RimNet: A Deep Neural Network Pipeline for Automated Identification of the Optic Disc Rim

Haroon Adam Rasheed, BS,[1] Tyler Davis, MS,[2] Esteban Morales, MS,[3] Zhe Fei, MS, PhD,[4,5] Lourdes Grassi, MD,[3] Agustina De Gainza, MD,[3] Kouros Nouri-Mahdavi, MD, MS,[3] Joseph Caprioli, MD[3]

Purpose: Accurate neural rim measurement based on optic disc imaging is important to glaucoma severity grading and often performed by trained glaucoma specialists. We aim to improve upon existing automated tools by building a fully automated system (RimNet) for direct rim identification in glaucomatous eyes and measurement of the minimum rim-to-disc ratio (mRDR) in intact rims, the angle of absent rim width (ARW) in incomplete rims, and the rim-to-disc-area ratio (RDAR) with the goal of optic disc damage grading.

Design: Retrospective cross sectional study.

Participants: One thousand and twenty-eight optic disc photographs with evidence of glaucomatous optic nerve damage from 1021 eyes of 903 patients with any form of primary glaucoma were included. The mean age was 63.7 (± 14.9) yrs. The average mean deviation of visual fields was −8.03 (± 8.59).

Methods: The images were required to be of adequate quality, have signs of glaucomatous damage, and be free of significant concurrent pathology as independently determined by glaucoma specialists. Rim and optic cup masks for each image were manually delineated by glaucoma specialists. The database was randomly split into 80/10/10 for training, validation, and testing, respectively. RimNet consists of a deep learning rim and cup segmentation model, a computer vision mRDR measurement tool for intact rims, and an ARW measurement tool for incomplete rims. The mRDR is calculated at the thinnest rim section while ARW is calculated in regions of total rim loss. The RDAR was also calculated. Evaluation on the Drishti-GS dataset provided external validation (Sivaswamy 2015).

Main Outcome Measures: Median Absolute Error (MAE) between glaucoma specialists and RimNet for mRDR and ARW.

Results: On the test set, RimNet achieved a mRDR MAE of 0.03 (0.05), ARW MAE of 31 (89)°, and an RDAR MAE of 0.09 (0.10). On the Drishti-GS dataset, an mRDR MAE of 0.03 (0.04) and an mRDAR MAE of 0.09 (0.10) was observed.

Conclusions: RimNet demonstrated acceptably accurate rim segmentation and mRDR and ARW measurements. The fully automated algorithm presented here would be a valuable component in an automated mRDR-based glaucoma grading system. Further improvements could be made by improving identification and segmentation performance on incomplete rims and expanding the number and variety of glaucomatous training images. *Ophthalmology Science 2023;3:100244 © 2022 by the American Academy of Ophthalmology. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).*

Glaucoma is the leading cause of irreversible blindness and the second leading cause of blindness worldwide.[1] Roughly half of glaucoma cases are undiagnosed according to population-based studies.[2,3] Early treatment preserves patient quality of life and reduces disease burden.[4] Therefore, identification of early glaucoma is a key to preventative care.

Glaucoma diagnosis and grading are performed, in part, by evaluation of the optic nerve head's neuroretinal rim of the optic disc. Metrics often include cup-to-disc ratio (CDR), minimum rim-to-disc ratio (mRDR), and the inferior > superior > nasal > temporal rule, which compares the regional width of the neuroretinal rim.[5] Recent studies have shown the advantages of mRDR compared to inferior > superior > nasal > temporal and CDR for glaucoma classification accuracy.[6]

The mRDR cannot adequately account for the degree of damage in optic discs with localized rim loss where the neuroretinal rim is noncontinuous or "incomplete". A solution can be found in the Disc Damage Likelihood Scale (DDLS) proposed by Spaeth et al.[5] The DDLS accounts for incomplete rims by measuring the angle for which a rim is absent. This is called the absent rim width (ARW). Additionally, the scale accounts for disc size which affects the significance of the mRDR or ARW.[5] It is commonly accepted and has been incorporated into eye health guidelines for optometrists and ophthalmologists.[7,8] The DDLS is limited as a diagnostic tool by the need for expert time to accurately grade images. Automated high-efficacy DDLS grading could offer a powerful screening method.

In recent years, a confluence of several factors has led to efforts in automated glaucoma diagnosis and grading. First, studies have shown that automated algorithms can offer more consistent and reliable grading than human graders.[8] Second, there has been a rapid advancement in image segmentation, image processing, and deep learning neural networks. In other fields, several neural networks outperformed human graders in image classification tasks.[9] This could allow for unprecedented accuracy in optic rim segmentation and glaucoma grading.[10] Finally, the optic disc exhibits characteristic alterations in glaucomatous patients, a prime candidate for automated segmentation and analysis. Together, these factors make automated glaucoma diagnosis and grading a possibility.

While DDLS also requires disc size analysis, automated rim segmentation with mRDR calculation for intact neuroretinal rims and ARW calculation for incomplete neuroretinal rims offers a step towards creating an efficacious, high-throughput diagnostic system for glaucomatous disc damage. Such a segmentation algorithm would need to be broadly applicable. Additionally, it would require an expansive learning capacity that could be applied to a variety of fundus images taken with different imaging modalities and with concurrent pathologies and normal variations. Convolutional neural networks offer such an approach.[11]

The goal of this paper is to present a novel convolutional neural network algorithm for neuroretinal rim segmentation, automated mRDR calculation for intact rims, and ARW calculation for incomplete neuroretinal rims. This neural network algorithm offers an important step towards building an automated DDLS screening tool.

Methods

The study adhered to the tenets of the Declaration of Helsinki, was approved by the University of California, Los Angeles Human Research Protection Program, and conformed to the Health Insurance Portability and Accountability Act policies.

Dataset

Optic disc photographs were taken from the UCLA Stein Eye Glaucoma database. The images were of varied magnifications and taken from slides and 3 different digital fundus cameras. All cameras were visible light cameras. No infrared, laser scanning, red-free, autofluorescence, or hand-held smartphone-based cameras were used. Slide films were scanned and digitized at a third-party location.

The enrolled images met the following inclusion and exclusion criteria as deemed by 2 board-certified glaucoma specialists (L.G. and A.D.G.). Inclusion criteria include: (1) evidence of glaucomatous damage in the posterior pole; (2) images had to be in focus, with discernible posterior pole and vasculature details. Exclusion criteria were concurrent nonglaucoma disease including optic neuritis, optic disc neovascularization, and vitreous hemorrhage that would impair visualization of the posterior pole. Globally, it was ensured that the full spectrum of glaucomatous damage, from early-stage intact neuroretinal rims to late-stage incomplete rims, were included while abiding by the inclusion and exclusion criteria. Figure 1 shows the mRDR distributions of our train, validate, and test set. The neuroretinal rim and optic cup were then manually segmented by 1 of 3 glaucoma specialists with a smart tablet and the image editing program GIMP. These masks were used as ground truth. The diagnostic categories for patients are shown in Table 1.

RimNet Model and Hyperparameter Architecture

A deep learning model for rim segmentation was developed as the centerpiece of the RimNet pipeline. The model was developed with Python 3.9.7.[12] Libraries used include TensorFlow 2.6.0, Segmentations Models 1.0.1, Keras Tuner 1.04, OpenCV Python 4.5.3, Numpy 1.19.5, Scipy 1.7.1, and Scikit-learn 0.24.2.[13–16]

Optimizing the deep learning model requires a careful choice of model architecture and hyperparameters. The choice of hyperparameters can greatly influence the prediction speed, processing requirements, and accuracy of a neural network model.[17] These hyperparameters include the decoder, learning rate, optimizer, and loss function as shown in Table 2. The optimal combination of these parameters is task-dependent. While trial and error has been used in the past, newer architecture search techniques allow for the rapid evaluation of combinations of hyperparameters with the goal of optimizing a selected metric.[17]

To narrow the search space, an encoder of InceptionV3 was chosen based on literature review and computational efficiency. InceptionV3 was first published in 2015, outperforming popular encoders at the time with a fraction of the computation costs.[18] It has previously been used for medical segmentation.[19,20] Our workstation uses NVIDIA 2080 RTX Ti graphics cards. Therefore, with limited computational efficiency, the selection of InceptionV3 was appropriate.

Transfer learning with ImageNet weights was used to initialize InceptionV3. No transfer learning was done for the decoder. Augmentations were used including a 20° rotation, a 10% vertical shift, a 10% horizontal shift, a horizontal flip, a vertical flip, up to a 30% random crop, a brightness change by $\pm$ 50 units, and a contrast limited adaptive histogram equalization filter. Image down sampling was completed via a nearest neighbor algorithm. Color information was encoded using Red, Green, Blue channels with 8 bits per channel. The encoder and decoder were coupled using the Segmentation Models 1.0.1 library. The total number of trainable parameters was 29 896 979. No dropout layers were manually added.

Finally, a random search was performed using the Keras Tuner library.[16] The search parameters included the decoder, loss function, learning rate, and the optimizer. The rim Intersection over Union (IoU) was used as the segmentation metric. The full search space is documented in Table 2.

End-To-End mRDR Calculation Procedure

The mRDR, ARW, and rim-to-disc-area (RDAR) measurements are the final output of RimNet, which can be accomplished by accurate rim segmentation followed by image analysis. These 2 steps, along with preprocessing, led to the final framework for RimNet as shown in Figure 2.

The optic disc photographs were first resized to 224 $\times$ 224 with nearest neighbor interpolation in order to meet model specifications. A contrast limited adaptive histogram equalization filter was then applied to highlight distinctive features. The preprocessed image was submitted to the neural network model which generated a segmentation mask of the optic rim and cup. While a segmentation of the optic cup is not directly needed for mRDR or RDAR calculations, it was found that training the model to identify and segment the optic cup improved identification of incomplete rims and ARW calculations. Finally, the rim segmentation mask was resized to the
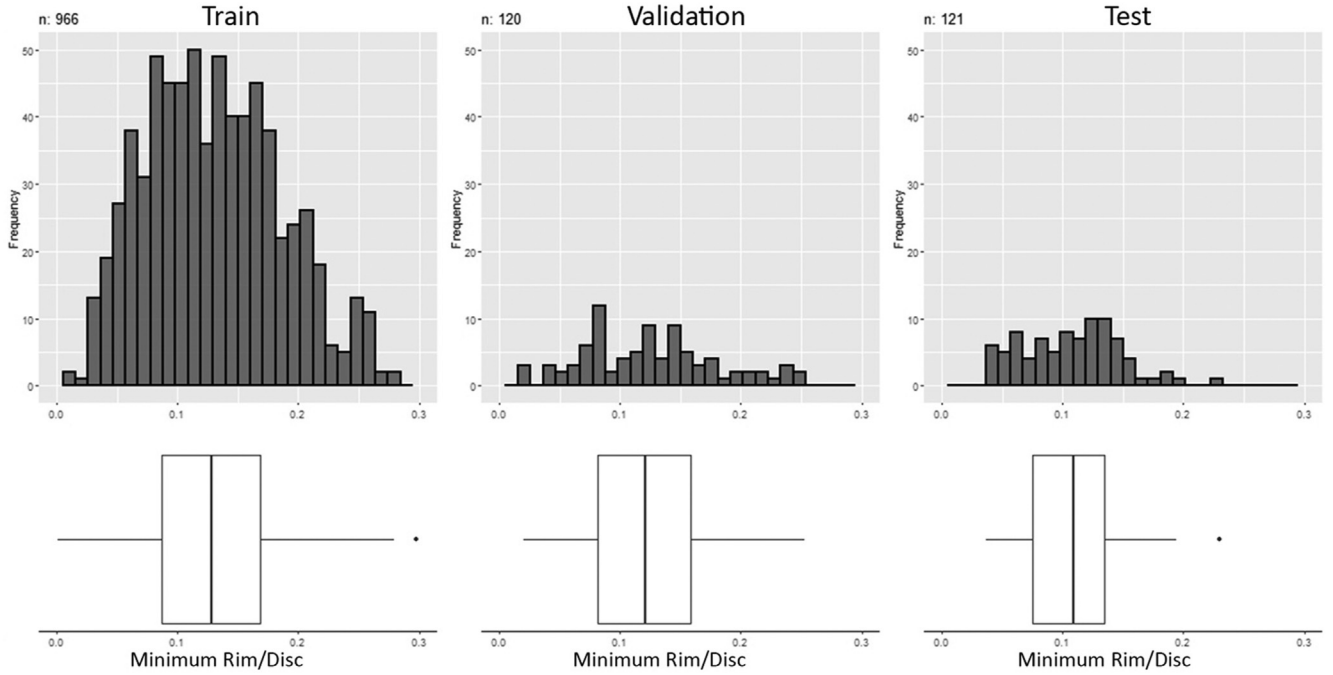
Figure 1. Distribution of minimum rim-to-disc ratios for Train, Validation, and Test Datasets. For each dataset, a frequency histogram is shown above with a box plot corresponding to the dataset below.

dimensions of the original image to allow for accurate mRDR calculation and submitted to image analysis algorithms.

For mRDR, the algorithm first identified the center of the segmented optic cup using OpenCV. Vectors were created from the center of the cup to the boundary points. Boundary points were found using OpenCV. The number of vectors depended on the number of boundary points detected in the segmented rim. The intersection between the vectors and the segmented rim was taken as the rim width. The shortest rim width was identified and, through boundary point analysis of the rim, the disc diameter was found. Hence, the mRDR was calculated by dividing the rim width by the diameter. The RDAR was calculated by dividing the number of pixels of the segmented rim by the number of pixels in optic disc.

The ARW was calculated by first applying contour hierarchies to identify shapes within the rim segmentation. We rely on the fact that intact rims will have a 'second shape' within the segmentation,

the elliptical or circular form of the optic cup. Incomplete rims will not have this second shape. If the rim is classified as broken, 360 radial segments from the center are drawn to the edge of the rim. The radial segments that do not intersect the rim are those within the "broken" segment of the neuroretinal rim. The number of radial segments within the incomplete segment is added to give the ARW, 1 radial segment for each degree. If there were 2 breaks in a neuroretinal rims, the angles were added together and reported as 1 ARW. Examples of this can be found in the neuroretinal rim segmentation shown in Figure 3.

External Validation

The Drishti-GS database is a publicly available dataset of retinal images of glaucomatous eyes with manual cup and disc segmentations.[21,22] The images were first cropped around the optic disk, as they are available at a field of view of 30°. Then, by subtracting the Drishti-GS cup segmentations from the disc segmentations, rim segmentations were acquired. These were used as 'ground truth' for validation testing. The database has been used to compare performance between published optic cup and disc segmentation models through metrics such as IoU and Dice coefficient.[10,23-27] Few investigators have attempted rim segmentations on the Drishti-GS database.[26] Therefore, RimNet rim segmentations were used to recreate cup segmentations to allow for comparison with other segmentation models. The IoU for cup segmentations (CupIoU) and disc segmentations (DiscIoU) were reported. Additionally, the Dice scores for the cup (CupDice) and disc (DiscDice) were reported.

Evaluation Criteria

The main outcome measures are the median absolute error (MAE) difference between the glaucoma specialists and RimNet for 3 metrics: mRDR, RDAR, and ARW. A secondary measure is the

Table 1. Glaucoma Diagnosis for All 1208 Patients Included in the RimNet Dataset

Diagnosis	Count
Primary open-angle glaucoma	530
Glaucoma suspect	403
Chronic angle-closure glaucoma	71
Low-tension glaucoma	47
Secondary open-angle glaucoma	35
Capsular glaucoma with psuedoexfoliation	33
Anatomical narrow angle	27
Glaucoma secondary to eye infection	24
Pigmentary glaucoma	15
Secondary angle closure	11
Congenital glaucoma	7
Juvenile glaucoma	3
Acute angle-closure glaucoma	2

Table 2. Hyperparameter Search Space for the RimNet Included the Encoders, Decoders, Loss Functions, Learning Rates, and Optimizer

	Hyperparameters
Encoders	MobileNetV2, ResNet34, EfficientnetB0, InceptionV3, ResNet101, VGG16, ResNet50
Decoders	U-Net, FPN, LinkNet, PSPnet
Loss Function	Categorical_Crossentropy, Categorical_Focal_Loss
Learning Rate	10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}
Optimizer	Adam, SGD

The optimized metric was intersection over union of the rim.

RimIoU, the IoU of the RimNet rim segmentation compared to that of the glaucoma specialists.

The mRDR, RDAR, and ARW were explained earlier. Two measures of segmentation accuracy are also reported: Intersection over Union and Dice scores. The IoU, also known as the Jaccard distance, is a measure of segmentation accuracy. It compares the ground truth with the segmentation by reporting the ratio of the intersection area over the union area. The Dice score for cup and disc segmentations are reported for the Drishti-GS dataset to compare segmentation performance. The Dice score compares the ground truth with the segmentation by reporting the ratio of 2 times the intersection area over the summed area of the ground truth and segmentation.

Results

A database of 1208 optic disc photographs of 1 to 21 eyes from 903 glaucoma patients were used for training, validation, and testing in an 80/10/10 split. Both scanned slides and original digital images were represented in the dataset. The average (± standard deviation) age of the patients was 63.7 (± 14.9) with a 43:57 male-to-female ratio. Full demographics including gender, age, and race are listed in Table 3. The average (± standard deviation) visual field mean deviation was −8.03 ± 8.59 dB (range: −31.64, 3.59). Of the 1208 optic disc photographs, 340 had incomplete neuroretinal rims. The diagnoses for the patients are listed in Table 1.

Hyperparameter Architecture

Optimized parameters were found through the random search of 64 model combinations, detailed in Table 2.[18,28–38] The combination of the InceptionV3 backbone and LinkNet architecture proved to be the most accurate.[18,29] LinkNet is a lightweight decoder first published in 2017.[29] Other parameters identified include the loss function of binary cross-entropy, learning rate of 10^{-3}, and the Adam optimizer.[38]

Segmentation Network Results

The code used to train, run, and evaluate RimNet can be found on our public repository at https://github.com/Tyler-ADavis/GlaucomaML. On the test set, an mRDR MAE (interquartile range [IQR]) of 0.03 (0.05) was achieved on the intact rims while a ARW MAE (IQR) of 31 (89)° was achieved on the incomplete rims. Twenty-two of 34 eyes with incomplete rims were correctly identified as incomplete on segmentation. A RDAR MAE (IQR) of 0.09 (0.10) was achieved on all images. A RimIoU of 0.68 was achieved on intact rims, while a RimIoU of 0.45 was achieved on incomplete rims. The results of RimNet are presented in Table 4. Figure 3 demonstrates examples of RimNet segmentation results. To better examine the accuracy of the mRDR and RDAR calculations, the difference between the estimated values and the ground truths were
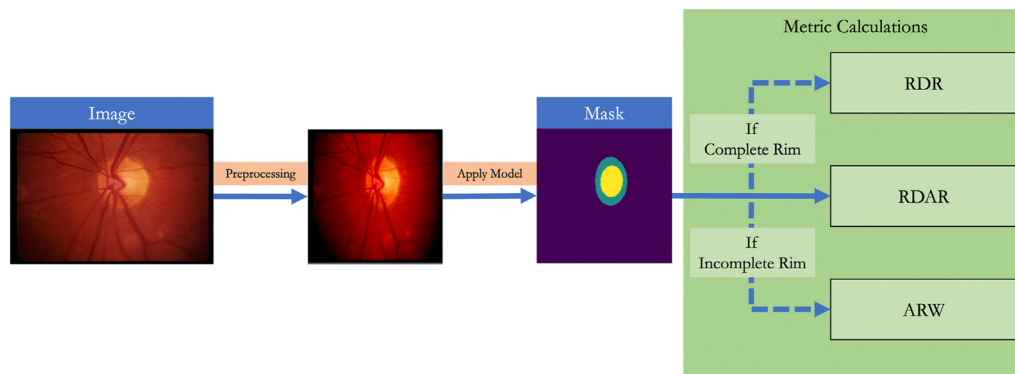

Figure 2. RimNet Pipeline. The raw image first undergoes preprocessing, where contrast limited adaptive histogram equalization is applied and the image is resized to model specifications. A mask is generated after applying the segmentation model to the preprocessed image. Image analysis then calculated the rim-to-disc-area ratio (RDAR). If the rim is intact, the minimum rim-to-disc ratio is calculated. If the rim is incomplete, absent rim width (ARW) is calculated. RDR = rim-to-disc ratio.
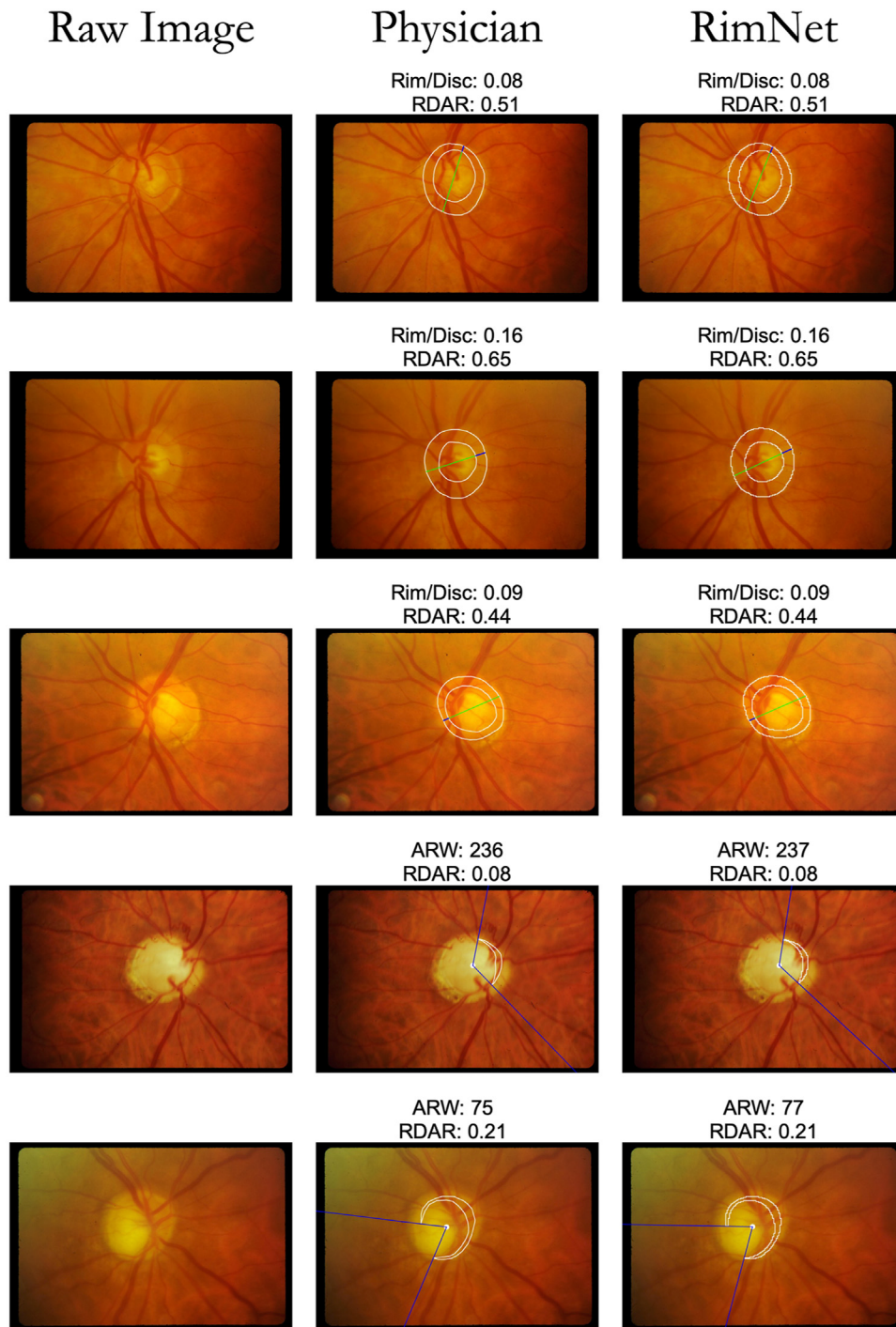
Figure 3. Segmentation Results. This figure demonstrates several examples of RimNet segmentation compared to physician segmentation. The left-most column shows the raw image. The middle column overlays the physician segmentation (white) over the raw image. The right-most column overlays the RimNet segmentation (white) over the raw image. In intact rims, green line shows the diameter and the dark blue shows the thinnest rim. In incomplete rims, the dark blue shows the edges of the segmentation. ARW = absent rim width; RDAR = rim-to-disc-area ratio.

calculated. Bland–Altman plots comparing the estimated and ground truth mRDR and RDAR are shown in Figure 4.

A comparison of RimNet segmentation on the Drishti-GS dataset to other published works is presented in Table 5. The mRDR MAE (IQR) was 0.03 (0.04) and the RDAR MAE (IQR) was 0.09 (0.10). The CupIoU was 0.77 and a DiscIoU was 0.91. The CupDice was 0.86 and DiscDice 0.95 was achieved.

Table 3. Demographic Data for Dataset

Variable	Slide Images	Digital Camera 1	Digital Camera 2	Digital Camera 3
Gender Distribution				
F	407	119	55	12
M	302	85	44	11
Age Distribution				
Mean	60.72	67.13	72.80	66.92
SD	13.48	17.43	12.75	17.71
Median	61.87	71.06	73.91	72.37
IQR	15.79	16.33	12.86	22.54
Min	9.36	6.92	16.19	17.48
Max	90.05	96.10	94.41	86.17
Race Distribution				
Asian	90	34	24	2
Black	63	22	8	1
Hispanic	66	20	16	6
White	366	100	45	12
Other	53	5	3	0
Unknown	71	22	3	2

F = female; IQR = interquartile range; M= male; SD = standard deviation.
Lists the gender distribution, age distribution, and racial distribution by camera type

Discussion

These results demonstrate that RimNet is capable of reasonably accurate segmentation and analysis of optic discs with both intact and incomplete rims. Spaeth et al. distinguished different DDLS grades by mRDR steps of 0.1.[5] The MAE of the mRDR is well within this value, showing that RimNet segmentations are clinically relevant. For more advanced glaucoma with DDLS grades of 6 and above, the neuroretinal rim is incomplete and Spaeth et al. uses the ARW to distinguish grades. The 5 categories are < 45°, 45° to 90°, 90° to 180°, 180° to 270°, and > 270°. The minimum step is 45°; the MAE falls slightly below that category at 31° with 22 of 34 total incomplete rims correctly identified as incomplete. However, the IQR demonstrates a broad range of ARW. The error echoes the difficulties faced by the glaucoma specialists. While segmenting these severely glaucomatous rims to create the

Table 4. RimNet Results on Internal Test Set and Drishti-GS Dataset

Variable	Internal	Drishti-GS
mRDR	0.04 (± 0.03)	0.04(±0.04)
ARW	48.9 (± 35.9)	N/A
RDAR	0.10 (± 0.09)	0.10(±0.08)
RimIoU	0.68	0.67
N	120 (87 intact, 33 Incomplete)	101 (101 intact, 0 Incomplete)

ARW = absent rim width; mRDR = minimum rim-to-disc ratio; N/A = not applicable; RDAR = rim-to-disc-area ratio; RimIoU = rim intersection over union.
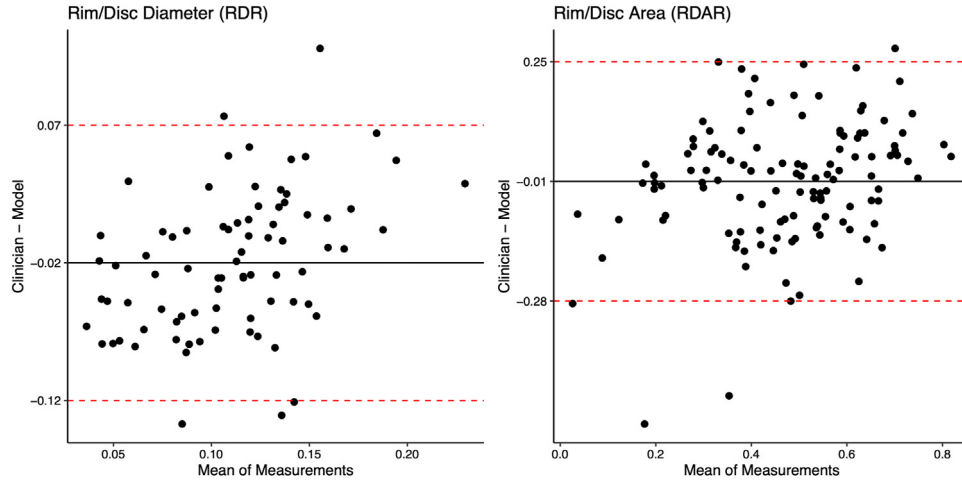The absent rim width cannot be calculated on the Drishti-GS dataset because all rims are intact.

"ground truth" masks, glaucoma specialists often differed regarding where rims were interrupted and if rims were incomplete or intact. Though a forced consensus was eventually reached, this demonstrates the difficulty of the task and the variability of this 'ground truth'. RimNet offers 65% accuracy in identifying incomplete rims and a relatively low ARW MAE. To our knowledge, RimNet is the first to offer such capabilities in published literature.

This work offers 3 improvements in the current landscape of optic disc segmentation. First, we utilized a dataset of 1208 images with external validation on Drishti-GS.[39] Second, while we have still reported IoU and Dice scores, we have focused on more clinically relevant metrics such as mRDR, RDAR, and ARW. Third, our study focuses on accurate segmentation of incomplete rims, an often overlooked aspect of optic disc segmentation. RimNet is a useful step towards completely automating the DDLS algorithm.

Automated segmentation of the optic disc and cup has been previously explored. The original studies were initially based on image processing functions such as thresholding, level set, active contour, clustering, and component extraction with success on local and publicly available datasets.[40] As early as 2001, Chrástek et al. offered an automated method of optic disc segmentation with filtering and edge detection, which achieved a segmentation accuracy of 82%.[41] In 2008, Liu and collaborators used level set and thresholding methods to achieve 97% accuracy when comparing algorithm-determined CDR ratio to clinical CDR ratio on a dataset of 73 images from the Singapore Eye Research Centre.[42] In 2015, Lotankar et al. used active contouring to achieve a 99% pixel-to-pixel accuracy on a private database of 150 images.[43] However, each of these approaches was limited in scope. Level-setting and thresholding would fail with images with decreased or increased intensity caused by pathological findings, which can be commonly seen on optic disc photographs such as peripapillary atrophy. This leads to

Figure 4. Bland-Altman plots showing the agreements in minimum rim-to-disc ratio and rim-to-disc-area ratio (RDAR) between clinician and RimNet in test images; each black point corresponds to one test set image. Red dashed lines indicate 95% confidence limits. RDR = rim-to-disc ratio.

overestimating or underestimating CDRs. Active contouring may similarly be affected by abnormal pathology or bright artifacts fixating on local maxima or minima within the image. Therefore, though these methods have proven efficacy, they can be improved upon.

An automated grading system for glaucoma diagnosis and progression needs a high efficiency, broadly applicable segmentation algorithm with an expansive learning capacity which could be applied to a variety of funduscopic images acquired with different imaging modalities with concurrent pathologies and variations. Though further work must be done, deep learning and convolutional neural networks may play an important role in the solution. They have an enormous learning capacity relative to their size.[11] Rapid advances in computational memory and processing speed have made neural networks more accessible for optic disc segmentation. Zilly et al. used ensemble learning to achieve 89% IoU on disc segmentation and 84% IoU on cup segmentation on the Drishti-GS dataset.[23] Sevastopolsky et al. furthered this work by using a modified U-Net to

achieve a comparable accuracy in less than a 10th of the time.[24] More on segmentation efforts, both image processing functions and neural network attempts, can be found on a review article by Thakur and Juneja et al.[40]

Several groups have pursued automated mRDR and RDAR calculations. In 2019, Kumar et al. proposed using an imaging processing technique called active discs to segment the optic disc and cup and perform general glaucoma classification (normal, moderate, severe) based on mRDR.[6] Though direct mRDR accuracy was not reported, an mRDR-based approach demonstrated high classification accuracy. In 2020, Martins et al. proposed a smartphone-based glaucoma diagnosis pipeline, which focuses on glaucoma classification and calculates RDAR.[44] However, RDAR results were not directly reported. More recently, Pachade et al. proposed an NENet model consisting of EfficentNetB4 and adversarial learning that achieved an area under the curve of 0.901 on RDAR calculation for Drishti-GS.[45]

To the best of our knowledge, RimNet is the first engineering attempt to pursue segmentation and glaucoma grading efforts with incomplete neuroretinal rims. Thus, direct comparison of RimNet to other segmentation models is difficult. However, through the Drishti-GS dataset, an artificially-derived segmentation comparison is possible by recreating cup and disc masks from the RimNet rim segmentations. Table 5 demonstrates that RimNet performed well overall compared to recent segmentation models on the Drishti-GS dataset. While it outperformed several other models in CupDice, DiscIoU, and DiscDice segmentations, it was below average in CupIoU. These results must be understood in the context of 3 factors. First, the Drishti-GS images were available as 30° field of views. However, RimNet requires images centered and cropped near the optic disc margin. Therefore, RimNet has a significant information loss compared to other models that use the 30° field of view. Second, RimNet is unique in that it has been trained on both complete and incomplete rims. The models compared to

Table 5. DRISHTI-GS Segmentation Performance of RimNet Compared to Published Segmentation Models[11,17−21]

Model	CupIoU	DiscIoU	CupDice	DiscDice
RimNet	0.77	**0.91**	0.86	0.95
Zilly et al. (2017)[23]	**0.85**	-	0.87	0.87
Sevastopolsky (2017)[24]	0.75	-	-	-
Edupuganti et al. (2018)[25]	0.81	0.69	-	-
Al-Bander and Zheng et al. (2018)[26]	-	-	0.83	0.95
Joshua et al. (2019)[10]	0.79	-	-	-
Yu et al. (2019)[27]	-	-	**0.88**	**0.97**

Bolded values represent the highest achieved value for each metric. CupIoU = intersection over union of the optic cup; CupDice = dice score of the optic cup; DiscDice = dice score of the optic disc; DiscIoU = intersection over union of the optic disc.

RimNet have been trained only on complete rims. It is reasonable to expect a higher segmentation accuracy in these cases. Finally, the cup and disc segmentations produced by RimNet were artificially derived from the RimNet's rim segmentation. By not directly predicting on the cup and disc, accuracy was lost. Considering these 3 factors, RimNet's performance on Drishti-GS is acceptable. This is corroborated by the Drishti-GS mRDR MAE of 0.03 (0.04) and RDAR MAE of 0.09 (0.10), both of which are low.

The findings of this study need to be interpreted with the shortcomings in mind. First, the hyperparameter architecture search was limited by the computing and memory limits of our workstation, which uses NVIDIA RTX 2080 Ti graphics cards. We could not include larger models such as ResNet152 into our search due to these memory constraints. Second, the number of ground truth masks and optic disc images, particularly those of more severe glaucoma, is limited. Greater numbers of diverse samples would allow RimNet to better learn mRDR and ARW calculations.

RimNet brings glaucomatous detection and DDLS grading a step closer to full automation.[5] Automated grading of disc size is a necessary step to fully autonomous DDLS grading. A future goal would be to not only pursue full automation of DDLS grading, but to test their capabilities as diagnostic tools. One promising avenue for further investigation would be screening with smartphone fundoscopy. The increasing quality of smartphone cameras have made smartphone fundoscopy viable as a screening method.[46,47] This, combined with automated DDLS grading, could provide a powerful screening tool to revolutionize glaucoma detection.

In conclusion, RimNet provides a method for high efficacy rim segmentation, mRDR, and ARW calculation. It also provides an example of how ophthalmic care be augmented by artificial intelligence. Though more work remains to be done, we believe that detection, diagnosis, and care of glaucoma can integrate with approaches such as these and aid ophthalmologists in decision making to provide higher quality care for a global population of patients.

Footnotes and Disclosures

[1] David Geffen School of Medicine, University of California Los Angeles, Los Angeles, California.

[2] Department of Computer Science, University of California Los Angeles, Los Angeles, California.

[3] Glaucoma Division, Jules Stein Eye Institute, Los Angeles, California.

[4] Jonathan and Karin Fielding School of Public Health, University of California Los Angeles, Los Angeles, California.

[5] Department of Biostatistics, University of California Los Angeles, Los Angeles, California.

Abbreviations and Acronyms:
ARW = absent rim width; **CDR** = cup-to-disc-ratio; **CupDice** = Dice score of the optic cup; **CupIoU** = IoU of the optic cup; **DDLS** = disc damage likelihood scale; **DiscDice** = dice score of the optic disc; **DiscIoU** = IoU of the optic disc; **IoU** = intersection over union; **IQR** = interquartile range; **MAE** = median average error; **mRDR** = minimum rim-to-disc ratio; **RDAR** = rim to disc area ratio.

Correspondence:
Joseph Caprioli, MD, 100 Stein Plaza 2-118, Los Angeles, CA 90095. E-mail: caprioli@jsei.ucla.edu.

References

1. Giangiacomo A, Coleman AL. The epidemiology of glaucoma. In: *Glaucoma*. Berlin, Heidelberg: Springer Berlin Heidelberg:13−21; 2009.

2. Michelson G, Hornegger J, Wärntges S, Lausen B. The papilla as screening parameter for early diagnosis of glaucoma. *Dtsch Arztebl Int*. 2008;105:583.

3. Shaikh Y, Yu F, Coleman AL. Burden of undetected and untreated glaucoma in the United States. *Am J Ophthalmol.* 2014;158:1121−1129.e1.

4. Cristina Leske M, Heijl A, Hussein M, et al. Factors for glaucoma progression and the effect of treatment: the early manifest glaucoma trial. *Arch Ophthalmol.* 2003;121:48−56.

5. Spaeth GL, Henderer J, Liu C, et al. The disc damage likelihood scale: reproducibility of a new method of estimating the amount of optic nerve damage caused by glaucoma. *Trans Am Ophthalmol Soc.* 2002;100:181.

6. Kumar JRH, Seelamantula CS, Kamath YS, Jampala R. Rim-to-disc ratio outperforms cup-to-disc ratio for glaucoma pre-screening. *Sci Rep.* 2019;9:7099.

7. Formichella P, Annoh R, Zeri F, Tatham AJ. The role of the disc damage likelihood scale in glaucoma detection by community optometrists. *Ophthalmic Physiol Opt.* 2020;40:752−759.

8. Tong W, Romero M, Lim V, et al. Reliability of graders and comparison with an automated algorithm for vertical cup-disc ratio grading in fundus photographs. *Ann Acad Med Singap.* 2019;48:282−289.

9. Sarvamangala DR, Kulkarni RV. Convolutional neural networks in medical image understanding: a survey. *Evol Intell.* 2021;15:1−22.

10. Joshua AO, Nelwamondo FV, Mabuza-Hocquet G. Segmentation of optic cup and disc for diagnosis of glaucoma on retinal fundus images. In: *Proceedings - 2019 Southern African Universities Power Engineering Conference/Robotics and Mechatronics/Pattern Recognition Association of South Africa, SAUPEC/RobMech/PRASA 2019.* United States: Institute of Electrical and Electronics Engineers Inc; 2019: 183−187.

11. Cybenko G. Approximation by superpositions of a sigmoidal function. *Math Control Signals Syst.* 1989;2:303−314.

12. van Rossum G, Drake FL. *Python 3 Reference Manual.* Scotts Valley, CA: CreateSpace; 2009:242.

13. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: machine learning in Python. *J Mach Learn Res.* 2011;12: 2825−2830.

14. Harris CR, Millman KJ, van der Walt SJ, et al. Array programming with NumPy. *Nature.* 2020;585:357−362.

15. Abadi M, Barham P, Chen J, et al. {TensorFlow}: a system for {large-scale} machine learning. https://tensorflow.org; 2016. Accessed September 8, 2022.

16. Tom O', Elie B, James L, et al. *Keras Tuner.* United States: GitHub Repository; 2019.

17. Elsken T, Metzen JH, Hutter F. Neural architecture search: a survey. *J Mach Learn Res.* 2019;20:1−21.

18. Szegedy C, Vanhoucke V, Ioffe S, et al. Rethinking the inception architecture for computer vision. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2015.* United States: IEEE Computer Society; 2016:2818−2826.

19. Shoaib M, Sayed N. YOLO object detector and inception-V3 convolutional neural network for improved brain tumor segmentation. *Trait Du Signal.* 2022;39:371−380.

20. Salama WM, Aly MH. Deep learning in mammography images segmentation and classification: automated CNN approach. *Alex Eng J.* 2021;60:4701−4709.

21. Sivaswamy J, Chakravarty A, Datt Joshi G, Abbas Syed T. A comprehensive retinal image dataset for the assessment of glaucoma from the optic nerve head analysis. *JSM Biomed Imaging Data Pap.* 2015;2:1004.

22. Sivaswamy J, Krishnadas SR, Joshi GD, et al. Drishti-GS: retinal image dataset for optic nerve head(ONH) segmentation. In: *2014 IEEE 11th International Symposium on Biomedical Imaging, ISBI 2014.* United States: IEEE Computer Society; 2014:53−56.

23. Zilly J, Buhmann JM, Mahapatra D. Glaucoma detection using entropy sampling and ensemble learning for automatic optic cup and disc segmentation. *Comput Med Imaging Graph.* 2017;55:28−41.

24. Sevastopolsky A. Optic disc and cup segmentation methods for glaucoma detection with modification of U-net convolutional neural network. *Pattern Recognit Image Anal.* 2017;27:618−624.

25. Edupuganti VG, Chawla A, Kale A. Automatic optic disk and cup segmentation of fundus images using deep learning. In: *Proceedings - International Conference on Image Processing, ICIP.* United States: IEEE Computer Society; 2018:2227−2231.

26. Al-Bander B, Williams BM, Al-Nuaimy W, et al. Dense fully convolutional segmentation of the optic disc and cup in colour fundus for glaucoma diagnosis. *Symmetry.* 2018;10:87.

27. Yu S, Xiao D, Frost S, Kanagasingam Y. Robust optic disc and cup segmentation with deep learning for glaucoma detection. *Comput Med Imaging Graph.* 2019;74:61−71.

28. Mannor S, Peleg B, Rubinstein R. The cross entropy method for classification. In: *ICML 2005 - Proceedings of the 22nd International Conference on Machine Learning.* United States: International Conference on Machine Learning; 2005:561−568.

29. Chaurasia A, Culurciello E. LinkNet: exploiting encoder representations for efficient semantic segmentation. In: *2017 IEEE Visual Communications and Image Processing, VCIP 2017.* United States: IEEE Computer Society; 2017:1−4.

30. Sandler M, Howard A, Zhu M, et al. MobileNetV2: inverted residuals and linear bottlenecks. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition.* United States: IEEE Computer Society; 2018:4510−4520.

31. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2015.* United States: IEEE Computer Society; 2016:770−778.

32. Tan M, Le QV. EfficientNet: rethinking model scaling for convolutional neural networks. In: *36th International Conference on Machine Learning, ICML 2019.* United States: International Conference on Machine Learning; 2019: 10691−10700.

33. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. In: *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings.* United States: International Conference on Learning Representations; 2014.

34. Weng W, Zhu X. U-net: convolutional networks for biomedical image segmentation. *IEEE Access.* 2015;9:16591−16603.

35. Lin T-Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection. https://arxiv.org/abs/1612.03144v2; 2016. Accessed June 10, 2022.

36. Zhao H, Shi J, Qi X, et al. Pyramid scene parsing network. In: *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017.* United States: IEEE Computer Society; 2016:6230−6239.

37. Ruder S. An overview of gradient descent optimization algorithms. https://arxiv.org/abs/1609.04747v2; 2016. Accessed June 10, 2022.

38. Kingma DP, Ba JL. Adam: a method for Stochastic optimization. In: *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings.* United States: International Conference on Learning Representations; 2014.

39. Sivaswamy J, Krishnadas SR, Joshi GD, et al. Drishti-GS: retinal image dataset for optic nerve head(ONH) segmentation.

In: *2014 IEEE 11th International Symposium on Biomedical Imaging, ISBI 2014*. United States: Institute of Electrical and Electronics Engineers Inc; 2014:53−56.

40. Thakur N, Juneja M. Survey on segmentation and classification approaches of optic cup and optic disc for diagnosis of glaucoma. *Biomed Signal Process Control*. 2018;42:162−189.

41. Chrástek R, Wolf M, Donath K, et al. Automated segmentation of the optic nerve head for diagnosis of glaucoma. *Med Image Anal*. 2005;9:297−314.

42. Liu J, Wong DWK, Lim JH, et al. Optic cup and disk extraction from retinal fundus images for determination of cup-to-disc ratio. In: *2008 3rd IEEE Conference on Industrial Electronics and Applications, ICIEA 2008*. United States: IEEE Computer Society; 2008:1828−1832.

43. Lotankar M, Noronha K, Koti J. Detection of optic disc and cup from color retinal images for automated diagnosis of glaucoma. In: *2015 IEEE UP Section Conference on Electrical Computer and Electronics, UPCON 2015*. United States: IEEE Computer Society; 2016.

44. Martins J, Cardoso JS, Soares F. Offline computer-aided diagnosis for Glaucoma detection using fundus images targeted at mobile devices. *Comput Methods Programs Biomed*. 2020;192:105341.

45. Pachade S, Porwal P, Kokare M, et al. NENet: nested EfficientNet and adversarial learning for joint optic disc and cup segmentation. *Med Image Anal*. 2021;74:102253.

46. Panwar N, Huang P, Lee J, et al. Fundus photography in the 21st century—a review of recent technological advances and their implications for worldwide healthcare. *Telemed J E Health*. 2016;22:198.

47. Nazari Khanamiri H, Nakatsuka A, El-Annan J. Smartphone fundus photography. *J Vis Exp*. 2017:e55958.