Automated Measurement of Ocular Movements Using Deep Learning-Based Image Analysis

Lixia Lou\*a, Yiming Sun\*a, Xingru Huang\b, Kai Jin\a, Xiajing Tang\a, Zhaoyang Xu\c, Qianni Zhang\b, Yaqi Wang\d, and Juan Ye\a

\a Department of Ophthalmology, The Second Affiliated Hospital of Zhejiang University, School of Medicine, Hangzhou, China; \b School of Electronic Engineering and Computer Science, Queen Mary University of London, London, The United Kingdom; \c Wellcome-MRC Cambridge Stem Cell Institute, University of Cambridge, Cambridge, The United Kingdom; \d College of Media Engineering, Communication University of Zhejiang, Hangzhou, China

ABSTRACT

Purpose: Clinical assessment of ocular movements is essential for the diagnosis and management of ocular motility disorders. This study aimed to propose a deep learning-based image analysis to automatically measure ocular movements based on photographs and to investigate the relationship between ocular movements and age.

Methods: 207 healthy volunteers (414 eyes) aged 5–60 years were enrolled in this study. Photographs were taken in the cardinal gaze positions. Ocular movements were manually measured based on a modified limbus test using ImageJ and automatically measured by our deep learning-based image analysis. Correlation analyses and Bland-Altman analyses were conducted to assess the agreement between manual and automated measurements. The relationship between ocular movements and age were analyzed using generalized estimating equations.

Results: The intraclass correlation coefficients between manual and automated measurements of six extraocular muscles ranged from 0.802 to 0.848 (P < 0.001), and the bias ranged from –0.63 mm to 0.71 mm. The average measurements were 8.62 ± 1.07 mm for superior rectus, 7.77 ± 1.24 mm for inferior oblique, 6.99 ± 1.23 mm for lateral rectus, 6.71 ± 1.22 mm for medial rectus, 6.81 ± 1.20 mm for inferior rectus, and 6.63 ± 1.37 mm for superior oblique, respectively. Ocular movements in each cardinal gaze position were negatively related to age (P < 0.05).

Conclusions: The automated measurements of ocular movements using a deep learning-based approach were in excellent agreement with the manual measurements. This new approach allows objective assessment of ocular movements and shows great potential in the diagnosis and management of ocular motility disorders.

ARTICLE HISTORY

Received 8 October 2021
Accepted 10 March 2022

KEYWORDS

Ocular movements; automated measurements; deep learning; image analysis; age

Introduction

Clinical assessment of ocular movements is essential for the diagnosis and management of ocular motility disorders, and particularly important in incomitant strabismus. Six cardinal positions of gaze are identified in which one muscle in each eye is principally responsible for moving the eye into that position as follows: adduction (medial rectus [MR]), abduction (lateral rectus [LR]), suprabduction (superior rectus [SR]), supraduction (inferior oblique [IO]), infraduction (inferior rectus [IR]), and infraduction (superior oblique [SO]). Traditionally, clinicians grade hyperfunction and hypofunction of extraocular muscles using a qualitative scale based on subjective criteria, so much of that depend on clinicians’ experience. To circumvent this problem, many quantitative methods either kinetic (e.g. the limbus test and the lateral version light-reflex test) or static (e.g. Hess and Lancaster screen) have been proposed, although no method has been advocated in the literature as the gold standard. Mai described a modified limbus test measuring maximal distances from the limbus to the eyelid margin at an angle of 45 degrees to the horizontal in the positions of suprabduction, supraduction, infraduction, and infraduction, as intuitive reflections of the function of vertical rectus and oblique muscles. This method avoids complicated measurement and calculation of the angle of ocular movements and thus could be easily and conveniently implemented in clinical practice.

Many previous studies have attempted to quantitatively measure ocular movements based on photographs taken in the cardinal positions of gaze. However, manual measurement in the photographic analysis was required in these studies, under which circumstances interobserver variability...
still existed. As is known to all, deep learning with convolutional neural networks (CNN) has reached ideal performance for automatic ophthalmological image segmentation.\textsuperscript{11} In a previous study, we have analyzed morphologic features of eyelids in normal participants based on photographs using CNN-based deep learning methods.\textsuperscript{12} That approach with excellent reliability and reproducibility showed great potential for automated evaluation of eyelid-related disorders. To further explore the application of deep learning in the field of ocular motility disorders, here we introduced a new technique to automatically measure ocular movements using deep learning-based image analysis, according to the modified limbus test proposed by Mai.\textsuperscript{5} In addition, the relationship between ocular movements and age in healthy volunteers was explored.

Materials and Methods

Study participants

207 healthy volunteers were recruited from the Department of Ophthalmology, the Second Affiliated Hospital of Zhejiang University, School of Medicine between November 2020 and April 2021. The exclusion criteria consisted of strabismus, eyelid diseases, orbital diseases, previous ocular or periocular surgery, history of neurological diseases, and age above 60 years old. Informed consent was obtained according to a protocol conforming to the Declaration of Helsinki and approved by the Institutional Review Board at the Second Affiliated Hospital of Zhejiang University.

In this study, 30,000 facial images from the CelebFaces Attributes Dataset were used to train the eye location network,\textsuperscript{13} facial images of 1862 volunteers were used to train the eye segmentation network, and facial images of 207 healthy volunteers in nine diagnostic positions of gaze were used as the test set.

Photography

Binocular movement testing was conducted by a single experienced ophthalmologist to ensure consistency in assessment. The volunteer was asked to follow an object presented by the examiner, from the primary position to dextroversion, levoversion, supraversion, infraversion, dextrosupraversion, levo-supraversion, dextroinfraversion, and levo-infraversion positions of gaze, in line with standard clinical practice as described by Vivian and Morris.\textsuperscript{2}

With the volunteer’s head aligned horizontally, photographs were taken in nine diagnostic positions of gaze using a digital camera (Canon 1500 D, Canon Corporation, Japan) which was placed 100 cm away at eye level (Figure 1). Verbal encouragement was given to ensure the stability of the head and maximum effort toward the extremes of gaze. In the infraversion, the upper eyelids were pulled for better observation. The measurements of IR and SO were relative to the lower eyelids, so the upper eyelids pulling did not impact the measurements. A circular marker with a diameter of 10 mm was attached to the forehead of the volunteer as a reference of distance.

Manual photographic measurement

After photographs of the nine diagnostic positions were collected, manual measurement of the images was conducted using ImageJ (version 1.52; National Institutes of Health, Bethesda, USA) by another experienced ophthalmologist, according to the grading system of extraocular muscles adapted by Mai (Figure 1).\textsuperscript{5} The six cardinal positions of gaze

Figure 1. Photographs of the nine diagnostic positions of gaze. Taking the right eye for example, measurements of six extraocular muscles: medial rectus = D1-D2; lateral rectus = D4-D3; superior rectus = D5; inferior oblique = D6; inferior rectus = D7; superior oblique = D8.
correspond respectively to six extraocular muscles which are principally responsible for moving the eye into that position. The distance from the medial canthus to the temporal limbus in the primary position minus the distance in the adducted position was regarded as the measurement of MR. The distance from the medial canthus to the nasal limbus in the abducted position minus the distance in the primary position was regarded as the measurement of LR. The maximal distances from the limbus to the eyelid margin at an angle of 45 degrees to the horizontal in the positions of suprabduction, supraduction, infraduction, and infraction were regarded as the measurements of SR, IO, IR, and SO respectively, with longer distances indicating less amount of ocular movement. Accurate angle estimation was achieved by prior practice.

**Automated photographic measurement**

Recurrent residual convolutional neural networks with attention gate connection based on U-Net (R2AU-Net) were adopted for eye location and eye segmentation in this study. R2AU-Net architecture was illustrated in Figure S1. Figure 2 showed the workflow of automated measurement of extraocular muscles using the deep learning method.

Step 1: A total of 30,000 facial images (60,000 eyes) with landmark locations from the CelebFaces Attributes Dataset, were used to train the eye location network via the first-stage R2AU-Net framework. Parameter settings of the network model: epoch = 200; batch size = 4; input image size = 512 × 512 pixels; logistic loss function: BCE loss; optimizer: Adam (lr = 0.00001).

Step 2: Facial images of 1862 volunteers (3724 eyes) who visited our Department of Ophthalmology during 2017 and 2021 were collected. All volunteers had no eyelid or cornea diseases. Two ophthalmologists were invited to outline the eyelid margin and the corneal limbus. These facial images were used to train the eye segmentation network via the second-stage R2AU-Net framework. Parameter settings of the network model: epoch = 200; batch size = 4; input image size = 256 × 256 pixels; logistic loss function: L1 loss; optimizer: Adam (lr = 0.00001).

![Figure 2](image.png)
Step 3: Facial images of 207 healthy volunteers in nine diagnostic positions of gaze were used as the test set. Two ophthalmologists were asked to outline the eyelid margin and the corneal limbus as manual eye segmentation. Pre-processing methods such as random multi-scale boosting, elastic transformation, color perturbation, and random rotation were adopted, in order to achieve higher robustness when segmenting the eye region. The output result of R2AU-Net was smoothed to obtain the eyelid mask and corneal limbus mask.

Step 4: The measurements of six extraocular muscles in pixels based on masked images were automatically conducted, according to the grading system of extraocular muscles adapted by Mai (Figure S2).\(^5\)

Step 5: After adaptive threshold segmentation of the circular marker (10 mm in diameter) on the volunteer’s forehead, the pixel/millimeter ratios were calculated and the measurements of six extraocular muscles were converted into millimeters.

**Statistical analyses**

The accuracy for automated eye segmentation tasks was evaluated using dice coefficients, by comparing automated to manual eye segmentation. Right and left extraocular muscles measurements by automated and manual methods respectively were compared using a \(T\) test. Pearson’s correlation coefficients were calculated to measure the strength of the linear relationship between automated and manual measurements of six extraocular muscles. The agreement between the two measurements was evaluated using intra-class correlation coefficients (ICCs).\(^15\) It was considered excellent agreement if 0.80 < ICC ≤ 1.00, substantial agreement if 0.60 < ICC ≤ 0.80, and moderate agreement if 0.41 < ICC ≤ 0.60. The agreement between automated and manual measurements was also represented in Bland-Altman plots showing the difference between the two measurements against the mean of the two measurements.\(^16\) Generalized estimating equations were used to evaluate the relationship between age and measurements of six extraocular muscles (mean measurements of the two methods), adjusting for the dependence of the intrindividul data. All statistical analyses were conducted using SPSS (version 23; IBM Corporation, Chicago, USA). \(P\) values of < 0.05 were considered statistically significant.

**Results**

In total, 414 eyes of 207 normal participants, including 88 males (42.5%) and 119 females (57.5%), were included in this study. All participants were of Asian ethnicity. The mean age was 23.2 ± 12.9 years old, ranging from 5 to 60 years old. The dice coefficients for automated eye segmentation tasks in the test set of 414 eyes were 0.947 for the eyelid and 0.952 for the cornea, respectively. The mean time of automated measurement for each participant was 4.5 ± 0.3 seconds.

Statistical analyses showed significant negative correlations between age and measurements of MR (\(\beta = -0.012, P < 0.05\)) and LR (\(\beta = -0.013, P < 0.05\)), and significant positive correlations between age and measurements of SR (\(\beta = 0.010, P < 0.05\)), IO (\(\beta = 0.039, P < 0.001\)), IR (\(\beta = 0.013, P < 0.05\)) and SO (\(\beta = 0.019, P < 0.01\)), which implied decreased ocular movements in the six cardinal positions of gaze with aging.

**Discussion**

In the present study, we proposed a deep learning-based image analysis to automatically measure ocular movements...
The automated measurements of six extraocular muscles were in excellent agreement with the manual measurements. This study measured the normative values of ocular movements in six cardinal gaze positions using a modified limbus test and found that ocular movements in each gaze position were negatively related to age.

Accurate and consistent assessment of ocular movements is particularly important in evaluating treatment effects when a patient sees different clinicians at different visits.

Figure 3. Scatterplots of automated and manual measurements of (A) superior rectus; (B) inferior oblique; (C) lateral rectus; (D) medial rectus; (E) inferior rectus; and (F) superior oblique.
Typically, ocular movements are subjectively graded using simple scales (e.g. −4 to 0 to 4). Such methods are apt to standardization errors and are less suitable for accurate quantification. The limbus test was firstly proposed by Kenstenbaum. He measured ocular movements with a transparent millimeter-scale ruler placed in front of the cornea, by comparing the position of the limbus from the primary to the secondary and tertiary positions of gaze, which is convenient to implement in clinics. However, the reliability of the test results depends on the experience of clinicians, due to the learning curve effect. Many ophthalmic devices have been used to quantify ocular movements, but they are either time consuming (e.g. manual perimeter), costly (e.g. scleral search coil), or have limited measurement range (e.g. synoptophore). Digital photography was regarded as an integral part of the examination in strabismus clinics. The photographic technique has several advantages over the above-mentioned techniques, including ease of acquisition, less effort of cooperation, and allowance of objective assessment.

In 2001, Holmes et al. put forward a photographic method to manually evaluate abduction deficit in 23 patients with sixth nerve palsy. The measurement was based on photographs taken in the primary and abductive gaze positions. The distance from the medial canthus to the nasal limbus was measured on each photograph, just as we did in this study. At the time, their method was considered simple and effective. However, it is time-consuming and not practical in clinical settings, if there are many photographs of patients with ocular motility disorders (including but not limited to sixth nerve palsy) to be addressed manually. More recently, some researchers have measured ocular movements

Figure 4. Bland-Altman plots analyzing the agreement between automated and manual measurements of (A) superior rectus; (B) inferior oblique; (C) lateral rectus; (D) medial rectus; (E) inferior rectus; and (F) superior oblique. Broken lines indicate arithmetic mean (µ); dotted lines indicate 95% limits of agreement (1.96σ).
in normal subjects, patients with Graves’ orbitopathy and patients with inferior oblique muscle overaction, by covering the primary position image with a semitransparent layer of another cardinal position image.\(^6\)\(^,\)\(^7\)\(^,\)\(^9\)\(^,\)\(^10\) This was done to identify the margin of the limbus. A limitation of the method was that the perfect overlap of two images could not be guaranteed. In addition, a manual process was required to quantify the distance of limbus-to-limbus, which might lead to interobserver variability. A distinct advantage of the modified limbus test conducted in this study was the ocular movement’s measurement in each image, much simpler and more intuitive than measurement in the overlapped images. Our deep learning method permitted rapid and accurate measurement for each patient within 5 seconds, without any need for a manual process by clinicians. Furthermore, R2AU-Net adopted in this study enhances the capability of integrating contextual information by using recurrent residual convolutional units instead of basic convolutional units in U-Net, and improves the representation ability of networks by adding attention gate in the skip connections.\(^14\) Thus, R2AU-Net has much better performance than other improved U-Net algorithms for medical image segmentation.

A few studies had reported no significant relationship between ocular movements and age.\(^6\)\(^,\)\(^19\) However, this study found that ocular movements were negatively related to age, which was consistent with the majority of previous reports of age-related decline in motility.\(^7\)\(^,\)\(^20\)–\(^22\) The primary utility of this finding was in the evaluation of elderly patients with presumed extraocular muscle palsy. Since symmetry of binocular movements had been found in healthy volunteers, a slight degree of bilaterally symmetric hypofunction of extraocular muscles might represent a normal aging phenomenon, which deserves more attention in clinical practice. Degenerative neuronal loss of motor neurons has been suggested as a possible mechanism of age-related changes.\(^7\) Further neurobiological research is needed to explore whether aging mechanisms of ocular movements are related to central or peripheral factors.

Several limitations should be noted in this study. Firstly, participants with eyelid diseases were excluded from this study, because that abnormality of eyelid function or morphology (e.g. eyelid retraction, blepharoptosis and epicanthus) would cause ocular measurements far from the real values. Considering a person’s eyelid morphology is relatively stable, this method still allows intraindividual comparison over time, especially when assessing treatment effects in patients with ocular motility disorders. Secondly, the effect of eyeball size had not been considered in this study, due to the lack of axial length measurements. Further estimation of angles of ocular movements would permit more comprehensive interindividual comparison. Thirdly, although this study found a negative relationship between ocular movements and age, only participants aged below 60 years were included in the analysis, for the purpose of avoiding the effects of dermatochalasis on ocular movements measurement, which is a common upper eyelid involutional change. Fourthly, our deep learning method had not been validated in populations with ocular motility disorders or populations of other ethnicities. When detecting eye movement abnormalities, clinicians might pay more attention to binocular symmetry, especially for patients with monocular abnormalities. However, considering that the normal range of ocular movements was \(>5\) mm among subjects and the maximum variability was close to 2 mm between automated and manual measurements, this approach is not intended to replace clinicians and should be cautiously used in clinical practice.

In summary, we presented a new image analysis technique to automatically measure ocular movements in healthy volunteers and found a negative relationship between ocular movements and age. Although only normal participants were included in the current study, this technique showed great potential in helping clinicians with the diagnosis and management of ocular motility disorders, such as paralytic or restrictive strabismus. Using only photographs, this technique could be easily implemented in clinical practice. Also, it would offer a possibility for telemedicine due to the simplicity of image transmission and storage.

**Disclosure statement**

No potential conflict of interest was reported by the author(s).

**Funding**

This work was supported by (1) National Natural Science Foundation of China (No. 82000948); (2) National Natural Science Foundation of China (No. 81870635); (3) Zhejiang Provincial Key Research and Development Plan (No. 2019C03020); (4) National Key Research and Development Program of China (No. 2019YFC0118400).

**Data availability statement**

The original image data cannot be made publicly available because they contain identifying patient information. Data are available upon request from the Second Affiliated Hospital of Zhejiang University, School of Medicine.

**References**

1. Bowling B. 2016. Kanski’s clinical ophthalmology. Edinburgh, England: Elsevier.
2. Vivian AJ, Morris RJ. 1993. Diagrammatic representation of strabismus. Eye. 7(4):565–571. doi:10.1038/eye.1993.123.
3. Kennard C. 2007. Examine eye movements. Pract Neurol. 7(5): 526–330. doi:10.1136/jnnp.2007.124388.
4. Hanif S, Rowe FJ, O’connor AR. 2009. A comparative review of methods to record ocular rotations. Br Ir Orthopt J. 6(0):47–51. doi:10.22599/bioj.8.
5. Mai G. 2005. Grading system for hyperfunction and hypofunction of extraocular muscles. Chin J Ophthalmol. 41(7):663–666.
6. Leite CA, Pereira TS, Chiang J, Pieroni Gonçalves AC, Monteiro MLR. 2020. Evaluation of ocular versions in Graves' orbitopathy: correlation between the qualitative clinical method and the quantitative photographic method. J Ophthalmol. 2020:1–7. doi:10.1155/2020/9758153.
7. Lee WJ, Kim JH, Shin YU, Hwang S, Lim HW. 2019. Differences in eye movement range based on age and gaze direction. Eye. 33(7):1145–1151. doi:10.1038/s41433-019-0376-4.

8. Yoon CK, Yang HK, Kim JS, Hwang JM. 2014. An objective photographic analysis of ocular oblique muscle dysfunction. Am J Ophthalmol. 158(5):924–931. doi:10.1016/j.ajo.2014.07.035.

9. Lim HW, Lee JW, Hong E, Song Y, Kang MH, Seong M, Cho HY, Oh SY. 2014. Quantitative assessment of inferior oblique muscle overaction using photographs of the cardinal positions of gaze. Am J Ophthalmol. 158(4):793–799. doi:10.1016/j.ajo.2014.06.016.

10. Lim HW, Lee DE, Lee JW, Kang MH, Seong M, Cho HY, Oh JE, Oh SY. 2014. Clinical measurement of the angle of ocular movements in the nine cardinal positions of gaze. Ophthalmology. 121(4):870–876. doi:10.1016/j.ophtha.2013.11.019.

11. Schmidt-Erfurth U, Sadeghipour A, Gerendas BS, Waldstein SM, Bogunovic H. 2018. Artificial intelligence in retina. Prog Retin Eye Res. 67:1–29. doi:10.1016/j.preteyeres.2018.07.004.

12. Cao J, Lou L, You K, Gao Z, Jin K, Shao J, Ye J. 2021. A novel automatic morphologic analysis of eyelids based on deep learning methods. Curr Eye Res. 46(10):1495–1502.

13. Liu ZW, Luo P, Wang XG, Tang XO. Deep learning face attributes in the wild. In: 2015 IEEE international conference on computer vision. New York: IEEE. 2015. 3730–8.

14. Žuo Q, Chen S, Wang Z. 2021. R2AU-Net: attention recurrent residual convolutional neural network for multimodal medical image segmentation. Secur Commun Netw. 2021:1–10. doi:10.1155/2021/6625688.

15. Chen CC, Barnhart HX. 2013. Assessing agreement with intraclass correlation coefficient and concordance correlation coefficient for data with repeated measures. Comput Stat Data Anal. 60:132–145. doi:10.1016/j.csda.2012.11.004.

16. Bunce C. 2009. Correlation, agreement, and Bland-Altman analysis: statistical analysis of method comparison studies. Am J Ophthalmol. 148(1):4–6. doi:10.1016/j.ajo.2008.09.032.

17. Kestenbaum A. 1947. Clinical methods of neuro-ophthalmologic examination. London, England: Butterworth-Heinemann.

18. Holmes JM, Hohberger GG, Leske DA. 2001. Photographic and clinical techniques for outcome assessment in sixth nerve palsy. Ophthalmology. 108(7):1300–1307. doi:10.1016/s0161-6420(01)00592-9.

19. Haggerty H, Richardson S, Mitchell KW, Dickinson AJ. 2005. A modified method for measuring uniconular fields of fixation: reliability in healthy subjects and in patients with Graves orbitopathy. Arch Ophthalmol. 123(3):356–362. doi:10.1001/archoph.123.3.356.

20. Clark RA, Isenberg SJ. 2001. The range of ocular movements decreases with aging. J Aapos. 5(1):26–30. doi:10.1067/mpa.2001.111016.

21. Oguro H, Okada K, Suyama N, Yamashita K, Yamaguchi S, Kobayashi S. 2004. Decline of vertical gaze and convergence with aging. Gerontology. 50(3):177–181. doi:10.1159/000076777.

22. Kang NY. 2009. Changes of ocular version with aging in normal Korean population. J Korean Med Sci. 24(4):708–714. doi:10.3346/jkms.2009.24.4.708.