibility analysis procedure. For eyes that have more than one valid scan, according to the capture time, we used the first scan of each eye as the Dark1 dataset and the second scan as the Dark2 dataset. Total 43 eyes were excluded in our study. One human expert grader masked to the identities and examination results of the participants, marked all the 36 SS in 18 scans per eye; these labels of SS locations were considered the reference standard.

Deep learning SS localisation

We developed a two-stage SS localisation network to automate the localisation of SS. Our two-stage SS localisation network also addressed the quantisation errors of heatmap-based methods¹⁴ to allow more reliable localisation of the SS. The framework of the neural network design is illustrated in online supplemental figure 1. The two-stage SS localisation network has adopted a coarse to fine strategy. The first stage S1 was designed to estimate the location of the SS roughly from a compressed AS-OCT image. The second stage S2 was to locate the SS from an uncompressed region of interest (ROI).

U-Net¹⁵ is a widely used deep learning network in the field of medical image analysis. U-Net consists of downsampling and upsampling convolutional blocks and skip connections between the convolutional blocks. The downsampling and upsampling convolutional blocks are convolution layers and transposed convolution layers followed by the maxpooling layer. To mitigate

the influence of information loss caused by downsampling, we applied the skip connections to concatenate the downsampled features and upsampled features on the same level. In this study, we used the original U-Net structure, which consists of four downsampling convolutional blocks and four upsampling convolutional blocks.

We used U-Net to generate the SS possibility map for both S1 and S2. Let $I \in \mathbb{R}^{h \times w \times c}$ be the original image which has c channels, w width in pixels and h height in pixels. We downsampled I to a lower resolution image $I_1 \in \mathbb{R}^{b_1 \times w_1 \times c}$ as the input of S1. The output of S1 is the SS heatmap $\mathfrak{S}_1 \in [0, 1]^{b_1 \times w_1 \times 1}$. \mathfrak{S}_1 locates left and right SS simultaneously. In this study, $w_1 = 800, b_1 = 800$ were used in all experiments. We obtained the predicted left and right SS coordinates P_{S1}^l and P_{S1}^r of S1 by finding the max value location of left and right peak in \mathfrak{S}_1 . P_{S1}^l and P_{S1}^r are roughly estimated SS locations due to the downsampled input data, but they are accurate enough to be used to extract the left and right ROIs, where the SS is located. The two ROIs I_2^l and I_2^r were cropped from I with the size 400×400 and centre at P_{S1}^l and P_{S1}^r , respectively. Similar to S1, an individual U-Net was used to generate the SS heatmap \mathfrak{S}_2^l and \mathfrak{S}_2^r . The difference between S1 and S2 is that S2 individually locates left and right SS. The coordinates of the SS in ROIs P_{S2}^l and P_{S2}^r are determined by the locations of the maximum value of \mathfrak{S}_2^l and \mathfrak{S}_2^r . Because I_2^l and I_2^r have the same resolution of the original image I , the effect of qualitative errors significantly reduced in the prediction of S2. The final coordinates of SS were obtained by mapping the ROIs back to I according to $P_{S2}^l, P_{S2}^r, P_{S1}^l$ and P_{S1}^r .

The L2 loss function was adopted between ground truth heatmap and the output heatmap \mathfrak{S} . The ground truth heatmap were generated with a two-dimensional (2D) Gaussian blob centred on the ground truth SS location. Similar to hourglass network,¹⁶ we use 2D Gaussian blob, which has an SD $\sigma = 1$. Deep learning model was trained up to 200 epochs with a separate training dataset, including 377 eyes (6786 images). During the training process, the validation loss was evaluated using the validation set (104 eyes and 1872 images) after each epoch and used as a reference for model selection. If the validation loss did not improve over 30 consecutive epochs, the training process was stopped. The model state where the validation loss was the lowest was saved as the final state of the model.

ACA parameters evaluation

The SS of angle location from all subjects was manually plotted by a single observer (KO) on the CASIA2 software (V3C.35, Tomey Corporation, Nagoya, Japan), and the ACA parameters are automatically calculated. We used degree to present results, and the definition of angle degree was shown in online supplemental figure 2. In this work, the clinical parameters (online supplemental figure 3) included¹⁷:

1. Angle opening distance 750 (AOD750): the perpendicular distance from the cornea at 750 μm from the SS to the anterior iris surface.
2. Trabecular-iris space area 750 (TISA750): the trapezoidal area bounds by the AOD750 and the perpendicular distance from the SS to the anterior iris surface.

Statistical analysis

The deep learning modules were developed with PyTorch library, V.1.4. Statistical analyses were performed with Python Scipy

package (Python V.3.7, Scipy V.1.2.1), SPSS software (statistics subscription, build 1.0.0.1406) and MedCalc software (V.19.4.1). We evaluated the difference between DLLSS (X_{DLLSS}, Y_{DLLSS}) and MPSS (X_{MPSS}, Y_{MPSS}) by calculating the Euclidean distance (ED)²: $ED = \sqrt{(X_{DLLSS} - X_{MPSS})^2 + (Y_{DLLSS} - Y_{MPSS})^2}$. For the evaluation of the repeatability and reproducibility of ACA measurements, we compared the ACA measurements between (1) DLLSS-based and MPSS-based measurements in Dark1, (2) DLLSS-based parameters in Dark1 and Dark2 and (3) MPSS-based parameters in Dark1 and Dark2 with a two-tailed t-test.¹⁸ We used Bland-Altman plots and repeatability coefficients to analyse parameters in all mentioned pairs before. The absolute difference of ACA measurements (including AOD750 and TISA750) between Dark1 and Dark2 was calculated as [*measurements on Dark1 - measurements on Dark2*]. The mean absolute difference of measurements between Dark1 and Dark2 based on DLLSS and based on MPSS was compared with two-tailed t-test. A p value of <0.05 was considered statistically significant for all comparisons.

Post hoc analysis of sample size

We reversely used the sample size estimation formula to check the precision of our SS localisation model powered by our samples size. There were 287 eyes included in the test dataset, that is, 10 332 SSs (36 angles for each eye). Based on our sample size and largest SD of the difference between DLLSS and MPSS associated with X and Y coordinates (74.07 μm for Y coordinate), our study was powered to locate the SS with a difference of 2.362 μm at α=0.05 and β=0.2. The detailed calculation was provided in online supplemental material 1. For the CASIA2 image resolution, 2.362 μm is much smaller than the width and height (8 μm) of one pixel in the original AS-OCT images. Therefore, we consider the sample size in this study was enough.

RESULTS

The demographics are given in table 1. A total of 287 eyes of 287 participants were included in the test dataset. The mean age was 60.8±12.3 years (range: 30–85 years) for the normal group (116 eyes: 87 Asians, 17 Caucasians, 2 African Americans, 1 native

	Normal (mean±SD)	PACD (mean±SD)	P value* (normal vs PACD)
Train set			
No. of subjects	284	86	–
Age (years)	40.8±16.9	57.7±10.5	<0.001
Gender (M/F)	108/176	27/59	–
Validation set			
No. of subjects	81	22	–
Age (years)	41.2±14.9	58.5±8.5	<0.001
Gender (M/F)	33/48	11/11	–
Test set			
No. of subjects	116	171	–
Age (years)	60.8±12.3	63.4±10.6	0.062
Gender (M/F)	47/69	48/123	–
AOD750 (mm)	0.385±0.193	0.129±0.079	<0.001
TISA750 (mm ²)	0.176±0.087	0.056±0.039	<0.001

*Independent t-test.
AOD, angle opening distance; PACD, primary angle-closure disease; TISA, trabecular iris space area.

Table 2 SS location difference between DLLSS and MPSS in Dark1 dataset

Angle	ED all (mean±SD, μm) (N=287)	ED normal (mean±SD, μm) (N=116)	ED PACD (mean±SD, μm) (N=171)
Mean	77.39±26.94	66.50±20.45	84.78±28.33
0°	63.91±47.90	53.21±33.34	71.17±54.56
30°	81.11±58.22	74.10±54.73	85.87±60.17
60°	98.60±102.51	78.39±59.37	112.31±121.76
90°	98.73±85.95	84.99±76.93	108.05±90.61
120°	80.44±67.71	67.72±52.13	89.06±75.43
150°	69.66±51.67	63.10±42.04	74.11±56.98
180°	63.73±50.01	56.96±42.88	68.32±53.95
210°	70.23±56.47	54.87±38.04	80.64±64.15
240°	84.07±62.41	70.70±50.69	93.14±67.90
270°	91.33±80.59	67.64±52.46	107.40±91.77
300°	78.74±63.63	59.70±46.56	91.66±70.22
330°	66.81±57.52	58.55±40.44	72.42±66.19

DLLSS, deep learning located scleral spur; ED, Euclidean distance; MPSS, manually plotted scleral spur; PACD, primary angle-closure disease; SS, scleral spur.

American, 5 Hispanics and 4 unknown ethnicity) and 63.4±10.6 years (range: 40–91 years) for the PACD group (171 eyes: 144 Asians, 14 Caucasians, 6 African Americans, 3 Hispanics and 4 unknown ethnicity). There were 139 right eyes and 148 left eyes in the test set. The deep learning model was trained with a separate dataset, including 377 eyes (6786 images), and the final model state was selected with the validation set, which included 104 eyes (1872 images) from ZOC. All the participants in train and validation set are Asian. The mean age was 45.1±17.3 years (range: 7–84 years) for the training dataset (284 normal eyes, 86 PACD eyes and 7 eyes undiagnosed) and 45.0±15.5 years (range: 19–74 years) for the validation set (81 normal eyes, 22 PACD eyes and 1 eye undiagnosed). In this study, all the 36 angles of each eye were included in the model training, validation and testing. The mean values in the following tables were calculated with 36 angles, while we reported specific angle results every 30° for brevity. Full 36 angles version results were provided in online supplemental material 2.

Performance of SS localisation

Taking all eyes into consideration (n=287), the mean ED of all angles between DLLSS and MPSS was 77.39±26.94 μm in the Dark1 dataset (table 2). The mean ED was 84.78±28.33 μm in the PACD group and 66.50±20.54 μm in the normal group. The ED was smallest at 180° and 0°, and greatest at 60° and 90°. In other words, the agreement between DLLSS and MPSS was better at the nasal and temporal quadrants than the superior and inferior quadrants.

Agreement of ACA width between DLLSS and MPSS

ACA measurements derived from DLLSS were similar to those derived from MPSS in the PACD group, although the former was generally greater than the latter in the normal group (table 3). The span of 95% limits of agreement was 0.064 mm and 0.034 mm² for AOD750 and TISA750, respectively (figure 1, left column). The R² was 0.993 and 0.991 for AOD750 and TISA750, respectively.

Repeatability of ACA width

The comparisons of ACA width between two AS-OCT scans capture in the same visit derived from DLLSS and MPSS are

Table 3 Comparisons of parameters (mean±SD) between DLLSS based and MPSS based in Dark1 dataset

AOD750 (mm)						
Angles	Normal			PACD		
	DLLSS	MPSS	P value*	DLLSS	MPSS	P value**
Mean	0.390±0.201	0.385±0.193	0.005	0.129±0.077	0.129±0.079	0.968
0°	0.434±0.250	0.433±0.242	0.803	0.150±0.091	0.148±0.093	0.297
30°	0.379±0.205	0.379±0.201	0.980	0.112±0.083	0.112±0.086	0.959
60°	0.340±0.211	0.343±0.200	0.624	0.098±0.082	0.102±0.079	0.312
90°	0.334±0.198	0.318±0.196	0.008	0.103±0.079	0.098±0.082	0.095
120°	0.327±0.194	0.318±0.193	0.023	0.086±0.075	0.085±0.078	0.688
150°	0.341±0.200	0.333±0.193	0.023	0.098±0.083	0.095±0.080	0.193
180°	0.419±0.206	0.409±0.196	0.004	0.143±0.100	0.140±0.104	0.164
210°	0.467±0.243	0.455±0.232	0.003	0.166±0.109	0.170±0.111	0.245
240°	0.381±0.205	0.375±0.185	0.308	0.124±0.104	0.128±0.104	0.229
270°	0.388±0.233	0.385±0.228	0.498	0.124±0.103	0.116±0.105	0.018
300°	0.417±0.226	0.414±0.204	0.601	0.145±0.097	0.152±0.103	0.014
330°	0.464±0.238	0.461±0.230	0.406	0.177±0.101	0.184±0.109	0.019
TISA750 (mm ²)						
Angles	Normal			PACD		
	DLLSS	MPSS	P value*	DLLSS	MPSS	P value*
Mean	0.177±0.090	0.176±0.087	0.161	0.055±0.038	0.056±0.039	<0.001
0°	0.205±0.115	0.203±0.109	0.542	0.075±0.056	0.076±0.057	0.394
30°	0.175±0.094	0.176±0.094	0.703	0.049±0.049	0.051±0.049	0.056
60°	0.142±0.094	0.147±0.093	0.084	0.034±0.038	0.035±0.037	0.217
90°	0.140±0.091	0.135±0.092	0.065	0.036±0.039	0.037±0.041	0.929
120°	0.144±0.092	0.140±0.093	0.052	0.032±0.039	0.034±0.040	0.077
150°	0.157±0.099	0.152±0.093	0.007	0.042±0.044	0.043±0.044	0.539
180°	0.200±0.098	0.196±0.093	0.047	0.066±0.054	0.065±0.057	0.446
210°	0.211±0.106	0.208±0.105	0.133	0.074±0.057	0.078±0.059	0.009
240°	0.171±0.095	0.170±0.089	0.730	0.049±0.052	0.052±0.053	0.044
270°	0.170±0.106	0.171±0.105	0.613	0.046±0.046	0.043±0.047	0.111
300°	0.191±0.099	0.191±0.093	0.893	0.063±0.052	0.068±0.054	<0.001
330°	0.218±0.105	0.217±0.101	0.786	0.084±0.057	0.089±0.060	<0.001

*Independent t-test.

AOD, angle opening distance; DLLSS, deep learning located scleral spur; MPSS, manually plotted scleral spur; PACD, primary angle-closure disease; TISA, trabecular iris space area.

summarised in online supplemental tables 1 and 2, respectively. For DLLSS, the repeatability coefficients were 0.049 mm and 0.026 mm² for AOD750 and TISA750 (figure 1, middle column), respectively. For MPSS, the repeatability coefficients were 0.058 mm and 0.030 mm² for AOD750 and TISA750 (figure 1, right column), respectively. The mean absolute difference of ACA width between Dark1 and Dark2 based on DLLSS was 0.0167±0.0188 mm for AOD750 and 0.0093±0.0093 mm² for TISA750 and based on MPSS was 0.0210±0.0212 mm for AOD750 and 0.0109±0.0110 mm² for TISA750. AOD750 and TISA750 showed strong association between the two AS-OCT scans (figure 2, middle column). The R² was 0.986 and 0.982 for AOD750 and TISA750, respectively, for DLLSS and 0.978 and 0.974, respectively, for MPSS.

DISCUSSION

Gonioscopy is the current gold standard for the assessment and diagnosis of angle closure. Gonioscopy remains useful for evaluation of pigment and vascularisation of the angle, along with assessment of PAS with indentation. However, gonioscopy is a subjective and contact assessment, while AS-OCT provides an objective and quantitative approach to evaluate the ACA. Currently, the measurements of ACA dimensions rely on the localisation of the SS, which still requires subjective judgement.

In this study, we developed a deep learning model to localise the SS automatically in AS-OCT images. We demonstrated the reproducible SS localisation, and the reproducible ACA width measurements (AOD750 and TISA750) could be achieved with CASIA2 AS-OCT images by our deep learning model.

We compared the difference in SS localisation between DLLSS and MPSS by calculating the mean ED (77.39±26.94 µm) of all angles in the Dark1 dataset. Xu *et al*² reported that the intra-grader variability of SS localisation on their test dataset with 921 images (698 open angle images and 223 images with PACD) was 73.08±52.06 µm (normal eyes: 67.37±56.82 µm, and eyes with PACD: 94.44±67.70 µm), and the inter-grader variability was 97.34±73.29 µm (normal eyes 96.10±70.37 µm, eyes with PACD 101.21±81.81 µm). We reimplemented the method of Xu *et al*'s,² trained and tested the ResNet-50 model with our dataset; the detailed results were presented in online supplemental table 3. Including all eyes, our model had higher repeatability compared with the study by Xu *et al*² in Dark1 dataset (for open angle, ours 66.50±20.54 µm vs Xu *et al*'s 108.68±91.08 µm; for no-open angle, ours 84.78±28.33 µm vs Xu *et al*'s 125.96±87.94 µm). The data in Xu *et al*² were collected by SS1000 (CASIA1) in only one instrument, which is the first-generation AS-OCT of Tomey Corporation, Japan. Our data were collected by CASIA2, the second generation, from five

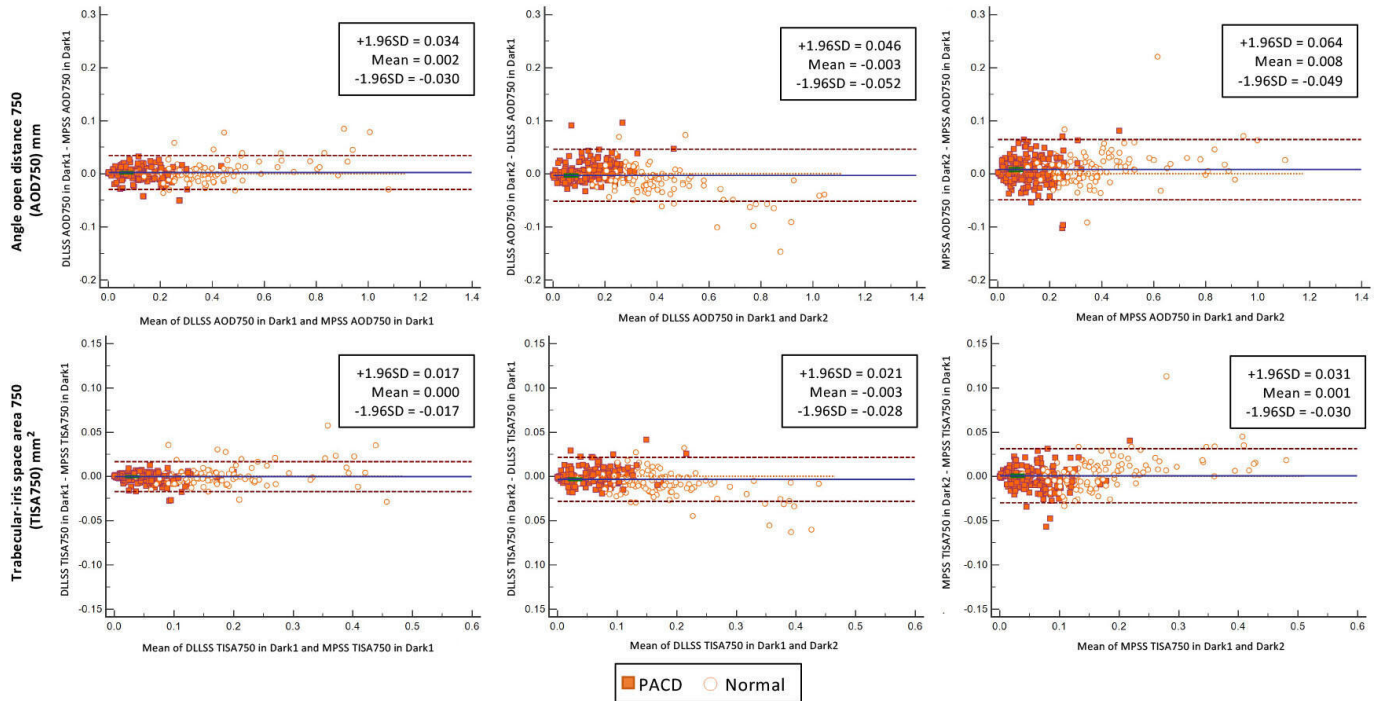


Figure 1 Bland-Altman plot of AOD750 and TISA750 with DLLSS and MPSS in Dark1 and Dark2 dataset. AOD, angle opening distance; DLLSS, deep learning located scleral spur; MPSS, manually plotted scleral spur; PACD, primary angle-closure disease; TISA, trabecular-iris space area.

different instruments of four clinical sites. Even if the resolution of CASIA2 image is better than of CASIA1 slightly, the performance of Xu *et al*² proposed that ResNet-50 model was significantly poor on our data. ResNet-50 model is a regression-based model, which is less robust in key points localisation tasks.¹⁴ The downsampling processing greatly affects the performance of deep learning models. The experimental results verified the robustness of our heatmap-based model in the Dark1 dataset. Our two-stage model addressed the effect from downsampling and greatly elevated the SS localisation accuracy. In addition, the agreement between DLLSS and MPSS was better at the nasal and

temporal quadrants than the superior and inferior quadrants, which is similar to the result of Cumba *et al*.⁴

Our deep learning model prediction errors are with 76% (7818/10 332) in 100 µm compared with manual SS points and with 90% (9248/10 332) in 150 µm, which were similar to Xu *et al*'s² (75% fell within 100 µm and 90% within 150 µm in their test dataset). Cumba *et al*⁴ reported that the SS location within 80 µm for both X-axis and Y-axis supports a 95% CI limit of TISA750. Their study reported interobserver and intraobserver reproducibility of 77% and 84% within 80 µm for both the X-axis and Y-axis. Our deep learning model

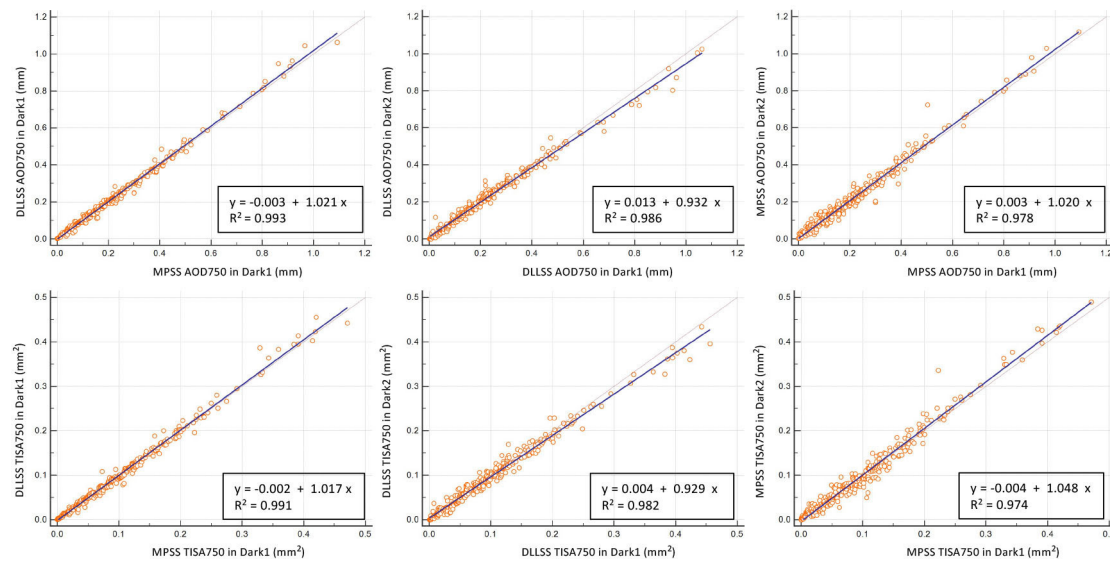


Figure 2 Correlation analysis of AOD750 and TISA750 with DLLSS and MPSS in Dark1 and Dark2 dataset. AOD, angle opening distance; DLLSS, deep learning located scleral spur; MPSS, manual plotted scleral spur; TISA, trabecular-iris space area.

SS reproducibility of 79% (3317/4176) within 80 μm for both the X-axis and Y-axis with normal angles and 68% (4177/6156) with closure angles. Our study was conducted on a large-scale dataset of multiple ethnicities, the images are from four clinical centres and obtained with five AS-OCT machines. Our data diversity and image quality variety and number of participants exceed the data used by Xu *et al*² and Cumba *et al*.⁴ Pham *et al*¹⁹ developed a deep learning model for both the SS localisation and the anterior segment structures segmentation, and the angle width measurements were evaluated. Pham *et al*¹⁹ reported that the SS locations by convolutional neural network (CNN) model was similar to the graders. Their results are consistent with ours. The mean error for SS localisation of our deep learning model (evaluated with ED) was $77.39 \pm 26.94 \mu\text{m}$. Pham *et al*¹⁹ reported strong correlation between CNN and human grader SS locations with intracorrelation coefficient (ICC), and the ED was not reported. The reproducibility of angle width measurements were evaluated with ICC in¹⁹ and with independent t-test in our study. Both reproducibility of the measurements obtained with the pipeline in¹⁹ and measurements obtained with our DLLSS were good. Instead of using self-developed software, we evaluated the measurements with CASIA2 embedded software, which could be a better reference for the CASIA2 users.

A trained and fixed deep learning model outputs the same SS location when fed the same AS-OCT image, that is, the deep learning model has no intra-grader variability when facing only one dataset. We used two AS-OCT datasets capture in the same visit (Dark1 and Dark2) to evaluate the repeatability of DLLSS-based and MPSS-based ACA parameters (AOD750 and TISA750). In such short capture interval, the AS-OCT images of one eye would be slightly different (eg, pupil diameter changes), while the ACA parameters should still be consistent. To access how the pupil diameter changes would impact the ACA measurements (AOD750 and TISA750), a linear regression analysis was conducted, and indicated that the pupil diameter differences between Dark1 and Dark2 fail to support a statistically significant change in ACA measurements, as presented in online supplemental material 3. The repeatability of ACA parameters with DLLSS (0.049 mm and 0.026 mm² for AOD750 and TISA750) was better than ACA parameters based on MPSS (0.058 mm and 0.030 mm² for AOD750 and TISA750). The mean absolute difference of ACA width between Dark1 and Dark2 based on DLLSS was statistically significant and smaller than that based on MPSS ($0.0167 \pm 0.0188 \text{ mm}$ vs $0.0210 \pm 0.0212 \text{ mm}$, $p=0.003$ for AOD750, and $0.0093 \pm 0.0093 \text{ mm}^2$ vs $0.0109 \pm 0.0110 \text{ mm}^2$, $p=0.018$ for TISA750). Our DLLSS was more reproducible compared with MPSS and enables the reproducible evaluation of ACA parameters. Additionally, our deep learning model can localise the SS in 489 ms on Windows 10 machine with Nvidia GeForce RTX 2080 8G graphic card, 32 GB RAM, Intel Core i7-9700 CPU, which was much faster than a human observer.

Our study is limited by examining individual OCT B-scans without taking the locations of SS in other B-scans in the same eye into consideration. Including three-dimensional volume structural information of the anterior segment may further improve the SS localisation performance. Another limitation is that only one grader could lead to biased SS localisation. In Fu *et al*,²⁰ the reference SS locations were the mean position of multiple graders. However, the mean SS locations could introduce confusions into the ground truth. In our study, the reference SS locations were provided by one experienced human grader, KO, who had been trained by two experts of ophthalmology, and the SS locations verified by RH. This one of our limitations in this study, and how to well use the labels from multiple experts, is still an important research topic in AI.

In summary, we developed a deep learning model for SS localisation and demonstrated that our deep learning model is able to locate the SS in CASIA2 AS-OCT images with high repeatability. The reproducibility of ACA parameters (AOD750 and TISA750) based on DLLSS was comparable with that based on MPSS. Deep learning model would provide a more efficient approach to localise the SS for measurement of ACA in the clinical evaluation of angle closure.

Author affiliations

¹School of Computer Science and Engineering & Shenzhen Institute for Advanced Study, University of Electronic Science and Technology of China, Chengdu, China
²Department of Computer Science and Engineering, Southern University of Science and Technology, Shenzhen, China
³Tomey Corporation, Nagoya, Japan
⁴Department of Ophthalmology and Visual Sciences, The Chinese University of Hong Kong, Hong Kong, China
⁵Zhongshan Ophthalmic Center, State Key Laboratory of Ophthalmology, Sun Yat-Sen University, Guangzhou, China
⁶Department of Ophthalmology, University of California San Francisco, San Francisco, California, USA
⁷Department of Ophthalmology, The University of Tokyo, Tokyo, Japan
⁸Glaucoma Center of San Francisco, San Francisco, California, USA
⁹Cixi Institute of Biomedical Engineering, Chinese Academy of Sciences, Ningbo, China
¹⁰Research Institute of Trustworthy Autonomous Systems, Southern University of Science and Technology, Shenzhen, China
¹¹Guangdong Provincial Key Laboratory of Brain-inspired Intelligent Computation, Department of Computer Science and Engineering, Southern University of Science and Technology, Shenzhen, China

Contributors PL and RH designed the study. RH, PYG, FL, AN, RS, MA, SL, XZ and CK-SL collected the data. KO labelled the data. RH verified the raw dataset and labels. PL and RH developed the algorithm, analysed the data, drafted the manuscript and interpreted the results. PL, RH, LD, CK-SL and JL contributed to the manuscript revision. All authors approved the submitted version. JL is the guarantor for this study.

Funding This work was supported in part by Guangdong Provincial Department of Education (2020ZDZX3043), Guangdong Provincial Key Laboratory (2020B121201001), and Shenzhen Natural Science Fund (JCYJ20200109140820699 and the Stable Support Plan Program 20200925174052004).

Disclaimer The funders had no role in the study design, data collection, data analysis, data interpretation or writing of the report. All authors had full access to all the data and were able to verify the data. The corresponding author had the final responsibility for the decision to submit for publication.

Competing interests XZ, AN, SL and CK-SL have received support from Tomey in the form of instrument and speaker honorarium. SL is a consultant to Aerie, Iridex, Bausch & Lomb and Allergan. CK-SL has received research support in the form of research grant and instruments from Carl Zeiss Meditec, Topcon and Heidelberg Engineering, and has been on the advisory board as a member for Alcon, Novartis, Santen, Aerie and Allergan.

Patient consent for publication Consent obtained directly from patient(s).

Ethics approval This study involves human participants. All research procedures adhered to the tenets of the Declaration of Helsinki. Informed consent was obtained from all participants. The study was approved by the respective Institutional Review Boards of The Chinese University of Hong Kong (CUHK), Tokyo University (TU), University of California, San Francisco (UCSF) and Zhongshan Ophthalmic Center (ZOC). CUHK's Research Ethics Committee (Kowloon Central/Kowloon East; document number: KCKE SOP001F6a), TU's Human Research Studies (OHRs; Graduate School of Medicine and Faculty of Medicine, ID: 11517-(2)UMIN: UMIN00002754), UCSF's Human Research Protection Program (reference numbers: 12-09705 and 167661) and ZOC's Ethics Committee (ID: 2018KYPJ074).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request. The full study protocol and study data can be obtained upon request from the corresponding author. Data in this study cannot be shared publicly due to regulations of local ethical committees. Data might be made available to researchers who meet the criteria for access to confidential data and upon institutional review board's approval. Access requests can be made to professor Jiang Liu (liuj@sustech.edu.cn).

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have

been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Peng Liu <http://orcid.org/0000-0002-4441-3638>

Anwell Nguyen <http://orcid.org/0000-0002-1104-4194>

Xiulan Zhang <http://orcid.org/0000-0002-9288-0767>

Christopher Kai-Shun Leung <http://orcid.org/0000-0003-4862-777X>

REFERENCES

- 1 Tham Y-C, Li X, Wong TY, *et al.* Global prevalence of glaucoma and projections of glaucoma burden through 2040: a systematic review and meta-analysis. *Ophthalmology* 2014;121:2081–90.
- 2 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;9:18.
- 3 Rigi M, Bell NP, Lee DA, *et al.* Agreement between gonioscopic examination and swept source Fourier domain anterior segment optical coherence tomography imaging. *J Ophthalmol* 2016;2016:1–8.
- 4 Cumba RJ, Radhakrishnan S, Bell NP, *et al.* Reproducibility of scleral Spur identification and angle measurements using Fourier domain anterior segment optical coherence tomography. *J Ophthalmol* 2012;2012:1–14.
- 5 He K, Zhang X, Ren S. Deep residual learning for image recognition. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016: 770–8.
- 6 Chen L-C, Papandreou G, Kokkinos I, *et al.* Deeplab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE Trans Pattern Anal Mach Intell* 2018;40:834–48.
- 7 Zhou X, Wang D, Krähenbühl P. Objects as Points. arXiv:1904.07850 [cs]. Available: <http://arxiv.org/abs/1904.07850> [Accessed 20 Apr 2020].
- 8 Fu H, Xu Y, Lin S, *et al.* Angle-Closure detection in anterior segment OCT based on multilevel deep network. *IEEE Trans Cybern* 2020;50:3358–66.
- 9 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.
- 10 Hao H, Zhao Y, Fu H, *et al.* Anterior chamber angles classification in anterior segment OCT images via multi-scale regions convolutional neural networks. *Annu Int Conf IEEE Eng Med Biol Soc* 2019;2019:849–52.
- 11 Xu Y, Liu J, Cheng J, *et al.* Automated anterior chamber angle localization and glaucoma type classification in OCT images. *Annu Int Conf IEEE Eng Med Biol Soc* 2013;2013:7380–3.
- 12 Xu Y, Liu J, Tan NM, *et al.* Anterior chamber angle classification using multiscale histograms of oriented gradients for glaucoma subtype identification. *Annu Int Conf IEEE Eng Med Biol Soc* 2012;2012:3167–70.
- 13 Foster PJ, Buhrmann R, Quigley HA, *et al.* The definition and classification of glaucoma in prevalence surveys. *Br J Ophthalmol* 2002;86:238–42.
- 14 Sun X, Xiao B, Wei F. Integral human pose regression. *Proceedings of the European Conference on Computer Vision* 2018:529–45.
- 15 Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. In: *International Conference on medical image computing and computer-assisted intervention*. Springer, 2015: 234–41.
- 16 Newell A, Yang K, Deng J. Stacked Hourglass Networks for Human Pose Estimation. In: Leibe B, Matas J, Sebe N, eds. *Computer vision – ECCV 2016*. Cham: Springer International Publishing, 2016: 483–99.
- 17 Nongpiur ME, Haaland BA, Friedman DS, *et al.* Classification algorithms based on anterior segment optical coherence tomography measurements for detection of angle closure. *Ophthalmology* 2013;120:48–54.
- 18 Chan PPM, Lai G, Leung CK-S. Comparison of test-retest variabilities of anterior segment parameters measured by two anterior segment swept-source optical coherence tomography (SS-OCT) – the ANTERION and the CASIA II. *Invest Ophthalmol Vis Sci* 2020;61:3898.
- 19 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;105:1231–7.
- 20 Fu H, Li F, Sun X, *et al.* Age challenge: angle closure glaucoma evaluation in anterior segment optical coherence tomography. *Med Image Anal* 2020;66:101798.