tvst

Fully Automated Segmentation of Human Eyeball Using Three-Dimensional U-Net in T2 Magnetic Resonance Imaging

Jin-Ju Yang[1,2,*], Kyeong Ho Kim[5,*], Jinwoo Hong[4], Yeji Yeon[1,2], Ji Young Lee[6], Won June Lee[1,2], Yu Jeong Kim[1,2], Jong-Min Lee[3,‡], and Han Woong Lim[1,2,‡]

[1] Department of Ophthalmology, Hanyang University College of Medicine, Seoul, Korea
[2] Hanyang Vision Research Center, Hanyang University, Seoul, Korea
[3] Department of Biomedical Engineering, Hanyang University, Seoul, Korea
[4] Department of Electronic Engineering, Hanyang University, Seoul, Korea
[5] Department of Artificial Intelligence, Hanyang University, Seoul, Korea
[6] Department of Radiology, Seoul St. Mary's Hospital, Seoul, Korea

Correspondence: Han Woong Lim, Department of Ophthalmology, Hanyang University College of Medicine, 222-1, Wangsimni-ro, Seongdong-gu, Seoul 04763, Korea. e-mail: limhw@nate.com
Jong-Min Lee, Department of Biomedical Engineering, Hanyang University, 222 Wangsimni-ro, Sungdong-gu, Seoul 04763, Republic of Korea. e-mail: ljm@hanyang.ac.kr

Purpose: To develop and validate a fully automated deep-learning-based tool for segmentation of the human eyeball using a three-dimensional (3D) U-Net, compare its performance to semiautomatic segmentation ground truth and a two-dimensional (2D) U-Net, and analyze age and sex differences in eyeball volume, as well as gaze-dependent volume consistency in normal subjects.

Methods: We retrospectively collected 474 magnetic resonance imaging (MRI) scans, including different gazing scans, from 119 patients. A 10-fold cross-validation was applied to separate the dataset into training, test, and validation sets for both the 3D U-Net and 2D U-Net. Performance accuracy was measured using four quantitative metrics compared to the ground truth, and Bland–Altman plot analysis was conducted. Age and sex differences in eyeball volume and variability in eyeball volume differences across gazing directions were analyzed.

Results: The 3D U-Net outperformed the 2D U-Net with mean accuracy scores >0.95, showing acceptable agreement in the Bland–Altman plot analysis despite a tendency for slight overestimation (mean difference $= -0.172$ cm^3). Significant sex differences and age effects on eyeball volume were observed for both methods ($P < 0.05$). No significant volume differences were found between the segmentation methods or within each method for the different gazing directions. Significant differences in performance accuracy were identified among the five gazing directions, with the upward direction showing a notably lower performance.

Conclusions: Our study demonstrated the effectiveness of 3D U-Net human eyeball volume segmentation using T2-weighted MRI. The robustness and reliability of 3D U-Net across diverse populations and gaze directions support enhanced ophthalmic diagnosis and treatment strategies.

Translational Relevance: Our findings demonstrate the feasibility of using the proposed 3D U-Net model for the automatic segmentation of the human eyeball, with potential applications in various ophthalmic research fields that require the analysis of 3D geometric eye globe shapes or eye movement detection.

1

Introduction

The ocular globes, which comprise paired spherical sensory organs located anteriorly on the face within the orbit, house the visual apparatus.[1] Numerous studies analyzing ocular structures and measuring ocular volumes have demonstrated the importance of magnetic resonance imaging (MRI).[2–7] Although the structure of the human eyeball appears simple because of its spherical shape, precise, robust, and efficient measurements are required. Ocular segmentation forms the basis for further geometric and volumetric evaluation of three-dimensional (3D) ocular structures.

In recent years, several conventional segmentation methods have been proposed for ocular structure analysis using MRI, primarily focusing on ocular volume or shape analysis using T2-weighted MRI.[2–5,8,9] However, there is a lack of advanced techniques, such as deep learning, for analyzing ocular structures in nonpathological eyes, such as emmetropic or ametropic eyes, and in pathological conditions. The potential benefits of incorporating deep learning models, which have recently demonstrated promising results in ocular tumor segmentation,[10,11] must be explored for more comprehensive analyses of ocular structures in various clinical scenarios.

The primary advantage of deep learning–based segmentation is its ability to process large datasets rapidly and provide robust results, which is often infeasible for manual segmentation, with manual presets necessary for training data.[12] Moreover, deep learning–based segmentation can improve performance by accommodating diverse variations in architecture and using hybrid approaches that utilize state-of-the-art deep-learning models.[13] We opted for the 3D U-Net model because it is better suited to spherical structures owing to its use of complete 3D spatial information, rather than a two-dimensional (2D) architecture's approach.[14–16] To compare the accuracy of the 3D U-Net, we used a 2D U-Net with multi-view aggregation as our baseline model.[17]

This study aimed to develop a fully automated segmentation tool for human eyeballs using a 3D U-Net deep learning model on a large dataset and to evaluate the segmentation performance. We retrospectively collected 474 T2-weighted MRI scans encompassing various gaze scans from 119 subjects. Additionally, we investigated age and sex differences in eyeball volume, as well as their variability across various gaze directions, including central, lateral, medial, upward, and downward gaze, in both the emmetropic and ametropic eyes of normal subjects. Our proposed method lays the groundwork for future studies requiring the analysis of 3D geometric eye globe shapes or eye movement detection, enabling its application in a wide range of ophthalmic research.

Methods

Dataset

The MRI data used in this study were collected retrospectively; further details on the data acquisition, experimental setup, and demographics are described in the literature.[18–20] The dataset contained MRI scans from 121 subjects, including 49 normal subjects, 39 high-myopia individuals with glaucoma, and 33 patients with strabismus, with varying gaze scans (front, left, right, up, and down). Different gaze directions were achieved using markers within the MRI equipment, positioned at 30° for horizontal movements and 20° for vertical movements. The targets were adjusted accordingly for patients with strabismus. Of the initial 494 images, 20 scans were excluded for the following reasons: seven scans had the eye region cropped during the preprocessing stage (one strabismus patient and one glaucoma patient), eight scans from emmetropic eyes lacked data for gaze directions other than the front gaze, and five scans were missing from the dataset despite being acquired. Consequently, 474 images from 119 participants were used to develop the deep learning model. A 10-fold cross-validation was used to split the data into training, validation, and test sets.[21] This study was approved by the Institutional Review Board of Hanyang University Medical Center and was conducted in accordance with the Declaration of Helsinki. The requirement for informed consent was waived owing to the retrospective nature of the fully anonymized images. The code for deep learning implementation can be accessed at (https://github.com/hoya1212/eyeball-segmentation).

MRI Acquisition

High-resolution 3D T2-weighted images were acquired using a 3T MRI scanner (Achieva; Philips Medical Systems, Best, Netherlands) with a 32-channel head coil. The image parameters included the following: repetition time, 2500 ms; echo time, 248 ms; flip angle, 90°; section thickness = 0.6 mm; field of view, 180 × 180 mm; matrix size, 256 × 256.

Preprocessing

MRI data were converted from DICOM to NIfTI format using MRIcron software (https://github.com/neurolabusc/MRIcron), and preprocessing

was performed using FSL software (https://fsl.fmrib.ox.ac.uk/fsl/fslwiki). Aligning images from different gazing directions with front-gazing images allowed for the inclusion of diverse gazing information within each individual's data, improving the training dataset and promoting more accurate data processing and analysis. To avoid high computational costs when training with 3D U-NET, the images were cropped to a consistent input size of $256 \times 130 \times 90$ before the training step, localizing the region of interest while resampling by an isovoxel.

Ground Truth

Manual delineation of the two eyeballs in slice-per-slice 3D MRI is considered the ground truth; however, it is labor-intensive and time-consuming. As an alternative, we adopted the widely used semiautomatic segmentation tool ITK-SNAP (version 3.6.0, http://itksnap.org), which encompasses the various medical image segmentation methods.[22] This semiautomatic method serves as an efficient alternative to manual segmentation, striking a balance between minimizing user intervention and maintaining the reliability and appropriateness of our methodology for 3D MRI. The benefits of semiautomatic segmentation include reduced time consumption, minimized human error compared to manual segmentation, and improved consistency between slices. Our primary focus is on proposing a fully automated deep-learning-based method that eliminates the need for researcher intervention. An overview of these steps is provided in Supplementary Figure S1. The process involves three main steps: presegmentation by selecting an intensity threshold, initializing a deformable object, and setting the parameters for an active contour model. After the semiautomatic segmentation process, we manually fine-tuned the results by adjusting the gray-level intensity of the MRI data to binarize the hyperintense regions of the eyeballs. The graphical user interface of the tool allowed us to review and edit the segmentation of the axial, sagittal, and coronal slices by adjusting the gray intensity threshold for binarization. This manual intervention was crucial for ensuring the precision of the segmentation results, ultimately leading to accurate identification of the ocular volume.

Deep Learning Model

Similar to the original U-Net, our 3D U-Net consists of a contracting path and an expansive path, which down-sample and up-sample the 3D input image, respectively.[14,15] To compare the accuracy of the proposed 3D U-Net, we generated a 2D U-Net with multi-view aggregation as the baseline method.[17] It consists of fully connected CNNs operating on 2D slices for coronal, axial, and sagittal views, followed by a view-aggregation step to infer the final segmentation. This model is simpler, adopts the canonical Dice loss instead of the previously used hybrid Dice loss, and excludes the test-time augmentation technique. The details of our deep learning architecture are described in the Supplementary Materials and a schematic illustration is provided in Supplementary Figure S1. A representative example of our 3D U-Net results and the ground truth are shown in Figure 1. This figure presents the original MRI, ground truth, and 3D U-Net segmentation results from various perspectives.

Performance Evaluation

The deep learning output was compared with the ground truth using four quantitative metrics that are commonly used to validate medical volume segmentation because of their reproducibility and repeatability.[23] These metrics include the following: (1) Dice Similarity Coefficient, which evaluates the spatial overlap between the predicted and ground truth segmentations. A value of 1 indicates a perfect match, whereas a value of 0 signifies no overlap. (2) Precision, also known as *positive predictive value*, measures the proportion of true positive predictions within the total predicted positive instances. High precision indicates that the model accurately predicted positive cases. (3) Recall, also referred to as *sensitivity* or *true-positive rate*, calculates the proportion of true positive predictions within actual positive instances. High recall indicates that the model could detect most of the positive cases. (4) Jaccard Similarity Index gauges the similarity between two sets by calculating the intersection divided by the union of the predicted and ground-truth segmentations.

Statistical Analyses

In the statistical analysis, we performed the χ^2 test for categorical variables (such as sex) and the Kruskal–Wallis test for continuous variables (such as age and eyeball volume) to evaluate the differences among the three groups. Dunn's test was used for post hoc multiple comparisons when significant differences were observed ($P < 0.05$). A general linear model was used to analyze the association between age and eyeball volume. A Bland–Altman plot was used to assess the level of agreement between the ground truth and 3D U-Net segmented volumes. Gaze terms were categorized according to five different
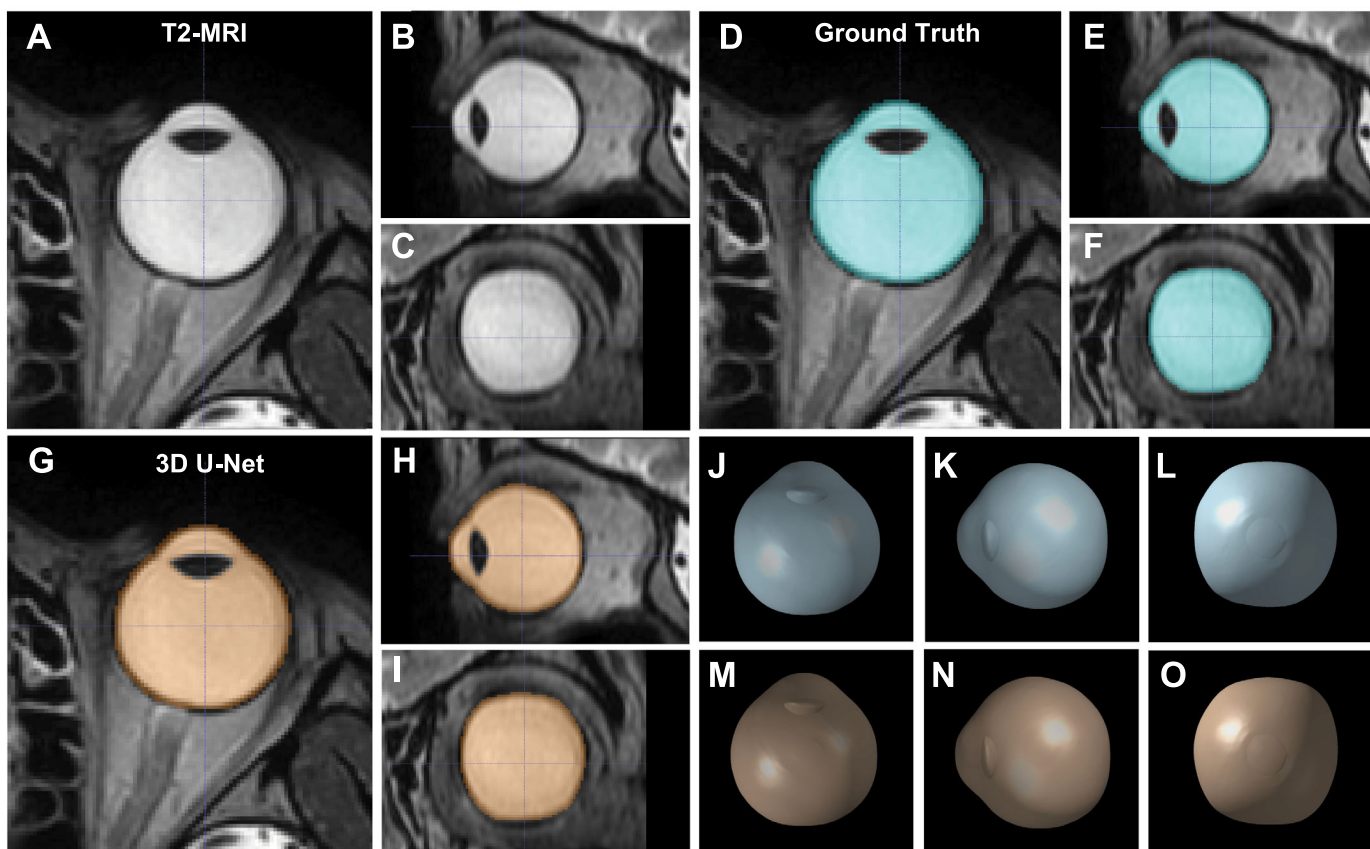
Figure 1. Human eyeball segmentation using 3D U-Net on T2-weighted MRI. (**A–C**) Original T2-weighted MRI, (**D–F**) ground truth segmentation results, (**G–I**) 3D U-Net segmentation results, and (**J–O**) volume rendering for both methods. An MRI from Group 1 data was randomly selected and cropped around the eyeball center of the left eye. (**A**) Axial, (**B**) sagittal, and (**C**) coronal views of the original image. (**D**) Axial, (**E**) sagittal, and (**F**) coronal views of the ground truth with a light blue color overlaid on the original T2-weighted MRI. (**G**) Axial, (**H**) sagittal, and (**I**) coronal views of the 3D U-Net segmentation results with an orange color overlaid on the original T2-weighted MRI. Volume-rendered eyeball for (**J**) axial, (**K**) sagittal (lateral), and (**L**) coronal (posterior) views of the ground truth. Volume-rendered eyeball for (**M**) axial, (**N**) sagittal (lateral), and (**O**) coronal (posterior) views of the 3D U-Net.

directions: front, abduction, adduction, up, and down, which encompassed both eyes of normal subjects. All statistical analyses and visualizations were performed using Python 3.7 (https://www.python.org/) in the Jupyter Notebook environment (https://jupyter.org/). The following Python packages were used for the analysis and visualization: NumPy, pandas, scipy, matplotlib, seaborn, and state models.

Results

Demographics and Characteristics of Study Participants

The demographics and characteristics of the study participants are presented in the Table. Statistically significant differences in age were observed among Groups 1, 2, and 3, whereas no significant differences

in sex were found. Eyeball volumes measured by the ground truth, 3D U-Net, and 2D U-Net demonstrated statistically significant differences between the three groups, as well as between males and females. Because of a lack of ophthalmic clinical information for Groups 2 and 3, the eyeball volume analysis results regarding age, sex differences, and gaze direction were reported only for Group 1.

Evaluation of Segmentation Performance

A performance evaluation was conducted on 948 eyes from all subjects, with both eyes included in the test set optimized for the model parameters. Four quantitative metrics were used in the assessment. The 3D U-Net outperformed the 2D U-Net across all evaluation metrics, achieving average scores above 0.95 (Fig. 2A). A comparative analysis was performed based on group (Fig. 2B) and gaze direction (Supplementary

Table. Demographics and the Characteristics of Study Participants

Characteristic	Total	Group1	Group2	Group3	Statistics[*]
Number of subjects	119	49	32	38	
Number of scans (missing)	474	245 (−13)	128	114	
Number of eyes	948	464	256	228	
Age	32.303 ± 12.521	30.163 ± 9.526	30.062 ± 17.031	36.947 ± 9.899	23.642[†,‡]
Age range (min - max)	9–81	16–56	9–81	20–55	
M/F, (Female ratio %)	69/50 (42%)	29/20 (41%)	19/13 (41%)	21/17 (45%)	0.170
Eyeball volume (cm³) for all eyes					
Ground truth	9.887 ± 1.689	9.026 ± 1.350	10.447 ± 1.179	11.009 ± 1.863	260.200[†,§]
3D U-Net	10.059 ± 1.684	9.085 ± 1.304	10.669 ± 1.091	11.356 ± 1.723	343.629[†,‡,§]
2D U-Net	9.505 ± 1.350	9.241 ± 1.259	9.423 ± 1.120	10.154 ± 1.554	56.280[†,‡]
Eyeball volume (cm³) for male eyes					
Ground truth	10.147 ± 1.672	9.371 ± 1.309	10.641 ± 1.157	11.240 ± 2.026	136.158[†,§]
3D U-Net	10.341 ± 1.645	9.438 ± 1.259	10.901 ± 1.011	11.627 ± 1.835	191.050[†,‡,§]
Eyeball volume (cm³) for female eyes					
Ground truth	9.524 ± 1.645	8.529 ± 1.249	10.163 ± 1.154	10.725 ± 1.594	141.975[†,§]
3D U-Net	9.666 ± 1.660	8.574 ± 1.194	10.331 ± 1.114	11.021 ± 1.506	172.298[†,‡,§]

F, female; M, male.

Continuous variables in each group are presented as the mean ± standard deviation, N, number of subjects, Group1, 2, and 3 represent normal subjects, patients with strabismus, and glaucoma.

[*]χ^2 test for sex distribution among groups.

[†]Kruskal-Wallis test for continuous variables and multiple comparisons using Dunn's test: Group1 vs. Group 3 for significant $P < 0.05$.

[‡]Kruskal-Wallis test for continuous variables and multiple comparisons using Dunn's test: Group 2 vs. Group 3 for significant $P < 0.05$.

[§]Kruskal-Wallis test for continuous variables and multiple comparisons using Dunn's test: Group 1 vs. Group 2 for significant $P < 0.05$.
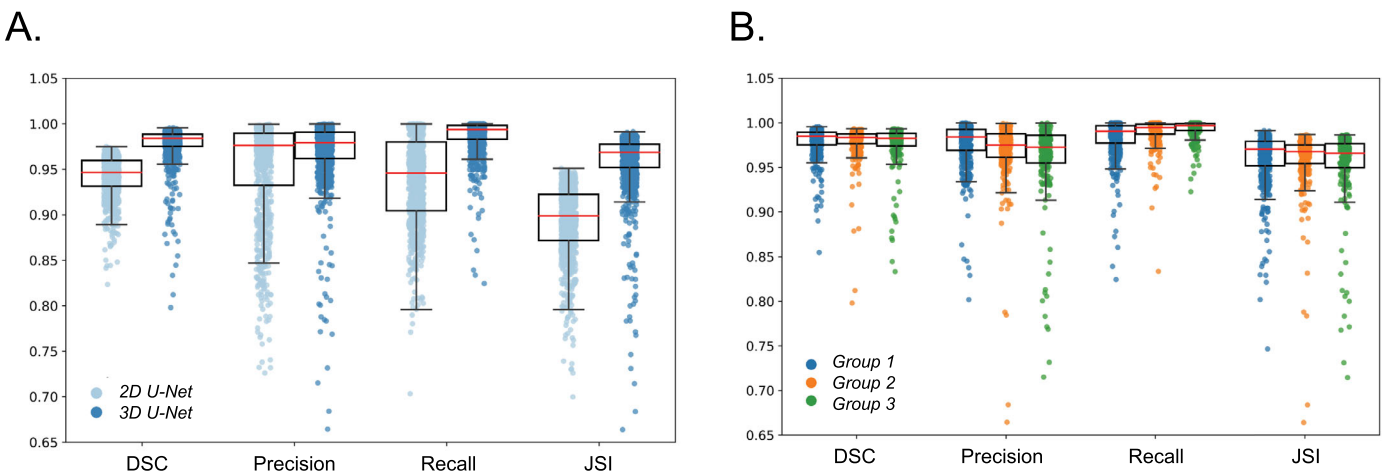

Figure 2. Comparison of deep learning performance evaluation. (**A**) Four evaluation metrics, including the Dice Similarity Coefficient (DSC), Precision, Recall, and Jaccard Similarity Index (JSI), ranged from 0 to 1, are plotted with boxplots and individual data points for 3D U-Net (*blue points*, N = 948 eyes from 474 MRIs) and 2D U-Net (*light blue points*, N = 942 eyes from 471 MRIs). (**B**) The same metrics are plotted only for the 3D U-Net in each group (Group 1, *blue*; Group 2, *orange*; Group 3, *green*). On each box, the *central red mark* indicates the median, and the edges of the box represent the 25th and 75th percentiles.
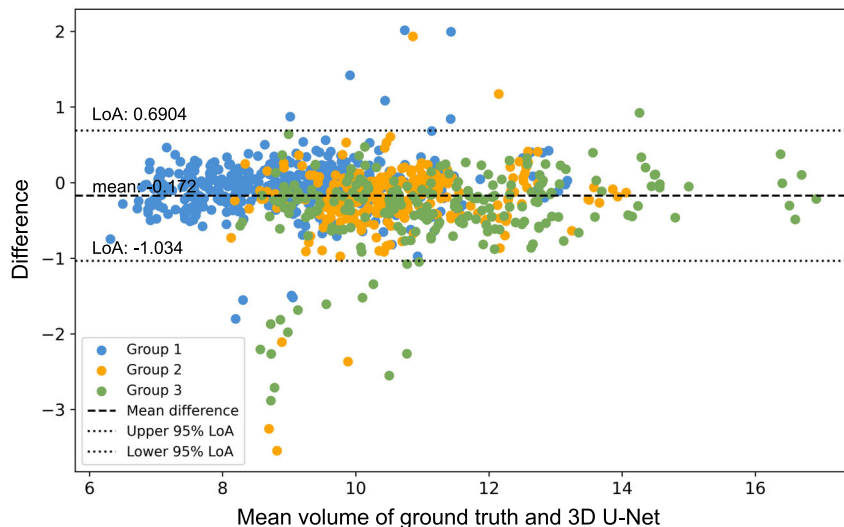
Figure 3. Bland-Altman plot for eyeball volume agreement between ground truth and 3D U-Net. The points with three colors for each group (Group 1, *blue*; Group 2, *orange*; Group 3, *green*) represent the individual deviation of each volume compared to the ground truth (N = 948 eyes). The *central solid line* indicates a mean difference of -0.172 cm^3 (ground truth $-$ 3D U-Net), and the *dashed lines* represent the lower and upper limits of agreement ($\pm 1.96*$ Standard Deviation). The points above the upper limits accounted for 0.949% (nine eyes), whereas those below the lower limits accounted for 2.426% (23 eyes).
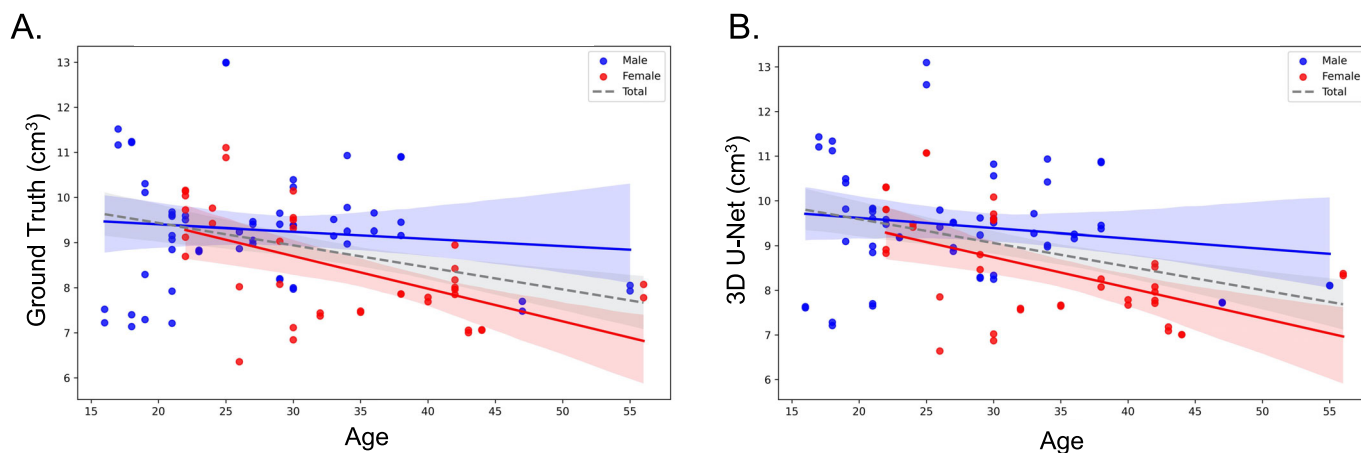

Figure 4. Age and sex differences in eyeball volume. (**A**) The x-axis represents age, and the y-axis represents the eye volume by ground truth. (**B**) The x-axis represents age, and the y-axis represents the eye volume by 3D U-Net. Scatter points are displayed in *blue* for males and *red* for females, with least square lines and confidence interval for each group and combined group using a *gray dotted line* with the confidence interval.

Fig. S2). Statistically significant differences were observed for all metrics (Supplementary Table S1).

Bland-Altman Analysis

The Bland-Altman plot demonstrated a mean difference (mean bias) of -0.172 cm^3, indicating an overestimation of the volume in the segmentation. The lower and upper confidence interval lines were positioned at -1.034 cm^3 and 0.690 cm^3, respectively (Fig. 3). Points above the upper limit accounted

for 0.949% (N = 9 eyes: six in Group 1, two in Group 2, and one in Group 3), whereas those below the lower limit accounted for 2.426% (N = 23 eyes: four in Group 1, five in Group 2, and 14 in Group 3).

Age and Sex Differences in Eyeball Volume

Age and sex differences in eyeball volume were examined using ground truth and 3D U-Net data from 98 emmetropic and ametropic eyes of 49 normal

subjects with forward-gaze images (Fig. 4). Significant sex differences in eyeball volumes were observed for both the ground truth and the 3D U-Net, as summarized in the Table. Ground truth showed average male and female values of 9.277 ± 1.318 cm^3 and 8.434 ± 1.206 cm^3, respectively ($t = 3.22$, $P < 0.05$). Meanwhile, 3D U-Net revealed average male and female values of 9.439 ± 1.252 cm^3 and 8.484 ± 1.183 cm^3, respectively ($t = 782$, $P < 0.001$). A significant age effect on eyeball volume was found in both the ground truth ($t = -2.866$, $P < 0.05$) and 3D U-Net ($t = -3.188$, $P < 0.05$) after controlling for sex. However, no significant correlation between age and eyeball volume was observed in the male group, whereas a significant correlation was observed in the female group (ground truth: $r = -0.608$, $P < 0.001$; 3D U-Net: $r = -0.587$, $P < 0.001$).

Eyeball Volume Consistency Across Gaze Directions

Gaze-dependent eyeball volume was analyzed in 464 eyes, including emmetropic and ametropic eyes, from 49 normal subjects across five different gazing directions. No significant volume differences were observed between the two segmentation methods for either gaze direction, nor were there any significant volume differences with respect to the gaze direction within each segmentation method. However, statistically significant differences were found among the five gazing directions in the performance evaluation metrics when comparing the ground truth and 3D U-Net segmentation (Supplementary Table S1). For a visual representation of the volume analysis, please refer to Supplementary Figure S2B.

Discussion

In this study, we developed a fully automated segmentation tool based on a 3D U-Net model, and evaluated its performance using a large dataset of 474 T2-weighted MRIs. Our results showed that 3D U-Net outperformed 2D U-Net in terms of segmentation performance across all evaluation metrics. Moreover, we observed significant sex differences in eyeball volume and an effect of age on eyeball volume, consistent with previous literature.[4] Our investigation highlights the robustness of the 3D U-Net segmentation method for different gaze directions, including nonpathological eyes, such as emmetropic or ametropic eyes, and pathological conditions.

We assessed the robustness and consistency of eyeball volume estimation using the 3D U-Net method across different gaze directions and explored the potential impact of gaze direction on segmentation performance compared with a custom 2D U-Net using multiview aggregation. However, this approach did not yield a significant improvement in this study, suggesting that the multi-view method might not be optimal because of missing connected voxels caused by eye-motion artifacts. In contrast, 3D U-Net successfully connected these voxels without creating fragments. Our findings emphasize the importance of considering gaze-direction variations when evaluating ocular structures and their impact on segmentation performance. Despite the challenges in comparing our results with those of previous studies, recent studies have reported high accuracy using similar techniques.[10,11] The clinical applicability of the 3D U-Net approach was demonstrated through a Bland–Altman plot analysis, which revealed an acceptable level of agreement between the ground truth and 3D U-Net predictions in terms of eyeball volumes, even though the 3D U-Net exhibited a tendency to slightly overestimate the total eye volume.

We detected significant differences among the five gazing directions in the performance evaluation metrics, indicating spatial overlap discrepancies between the ground truth and the 3D U-Net segmentation. Notably, the upward gazing direction showed a significantly lower performance than the other directions in normal subjects (Group 1). A possible explanation for this low spatial agreement in the vertical gaze could be related to the range of eye movement.[24,25] Previous studies have shown that vertical gaze ranges are asymmetric, with the upward gaze being significantly smaller than the downward gaze angle, whereas horizontal gazes remained consistent. Maintaining a specific gaze direction was challenging for participants during scanning. In particular, upward gazing could be uncomfortable owing to the restricted angle range of eye movement, which may lead to movement artifacts in MRIs.

This study has several limitations. First, instead of using traditional manual segmentation for the ground truth by drawing on a slice-by-slice basis, we opted for a semiautomatic tool. Consequently, we could not evaluate the intraobserver and interobserver variability of the semiautomatic tool. The segmentation procedure was performed carefully, and all MRIs were manually inspected. Second, owing to the limited clinical information available for the patient group, our eyeball volume analysis of age, sex, and gazing direction was restricted to normal subjects (Group 1). This limits the generalizability of our findings to the patient group. Third, despite using a large dataset from a single

center and using 10-fold cross-validation to improve generalizability, we did not conduct external validation. Further research is needed to explore the utility of our model in different contexts, such as different imaging modalities and ocular conditions. In particular, multi-class segmentation approaches incorporating both T1 and T2 images may improve the segmentation of the cornea and vitreous humor by providing clearer information on their boundary lines. By incorporating this information into our method, we hope to expand its applicability to situations requiring separate analyses of these structures.

In conclusion, this study presents a pioneering contribution to the field of human eyeball volume segmentation by developing a fully automated tool using a 3D U-Net for T2-weighted MRI. Our findings demonstrate the robustness and reliability of 3D U-Net across various gaze directions and diverse study populations, including different age groups, pathological patients, and normal subjects. An improved understanding of the impact of gaze direction on segmentation performance will enable researchers and clinicians to better utilize segmentation approaches for a more accurate and efficient assessment of the human eyeball, ultimately leading to enhanced ophthalmic diagnoses and treatment strategies. Future studies should focus on refining the model and exploring its utility in different contexts, such as different imaging modalities, and the impact of various ocular conditions on model performance.

Acknowledgments

Disclosure: **J.-J. Yang,** None; **K.H. Kim,** None; **J. Hong,** None; **Y. Yeon,** None; **J.Y. Lee,** None; **W.J. Lee,** None; **Y.J. Kim,** None; **J.-M. Lee,** None; **H.W. Lim,** None

* JJ Yang and KH Kim authors contributed equally and share first authorship.
‡ JM Lee and HW Lim contributed equally and share corresponding authorship.

References

1. Hallinan JT, Pillay P, Koh LH, Goh KY, Yu WY. Eye globe abnormalities on MR and CT in adults: an anatomical approach. *Korean J Radiol.* 2016;17:664–673.
2. Lim LS, Yang X, Gazzard G, et al. Variations in eye volume, surface area, and shape with refractive error in young children by magnetic resonance imaging analysis. *Invest Ophthalmol Vis Sci.* 2011;52:8878–8883.
3. Singh KD, Logan NS, Gilmartin B. Three-dimensional modeling of the human eye based on magnetic resonance imaging. *Invest Ophthalmol Vis Sci.* 2006;47:2272–2279.
4. Chau A, Fung K, Pak K, Yap M. Is eye size related to orbit size in human subjects? *Ophthal Physl Opt.* 2004;24:35–40.
5. Tatewaki Y, Mutoh T, Omodaka K, et al. Morphological prediction of glaucoma by quantitative analyses of ocular shape and volume using 3-dimensional T2-weighted MR images. *Sci Rep.* 2019;9:15148.
6. Hoang QV, Chang S, Yu DJG, Yannuzzi LA, Freund KB, Grinband J. 3-D assessment of gaze-induced eye shape deformations and downgaze-induced vitreous chamber volume increase in highly myopic eyes with staphyloma. *Br J Ophthalmol.* 2021;105:1149–1154.
7. Sater SH, Sass AM, Rohr JJ, et al. Automated MRI-based quantification of posterior ocular globe flattening and recovery after long-duration spaceflight. *Eye (Lond).* 2021;35:1869–1878.
8. Xing Q, Wei Q. Human eyeball model reconstruction and quantitative analysis. *IEEE Eng Med Bio.* 2014;2014:2460–2463.
9. Yu X, Ma W, Liu B, et al. Morphological analysis and quantitative evaluation of myopic maculopathy by three-dimensional magnetic resonance imaging. *Eye (Lond).* 2018;32:782–787.
10. Strijbis VIJ, de Bloeme CM, Jansen RW, et al. Multi-view convolutional neural networks for automated ocular structure and tumor segmentation in retinoblastoma. *Sci Rep-Uk.* 2021;11:14590.
11. Nguyen HG, Pica A, Maeder P, et al. Ocular structures segmentation from multi-sequences MRI using 3D Unet with fully connected CRFs. *Lect Notes Comput Sc.* 2018;11039:167–175.
12. Niendorf T, Beenakker JM, Langner S, et al. Ophthalmic magnetic resonance imaging: where are we (heading to)? *Curr Eye Res.* 2021;46:1251–1270.

13. Punn NS, Agarwal S. Modality specific U-Net variants for biomedical image segmentation: a survey. *Artif Intell Rev*. 2022;55:5845–5889.

14. Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. *Med Image Comput Comp Assist Interv III*. 2015;9351:234–241.

15. Çiçek Ö, Abdulkadir A, Lienkamp SS, Brox T, Ronneberger O. 3D U-Net: learning dense volumetric segmentation from sparse annotation. *International conference on medical image computing and computer-assisted intervention*. Berlin: Springer; 2016:424–432.

16. Hwang H, Rehman HZU, Lee S. 3D U-Net for skull stripping in brain MRI. *Appl Sci*. 2019;9: 569.

17. Hong J, Yun HJ, Park G, et al. Fetal cortical plate segmentation using fully convolutional networks with multiple plane aggregation. *Front Neurosci-Switz*. 2020;14:591683.

18. Moon Y, Lee WJ, Shin SH, et al. Quantitative analysis of translatory movements in patients with horizontal strabismus. *Invest Ophthalmol Vis Sci*. 2021;62:24.

19. Moon Y, Shin SH, Lee JY, et al. Quantitative analysis of eyeball rotation during lateral gaze in intermittent exotropia: a magnetic resonance imaging study. *Transl Vis Sci Technol*. 2021;10:20.

20. Moon Y, Lee WJ, Shin SH, et al. Positional change of the eyeball during eye movements: evidence of translatory movement. *Front Neurol*. 2020;11:556441.

21. Yadav S, Shukla S. Analysis of k-fold cross-validation over hold-out validation on colossal datasets for quality classification. *Int Conf Adv Comput* 2016;78–83.

22. Yushkevich PA, Piven J, Hazlett HC, et al. User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. *Neuroimage*. 2006;31:1116–1128.

23. Taha AA, Hanbury A. Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool. *BMC Med Imaging*. 2015;15:29.

24. Lee WJ, Kim JH, Shin YU, Hwang S, Lim HW. Differences in eye movement range based on age and gaze direction. *Eye (Lond)*. 2019;33:1145–1151.

25. Lim HW, Lee DE, Lee JW, et al. Clinical measurement of the angle of ocular movements in the nine cardinal positions of gaze. *Ophthalmology*. 2014;121:870–876.