Detection of Meibomian Gland Dysfunction by in vivo Confocal Microscopy Based on Deep Convolutional Neural Network

Yi Shao
Nanchang University
Yichen Yang
Tianjin Medical University Cancer Institute and Hospital
Hui Zhao
Shanghai Jiao Tong University
Wen-Qing Shi
Nanchang University
Xu-Lin Liao
The Chinese University of Hong Kong https://orcid.org/0000-0002-2800-9838
Ting Su
Harvard Medical School
Rong-Bin Liang
Nanchang University
Qiu-Yu Li
Nanchang University
Qian-Min Ge
Nanchang University
Hui-Ye Shu
Nanchang University
Yi-Cong Pan
Nanchang University
Xiangchun Li (lixiangchun@tmu.edu.cn)
Tianjin Medical University Cancer Institute and Hospital https://orcid.org/0000-0002-2230-2152

Article

Keywords: meibomian glands, in vivo confocal microscopy, deep learning

Posted Date: October 25th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-936418/v1

Additional Declarations: There is **NO** Competing Interest.

Abstract

In vivo confocal microscopy IVCM is a real-time in vivo high-resolution and non-invasive imaging method that allows observation of morphological changes in the lid gland at the cellular level. We use IVCM to observe the meibomian glands and divide the 12,630 pictures obtained into six groups (normal group, normal with meibomian gland opening group, meibomian gland atrophy group, meibomian gland atrophy with obstruction group, meibomian gland obstruction group and meibomian gland obstruction with opening group), randomly select 70% of the pictures and use the ResNet34 deep learning network model for training, and use the remaining 30% of the pictures and another 12889 pictures collected from the top three hospitals as internal and external validation sets to verify the performance of the model. The results show that the model validation set recognition constructed using the deep learning method has achieved good performance, and the obtained AUROCs are all greater than 0.95. This model provides the possibility for future automatic classification and diagnosis of meibomian gland dysfuncton(MGD),and can be used for MGD related clinical auxiliary diagnosis and screening of diseases.

Introduction

Meibomian glands are the largest sebaceous glands in the human body. They are distributed in the upper and lower meibomian layers, with 30–40 on the upper eyelid and 20–30 on the lower eyelid and are perpendicular to the eyelid margin. The meibomian gland can synthesize, store, and secrete lipids, and these secreted lipids constitute the lipid layer of the tear film. Meibomian gland dysfunction (MGD) is a chronic, non-specific inflammation characterized by meibomian gland duct obstruction or abnormal meibomian gland secretions. In MGD, the glandular lipid secretion is impaired, the ocular surface cannot maintain the stability of the tear film, the tear film evaporates rapidly, and the tear osmotic pressure increases, which can lead to dry eye disease (DED)[1,2]. With increasing age, the acinar epithelial cells of the meibomian gland shrink, leading to large-scale atrophy of the meibomian glands. As these glands are non-renewable, MGD reduces lipid secretion[3]. With an increase in the aging population, environmental pollution, and widespread use of video terminals, the incidence of MGD has gradually increased and become a disease of global proportions. The clinical symptoms of MGD are similar to those of DED and include dry eyes, eye irritation, blurred vision, and increased secretions. As the symptoms of MGD are not specific, it is easy to be missed, misdiagnosed, and mistreated. Therefore, objective physical signs can provide a reference for diagnosis, including the loss of meibomian glands, abnormal secretion of meibomian glands, and changes in the morphology of the eyelid margin.

In vivo confocal microscopy (IVCM) is a new type of high-resolution, non-invasive ocular surface imaging detector. The laser focusing principle of IVCM is used to scan each point of the detected layer in the confocal plane. In IVCM, the formation of ocular surface tissues can be observed at multiple levels, and the cell level of tissue morphology and changes can be photographed in real time[4]. Hence, IVCM has real-time, non-invasive, three-dimensional advantages. In recent years, as MGD has received increasing attention, there have been an increasing number of researches on the application of IVCM in the

meibomian glands[5]. IVCM can facilitate clear visualization of the characteristics of the meibomian glands, such as acinus, acinar stroma, gland secretions, periadenositis cells, and meibomian gland openings and classify them for standardized treatment. Ibrahim et al.[6] obtained IVCM images and analyzed the density of meibomian acinar units and the maximum and minimum diameters of the meibomian acinars; they concluded that IVCM has a high level of specificity and sensitivity in the diagnosis of MGD. However, because of the limitations of clinical manpower and material resources, the diagnosis of MGD-related eye diseases still has limitations at present, such as the limited surgical skills of junior doctors and the huge workload of manual classification, which are not conducive to the development of individualized disease treatment and chronic disease management. Recent advancements in the field of artificial intelligence (AI) have led to its increased application in disease assessment and diagnosis; further, AI has also become a research focus for clinical professionals. Analyzing MGD IVCM images data with high-performance AI algorithm can help distinguish different types of patients, greatly increase the work efficiency of clinicians, and save time by formulating different treatment plans for patients, in order to reduce the social and medical burden and aid with the new trend of personalized and refined MGD diagnosis and treatment. In recent years, the combined deep convolutional neural network (DCNN) model has been applied to the diagnosis of various medical diseases[7,8]. We propose a method that combines the DCNN algorithm and IVCM diagnosis for early diagnosis and classification of MGD.

Methods

Study Subjects

The Institutional Review Board of the Ethics Committee of the First Affiliated Hospital of Nanchang University approved this retrospective study, and all methods complied with the tenets of the Helsinki Declaration.

This was a single-center, clinical study. From January 2018 to June 2021, we extracted 12,630 IVCM images from the IVCM database of three hospitals, namely the First Affiliated Hospital of Nanchang University, Xinhua Hospital Affiliated to Shanghai Jiaotong University School of Medicine, and Shanghai General Hospital. Three ophthalmologists checked the quality of the images and performed clinical diagnosis and classification based on the 2017 Chinese Expert Consensus on the Diagnosis and Treatment of MGD[9]. The extracted data included the following: normal group (n=2896), normal with meibomian gland opening (NMGO) group (n=830), meibomian gland atrophy (MGA) group (n=3585), meibomian gland atrophy with obstruction (MGAO) group (n=1745), meibomian gland obstruction (MGO) group (n=3086), and meibomian gland obstruction with opening (MGOO) group (n=488). The six types of IVCM classification pictures can be seen in Figure 1. According to random grouping, 70% (n=8,841) and 30% (n=3,789) images were allocated to the training and internal validation sets, respectively. Simultaneously, we used 12,289 images from Nanchang University Affiliated Eye Hospital and Shanghai AIER Eye Hospital as an external verification set to verify the performance of the model. The DCNN classification model that maps the input features (such as image pixels) to the corresponding output

labels (MGD of different classifications) was used to train the deep learning algorithm. We used the ResNet34 network model for classification and initial training(120 rounds of training) and developed and trained the DCNN model on the training set. We enhanced the images from the training data set using brightness, gamma correction, histogram equalization, noise addition, and inversion. The process flow chart of this research is shown in Figure 2.

Classification Standard

Normal group: meibomian glands and ducts are normal.

NMGO group: full meibomian glands with opening.

MGA group: tire-like meibomian gland epithelial cells are absent, the acinar wall disappears, becomes smaller, or the fiber cord changes.

MGO group: enlarged acinus and blocked meibum in the acinus.

MGAO group: shows the characteristics of both the atrophy group and obstruction group.

MGOO group: normal meibomian gland opening with highly dilated meibomian gland secretory tube and blocked meibomian gland acinus.

In Vivo Confocal Microscope Image Acquisition

The IVCM (HRT-3, RCM module, Hidelberg Engineering, Germany) parameters were as follows: laser wavelength, 670 nm; observation field, 400 um*400 um; resolution, 384 pixels * 384 pixels; magnification, 800 times; and axis resolution, is 1 um. Use 0.5% proparacaine eye drops (American Alcon company) for topical anesthesia,and apply Carbomer Eye Gel (Shandong Bausch & Lomb Freda Pharmaceutical Co., Ltd.) on the surface of the IVCM contact lens. The disposable corneal contact cover is attached to the contact lens. To make the image clearer, the outer layer of the contact sleeve is evenly smeared with Carbomer Eye Gel. The examinee's chin and forehead are fixed at the corresponding positions of the IVCM, and the eyelid of the examinee's eye is turned over. The examiner fixes the eyelid with one hand to keep it in a state of eversion, and pushes the IVCM lens forward with the other hand to make it touch the edge of the examinee's eyelid, and adjusts the scan depth, starting from the eyelid margin and gradually scanning the meibomian glands in the fornix, from the nasal side of the lesion to the temporal side, while taking IVCM images.

Model Building

Convolutional neural network (CNN) is a multi-layer neural network with fault tolerance and is easy to train and optimize[10]. The residual network (ResNet) is a type of CNN, the network is connected by adding one operation channel in every two layers to form a residual block to reach a jump in the whole network,

avoiding the gradient descent problem and network degradation problem caused by the multiplication of very small parameters in each layer[11].

In this study, we chose a RESNET network (ResNet34) composed of 34 convolution layers to classify the MGD images. Image features were extracted using sixty-four 7×7 convolution kernels, and then a 3×3 maximum down-sampling operation with a stride of 2 was performed. Then, input the result into four residual groups with three, four, six, and three residual units, where each residual unit has two 3×3 convolutional layers with a stride of 2. After each convolutional layer, the linear rectification function (ReLU) is used as the activation function. It is mapped to the probability of each category by the fully connected layer and the Softmax activation function.

In this experiment, a desktop computer was used as the experimental environment: the Central Processing Unit (CPU) is Ryzen5 1500X (Advanced Micro Devices, Inc.), the memory is 16 G, the graphic processing unit (GPU) is Nvidia GeForce GTX1060, and the graphics card memory is 6 G. We used the PyTorch (https://py-torch.org) Python framework based on the Windows 10 system to develop and train the MGD image recognition model. First, we used the pre-trained generation weight initialization network on the ImageNet (http://pytorch.org/docs/stable/torchvision/models.html) dataset, and then transferred the model network to MGD. On the data set, the modeling process is divided into three parts: image preprocessing, image feature extraction and classification of MGD IVCM, and active learning. The optimizer is Adam. The model training parameters are as follows: batch size=16, epoch=120, and learning rate=0.1; the other parameters are default.

Model Evaluation and Statistical Analysis

After the MGD IVCM image recognition model is trained, the performance of the trained model needs to be evaluated, and the results of classification by three ophthalmologists from the First Affiliated Hospital of Nanchang University were used as the gold standard to calculate the area under the receiver operating characteristic curve (AUROC), accuracy, sensitivity, specificity, negative predictive value (NPV), Kappa coefficient, and F1 scores. Four statistical parameters are involved: true positive (TP), false positive (FP), true negative (TN), and false negative (FN). The calculation formula is as follows: accuracy= (TP+TN)/(TP+FP+FN+TN), specificity=TN/N (negative), precision (P)=TP/(TP+FP), sensitivity=TP/P, and NPV=TN/(TN+FN). The ROC curve was plotted with specificity as the abscissa and sensitivity as the ordinate; the larger the AUROC value, the better the classification performance of the model. The Kappa test was used to compare the consistency of the results of the artificial intelligence diagnosis group and the expert group and to evaluate the accuracy of the multi-classification model. All the above indicators were obtained using the R software (version 6.078).

Results

Data Set Characteristics

A total of 24,919 images from five hospitals were used to train and verify the performance of the DCNN. From the First Affiliated Hospital of Nanchang University, Xinhua Hospital Affiliated to Shanghai Jiaotong University School of Medicine, and Shanghai General Hospital, the training set consisted of 8841 images (1993 images in the normal group, 578 images in the NMGO group, 2500 images in the MGA group, 1228 images in the MGAO group, 2191 images in the MGO group, and 351 images in the MGOO group). Additionally, 3789 internal validation sets (903 in the normal group, 252 in the NMGO group, 1085 in the MGA group, 517 in the MGAO group, 895 in the MGO group, and 137 in the MGOO group), and a total of 12289 images were collected from the external validation set of Nanchang University Affiliated Eye Hospital and Shanghai AIER Eye Hospital(2912 in the normal group, 838 in the NMGO group, 3493 in the MGA group, 1693 in the MGAO group, 2866 in the MGO group, and 487 in the MGOO group). The total number, training set, internal and external validation set cohorts of the six types of MGD are summarized in Table 1.

Performance of Deep Learning Algorithms

We used the DCNN algorithm ResNet34 to train the classification model on the same training data set. Based on the 30% of randomly selected pictures for internal verification, the accuracy, sensitivity, specificity, precision, and AUROC parameters of the classification were calculated. The accuracy of the internal validation set test of the six groups of pictures ranged from 88.1% (95%CI: 87.1–89.2) to 93.9% (95%CI: 93.1–94.7); sensitivity, from 87.4% (95%CI: 85.0–89.5) to 93.4% (95%CI: 85.0–89.5); specificity, from 86.2% (95%CI: 84.9–87.5) to 94.3% (95%CI: 93.5–95.0); precision, from 33.8% (95%CI: 28.9–39.0) to 73.1% (95%CI: 70.7–75.5); AUROC, from 0.947 (95%CI: 0.939–0.955) to 0.971 (95%CI: 0.965–0.977). The accuracy of the six-category external validation set to verify the classification model ranged from 88.1% (95%CI: 87.5–88.6) to 93.9% (95%CI: 93.4–94.3); sensitivity, from 82.5% (95%CI: 78.9–85.8) to 94.0% (95%CI: 93.1–94.8); specificity, from 86.3% (95%CI: 85.6–87.0) to 94.3% (95%CI: 93.9–94.7); and AUROC, from 0.951 (95%CI: 0.941–0.961) to 0.971 (95%CI: 0.966–0.976). The built DCNN model internal and external validation set had high negative predictive value, Kappa coefficient, and F1 score value. The internal validation set test results are shown in the following table (Table 2). The test results of the external validation set are shown in Table 3. The ROC curve of the six classification validation sets is shown in Figures 3 and 4.

Discussion

In this study, we aimed to evaluate the performance of convolutional neural networks for detecting MGD using a large number of confocal microscopy images with different classifications. The results show that the combination of the DCNN model and confocal microscopy images that we tested in the validation set can finely classify MGD patients in a realistic setting with high accuracy, specificity, and AUROC values, compared to manual classification by ophthalmologists.

In recent years, with the rapid development of AI, its application in clinical practice has become increasingly widespread[12,13]. At present, AI-assisted diagnosis of cataracts, early glaucoma, diabetic

retinopathy, age-related macular degeneration, and other ophthalmic diseases is developing rapidly (Table 4)[14-17]. However, studies on AI-assisted diagnosis or screening of dry eye or MGD are relatively rare. Compared with traditional manual diagnosis, the more mature AI diagnosis and segmentation has the following advantages: In terms of image characteristics such as brightness and image contrast, the recognition ability of AI is stronger than that of human eyes; moreover, it does not require the experience of and intervention by imaging physicians. The rate of misdiagnosis caused by subjective factors such as proficiency and fatigue can be avoided. Digital information storage facilitates the development of cases, data exchanges, and cooperation across different regions and fields.

Convolutional Neural Network is a deep learning model that has developed rapidly in recent years and received wide attention from various disciplines[18]. It is a pre-feedback neural network structure system that can be used for visual processing. The main structure is composed of multiple groups of units. It is an efficient computing network. In the CNN model, the ResNet model introduces the residual module[19], which solves the problem of gradient disappearance and model performance degradation as the network depth increases in deep learning. It has become the first choice for image classification tasks at present.

IVCM technology has been widely used in the diagnosis of ocular surface diseases. It can be used to observe and evaluate ocular surface inflammation, tissue damage, and nerve distribution at the biological level[20,21]. In 2005, Kobayashi et al.[22] used IVCM for the first time to observe the structural characteristics of the eyelid conjunctiva and lower eyelid meibomian glands in four healthy individuals. Subsequently, some international researchers used IVCM to observe the morphological changes of meibomian glands in healthy subjects and others with MGD, keratitis, Graves ophthalmopathy, and Sjogren's syndrome, and found that the diameter and density of the acinus, density of inflammatory cells in the acinus, and degree of meibomian gland fibrosis in patients with MGD were significantly abnormal compared with the healthy subjects[23-25]. The degree of keratosis of the meibomian gland openings and the distribution of acinar expansion and atrophy in patients with MGD are helpful indicators to explore the pathogenesis and influencing factors of MGD, and provide a basis for evaluating and detecting the therapeutic effect of MGD from the level of cell morphology. However, there may be some problems in the clinical diagnosis and treatment of MGD by IVCM, such as the large amount of work required for the examination, and low precision of the examination results resulting in a small correlation between the severity of the disease and treatment. These problems restrict the diagnosis and treatment level and the improvement of long-term management effect of DED or MGD.

At present, most of the research on AI in MGD mainly focuses on the application of an AI-based diagnosis system to the range of meibomian gland atrophy and to judge whether the patient has MGD according to the characteristics of the gland. Wang et al.[26] used deep learning to automatically segment and calculate the area of meibomian gland atrophy with an accuracy rate of 95.4% and provided quantitative information on the severity of gland atrophy. Koh et al.[27] used slit lamp microscopes equipped with infrared emission filters and infrared cameras combined with machine learning to train and classify based on the length and width characteristics of the glands, and reported a high specificity of 96.1%. In

this study, we chose to analyze IVCM images of outpatients from multiple hospitals through an AI-based diagnostic system to assess their consistency, accuracy, sensitivity, and specificity in an attempt to improve the diagnostic rate. Maruoka et al.[28] published a method for detecting MGD by IVCM based on a deep learning model, but the study was mainly a case-control study (MGD and non-MGD). To our knowledge, this study is the first to analyze a six-category MGD network model. We collected a larger number of IVCM images (12,630 images) than previous studies, trained the model with a large number of images, and validated it with data from several hospitals to make the data source more general and the amount of data sufficient. Using Resnet34 migration learning to the MGD dataset, the mean and standard deviation were selected from the parameters obtained by ImageNet dataset. The parameters obtained after extensive pre-training can better improve the accuracy of the model interpretation and the persuasiveness of the study results. The results obtained were as expected, with an accuracy rate of 86.1% or higher for both the internal and external validation sets, which is in high agreement with the ophthalmologist's diagnosis. The six groups of images had high specificity, ranging from 86.2% (95%CI: 84.9–87.5) to 94.3% (95%CI: 93.5–95.0), suggesting that each group had very low false negatives in the screening and classification diagnosis; this is clinically significant for the classification and diagnosis of MGD. The future integration of AI into confocal microscopy instruments will allow for rapid differentiation of disease types and better detection of disease progression, thus improving the efficiency of ophthalmologists and their ability to manage MGD, and promoting the further development of AI in aiding the diagnosis of ocular surface diseases.

However, in this study, our design has some limitations. Since multiple machine learning algorithms were not used (only ResNet34 was used in this study), we cannot guarantee whether the results we obtained are relevant to the chosen algorithm. In addition, although this study covers the data volume of five hospitals, it mainly focuses on the data from South China. In the future, we can collaborate with other hospitals across China to form a complete ' Data Model Diagnosis System'. Once the platform is up and running, doctors' diagnoses and patients' IVCM images will be reused as training sets to continuously update and optimize the diagnostic models.

In conclusion, the classification and evaluation of MGD using the DCNN model have good accuracy, which can help clinicians analyze the examination results better and faster, provide a more reliable basis for diagnosis, and support individualized treatment and chronic disease management of MGD and DED. This model is suitable for comprehensive ophthalmology clinics with a large number of visits and a shortage of physicians. It can also be applied to the screening of large populations, follow-up of MGD and DED patients, and observation of efficacy.

References

1. Liu Z*, et al.* Efficacy and Safety of Houttuynia Eye Drops Atomization Treatment for Meibomian Gland Dysfunction-Related Dry Eye Disease: A Randomized, Double-Blinded, Placebo-Controlled Clinical Trial. *J Clin Med* **9**, (2020).

2. Lemp MA, Crews LA, Bron AJ, Foulks GN, Sullivan BD. Distribution of aqueous-deficient and evaporative dry eye in a clinic-based patient cohort: a retrospective study. *Cornea* **31**, 472-478 (2012).

3. Kiyat P, Palamar M, Gerceker Turk B, Yagci A. Evaluation of dry eye and Meibomian gland dysfunction in female androgenetic alopecia patients. *Int Ophthalmol*, (2021).

4. Pedrotti E*, et al.* In Vivo Confocal Microscopy of the Corneal-Conjunctival Transition in the Evaluation of Epithelial Renewal after SLET. *J Clin Med* **9**, (2020).

5. Wei S, Ren X, Wang Y, Chou Y, Li X. Therapeutic Effect of Intense Pulsed Light (IPL) Combined with Meibomian Gland Expression (MGX) on Meibomian Gland Dysfunction (MGD). *J Ophthalmol* **2020**, 3684963 (2020).

6. Ibrahim OM*, et al.* The efficacy, sensitivity, and specificity of in vivo laser confocal microscopy in the diagnosis of meibomian gland dysfunction. *Ophthalmology* **117**, 665-672 (2010).

7. Jin C*, et al.* Development and evaluation of an artificial intelligence system for COVID-19 diagnosis. *Nat Commun* **11**, 5088 (2020).

8. Ting DSW*, et al.* Development and Validation of a Deep Learning System for Diabetic Retinopathy and Related Eye Diseases Using Retinal Images From Multiethnic Populations With Diabetes. *JAMA* **318**, 2211-2223 (2017).

9. HONG,X. Expert consensus on diagnosis and treatment of meibomian gland dysfunction in China. *Chin J Ophthalmol* **53**,657-661 (2017).

10. Llorella FR, Patow G, Azorin JM. Convolutional neural networks and genetic algorithm for visual imagery classification. *Phys Eng Sci Med* **43**, 973-983 (2020).

11. Zhang X, Jiang L, Yang D, Yan J, Lu X. Correction to: Urine Sediment Recognition Method Based on Multi-View Deep Residual Learning in Microscopic Image. *J Med Syst* **44**, 84 (2020).

12. Devalla SK*, et al.* Glaucoma management in the era of artificial intelligence. *Br J Ophthalmol* **104**, 301-311 (2020).

13. Moraru AD, Costin D, Moraru RL, Branisteanu DC. Artificial intelligence and deep learning in ophthalmology - present and future (Review). *Exp Ther Med* **20**, 3469-3473 (2020).

14. Long E*, et al.* An artificial intelligence platform for the multihospital collaborative management of congenital cataracts. *Nature Biomedical Engineering* **1**, (2017).

15. Panda R, Puhan NB, Rao A, Mandal B, Padhy D, Panda G. Deep convolutional neural network-based patch classification for retinal nerve fiber layer defect detection in early glaucoma. *J Med Imaging (Bellingham)* **5**, 044003 (2018).

16. Voets M, Mollersen K, Bongo LA. Reproduction study using public data of: Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *PLoS One* **14**, e0217541 (2019).

17. Li Z, He Y, Keel S, Meng W, Chang RT, He M. Efficacy of a Deep Learning System for Detecting Glaucomatous Optic Neuropathy Based on Color Fundus Photographs. *Ophthalmology* **125**, 1199-1206 (2018).

18. Torsten P, Jan F, Arthur K, Henner S. Using Local Convolutional Neural Networks for Genomic Prediction. *Frontiers in Genetics*, (2020).

19. Woldegiorgis S, Enqvist A, Baciak J. ResNet and CycleGAN for pulse shape discrimination of He-4 detector pulses: Recovering pulses conventional algorithms fail to label unanimously. *Appl Radiat Isot* **176**, 109819 (2021).

20. Belmonte C*, et al.* TFOS DEWS II pain and sensation report. *Ocul Surf* **15**, 404-437 (2017).

21. Ucar IC, Esen F, Turhan SA, Oguz H, Ulasoglu HC, Aykut V. Corneal neuropathic pain in irritable bowel syndrome: clinical findings and in vivo corneal confocal microscopy. *Graefes Arch Clin Exp Ophthalmol*, (2021).

22. Kobayashi A, Yoshita T, Sugiyama K. In vivo findings of the bulbar/palpebral conjunctiva and presumed meibomian glands by laser scanning confocal microscopy. *Cornea* **24**, 985-988 (2005).

23. F F*, et al.* Facial Seborrheic Keratosis with Unusual Dermoscopic Patterns can be differentiated from other skin malignancies by In Vivo Reflectance Confocal Microscopy. *Journal of the European Academy of Dermatology and Venereology : JEADV*, (2021).

24. Shengnan C, Yueqi Y, Jin C, Lin Y, Xinghua W, Fagang J. In vivo confocal microscopy assessment of meibomian glands microstructure in patients with Graves' orbitopathy. *BMC ophthalmology* **21**, (2021).

25. Szalai E, Szucs G, Szamosi S, Aszalos Z, Afra I, Kemeny-Beke A. An in vivo confocal microscopy study of corneal changes in patients with systemic sclerosis. *Sci Rep* **11**, 11111 (2021).

26. Wang J, Yeh TN, Chakraborty R, Yu SX, Lin MC. A Deep Learning Approach for Meibomian Gland Atrophy Evaluation in Meibography Images. *Transl Vis Sci Technol* **8**, 37 (2019).

27. Koh YW, Celik T, Lee HK, Petznick A, Tong L. Detection of meibomian glands and classification of meibography images. *J Biomed Opt* **17**, 086008 (2012).

28. Maruoka S*, et al.* Deep Neural Network-Based Method for Detecting Obstructive Meibomian Gland Dysfunction With in Vivo Laser Confocal Microscopy. *Cornea* **39**, 720-725 (2020).

29. Zhongwen L*, et al.* Preventing corneal blindness caused by keratitis using artificial intelligence. *Nature communications* **12**, (2021).

30. Zhang K*, et al.* Deep-learning models for the detection and incidence prediction of chronic kidney disease and type 2 diabetes from retinal fundus images. *Nat Biomed Eng* **5**, 533-545 (2021).

31. Yang HK, Kim YJ, Sung JY, Kim DH, Kim KG, Hwang J-M. Efficacy for Differentiating Nonglaucomatous Versus Glaucomatous Optic Neuropathy Using Deep Learning Systems. *American Journal of Ophthalmology* **216**, (2020).

32. Khan SMA, Jens H, Stefan S, E SM, Philipp S. Deep learning-based automatic meibomian gland segmentation and morphology assessment in infrared meibography. *Scientific reports* **11**, (2021).

33. Daisuke N*, et al.* Automated detection of a nonperfusion area caused by retinal vein occlusion in optical coherence tomography angiography images using deep learning. *PloS one* **14**, (2019).

Tables

Table 1 The total number of six types of MGD, training set ,internal and external validation set cohorts

Classification	ALL (n=12630)	Train set (n=8841)	Internal test set (n=3789)	External test set (n=12289)
Normal group	2896 (22.9%)	1993 (22.5%)	903 (23.8%)	2912 (23.7%)
NMGO group	830 (6.6%)	578 (6.5%)	252 (6.7%)	838 (6.8%)
MGA group	3585 (28.4%)	2500 (28.3%)	1085 (28.6%)	3493 (28.4%)
MGAO group	1745 (13.8%)	1228 (13.9%)	517 (13.6%)	1693 (13.8%)
MGO group	3086 (24.4%)	2191 (24.8%)	895 (23.6%)	2866 (23.3%)
MGOO group	488 (3.9%)	351 (4.0%)	137 (3.6%)	487 (4.0%)

Abbreviations: NMGO: normal with meibomian gland opening; MGA: meibomian gland atrophy; MGAO: meibomian gland atrophy with obstruction; MGO: meibomian gland obstruction; MGOO: meibomian gland obstruction with opening

Table 2 Performance metrics of the ensemble DCNN model evaluated on the internal validation set

Performance metrics	Internal test set (n=3789)					
	Normal group (n=903)	NMGO group (n=252)	MGA group (n=1085)	MGAO group (n=517)	MGO group (n=895)	MGOO group (n=137)
Accuracy (95% CI)	0.893 (0.883 - 0.903)	0.939 (0.931 - 0.947)	0.883 (0.872 - 0.893)	0.914 (0.905 - 0.923)	0.881 (0.871 - 0.892)	0.933 (0.925 - 0.941)
Sensitivity (95% CI)	0.893 (0.871 - 0.912)	0.889 (0.843 - 0.925)	0.934 (0.917 - 0.948)	0.919 (0.892 - 0.941)	0.874 (0.850 - 0.895)	0.876 (0.809 - 0.926)
Specificity (95% CI)	0.893 (0.881 - 0.904)	0.943 (0.935 - 0.950)	0.862 (0.849 - 0.875)	0.914 (0.903 - 0.923)	0.884 (0.872 - 0.895)	0.936 (0.927 - 0.943)
Precision (95% CI)	0.723 (0.696 - 0.749)	0.526 (0.477 - 0.574)	0.731 (0.707 - 0.755)	0.627 (0.591 - 0.661)	0.699 (0.672 - 0.726)	0.338 (0.289 - 0.390)
Negative predictive value (95% CI)	0.964 (0.956 - 0.970)	0.992 (0.988 - 0.994)	0.970 (0.962 - 0.976)	0.986 (0.981 - 0.990)	0.958 (0.949 - 0.965)	0.995 (0.992 - 0.997)
Kappa[§]	0.727	0.630	0.735	0.696	0.698	0.460
F_1[†]	0.799	0.661	0.820	0.745	0.777	0.488
AUROC (95% CI)	0.957 (0.951 - 0.964)	0.965 (0.954 - 0.976)	0.959 (0.954 - 0.965)	0.971 (0.965 - 0.977)	0.947 (0.939 - 0.955)	0.963 (0.947 - 0.978)

Abbreviations: NMGO: normal with meibomian gland opening; MGA: meibomian gland atrophy; MGAO: meibomian gland atrophy with obstruction; MGO: meibomian gland obstruction; MGOO: meibomian gland obstruction with opening;AUROC: Area Under the Receiver Operating Characteristic curve

Table 3 Performance metrics of the ensemble DCNN model evaluated on the external validation set

Performance metrics	External test set (n=12289)					
	Normal group (n=2912)	NMGO group (n=838)	MGA group (n=3493)	MGAO group (n=1693)	MGO group (n=2866)	MGOO group (n=487)
Accuracy (95% CI)	0.894 (0.889 - 0.900)	0.929 (0.925 - 0.934)	0.885 (0.879 - 0.890)	0.916 (0.911 - 0.921)	0.881 (0.875 - 0.886)	0.939 (0.934 - 0.943)
Sensitivity (95% CI)	0.901 (0.890 - 0.912)	0.906 (0.884 - 0.925)	0.940 (0.931 - 0.948)	0.926 (0.912 - 0.938)	0.893 (0.881 - 0.904)	0.825 (0.789 - 0.858)
Specificity (95% CI)	0.892 (0.886 - 0.898)	0.931 (0.926 - 0.936)	0.863 (0.856 - 0.870)	0.914 (0.909 - 0.920)	0.877 (0.870 - 0.884)	0.943 (0.939 - 0.947)
Precision (95% CI)	0.722 (0.707 - 0.736)	0.490 (0.465 - 0.515)	0.732 (0.718 - 0.744)	0.633 (0.614 - 0.652)	0.689 (0.673 - 0.703)	0.375 (0.346 - 0.405)
Negative predictive value (95% CI)	0.967 (0.963 - 0.970)	0.993 (0.991 - 0.994)	0.973 (0.969 - 0.977)	0.987 (0.985 - 0.989)	0.964 (0.960 - 0.968)	0.992 (0.991 - 0.994)
Kappa[§]	0.731	0.601	0.739	0.703	0.698	0.488
F_1[†]	0.802	0.636	0.823	0.752	0.777	0.516
AUROC (95% CI)	0.960 (0.957 - 0.963)	0.971 (0.966 - 0.976)	0.961 (0.958 - 0.964)	0.974 (0.971 - 0.977)	0.954 (0.950 - 0.958)	0.951 (0.941 - 0.961)

Abbreviations: NMGO: normal with meibomian gland opening; MGA: meibomian gland atrophy; MGAO: meibomian gland atrophy with obstruction; MGO: meibomian gland obstruction; MGOO: meibomian gland obstruction with opening; AUROC: Area Under the Receiver Operating Characteristic curve

Table 4 Application of deep learning in ophthalmic diseases

Author	Year	Diseases	Deep learning algorithms	Refs.
Li *et al*	2021	Keratitis	DenseNet121, Inception-v3, ResNet-50	29
Zhang *et al*	2021	Chronic kidney disease and diabetic fundus disease	ResNet-50	30
Yang *et al*	2020	Glaucomatous Optic Neuropathy	ResNet-50	31
Khan *et al*	2021	Meibomian gland dysfunction	ResNet-v2	32
Daisuke *et al*	2019	Branch Retinal Vein Occlusion	Visual Geometry Group (VGG)-16	33

Declarations

Acknowledgements

Not applicable.

Foundation item：National Natural Science Foundation No: 82160195； Central Government Guides Local Science and Technology Development Foundation(No: 20211ZDG02003); Key Research Foundation of Jiangxi Province (No: 20181BBG70004, 20203BBG73059); Excellent Talents Development Project of Jiangxi Province No: 20192BCBL23020；Natural Science Foundation of Jiangxi Province No: 20181BAB205034；Grassroots Health Appropriate Technology “Spark Promotion Plan” Project of Jiangxi Province(No:20188003); Health Development Planning Commission Science Foundation of Jiangxi Province (No: 20201032,202130210)

Availability of data and materials

The datasets used and/or analyzed during the present study are available from the corresponding author on reasonable request.

Ethical approval and consent to participate

The study methods and protocols were approved by the Medical Ethics Committee of the First Affiliated Hospital of Nanchang University (Nanchang, China) and followed the principles of the Declaration of Helsinki. All subjects were notified of the objectives and content of the study and latent risks, and then provided written informed consent to participate.

Patient consent for publication

Not applicable.

Competing interests

This study did not receive any industrial support. The authors have no competing interests to declare regarding this study

Author contributions All the authors contributed to this manuscript. Yi Shao and Yi-Chen Yang are responsible for conceiving and designing the work, acquiring data and writing the manuscript; Hui Zhao and Wen-Qing Shi played an important role in interpreting the results; Xu-Lin Liao, Ting Su and Rong-Bin Liang Min helped in acquiring data and giving some advice; Qiu-Yu Li and Qian-Min Ge contributed to the application to the ethics committee; Hui-Ye Su and Yi-Cong Pan revised the manuscript, and Xiang-Chun Li gave valuable guidance at every stage and approved the final version.

Figures

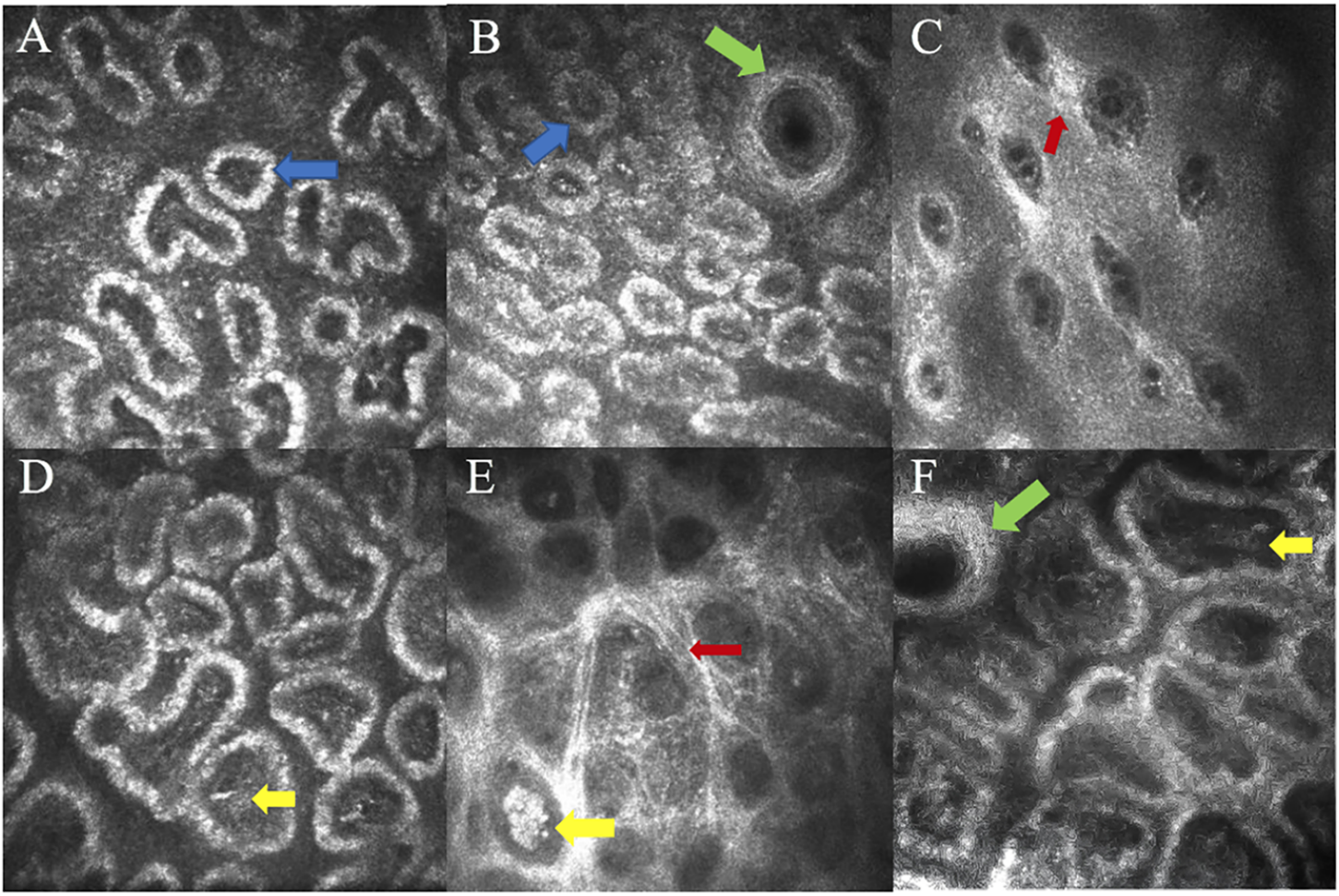

Figure 1

Representative MG images obtained under in vivo confocal microscope. Notes: A:Normal group;B:Normal with meibomian gland opening group .C:Meibomian gland atrophy group.D:Meibomian gland obstruction group.E:Meibomian gland atrophy with obstruction group .F:Meibomian gland obstruction with opening group.Blue arrow: Normal acinar; Green arrow: Meibomian gland opening; Red arrow: Atrophy of acinar wall; Yellow arrow: Acinar containing blocked lipid.

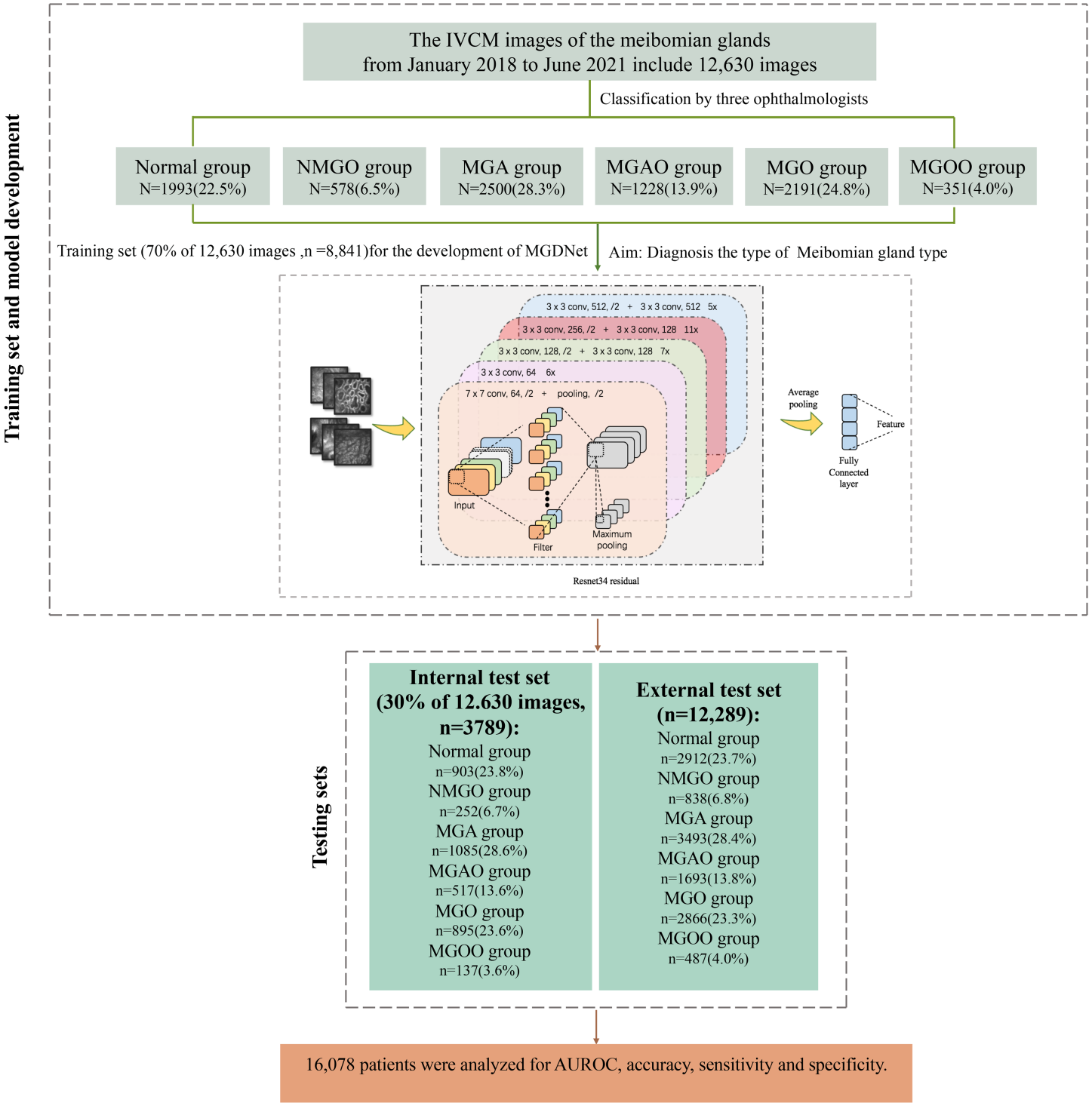

Figure 2

Flow chart of deep learning model development and internal validation process for automatic diagnosis of meibomian gland dysfunction. NMGO: Normal with meibomian gland opening; MGA: Meibomian gland atrophy; MGO: Meibomian gland obstruction; MGAO: Meibomian gland atrophy with obstruction; MGOO: Meibomian gland obstruction with opening; MGD: Meibomian gland dysfunction; AUROC: Area under the receiver operating characteristic curve.

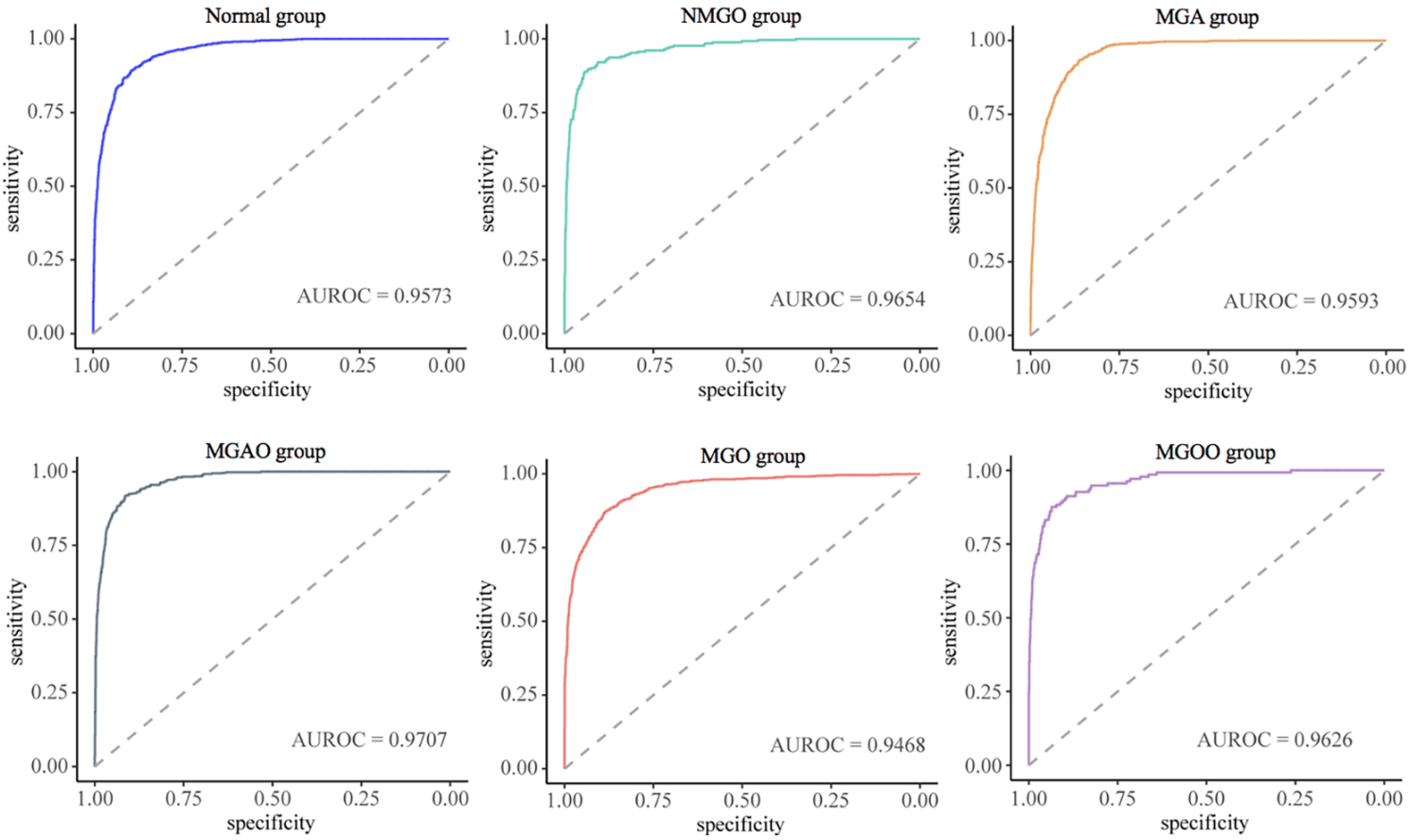

Figure 3

The ROC curve of the six classification internal validation sets. NMGO: Normal with meibomian gland opening; MGA: Meibomian gland atrophy; MGO: Meibomian gland obstruction; MGAO: Meibomian gland atrophy with obstruction; MGOO: Meibomian gland obstruction with opening; AUROC: Area under the receiver operating characteristic curve.

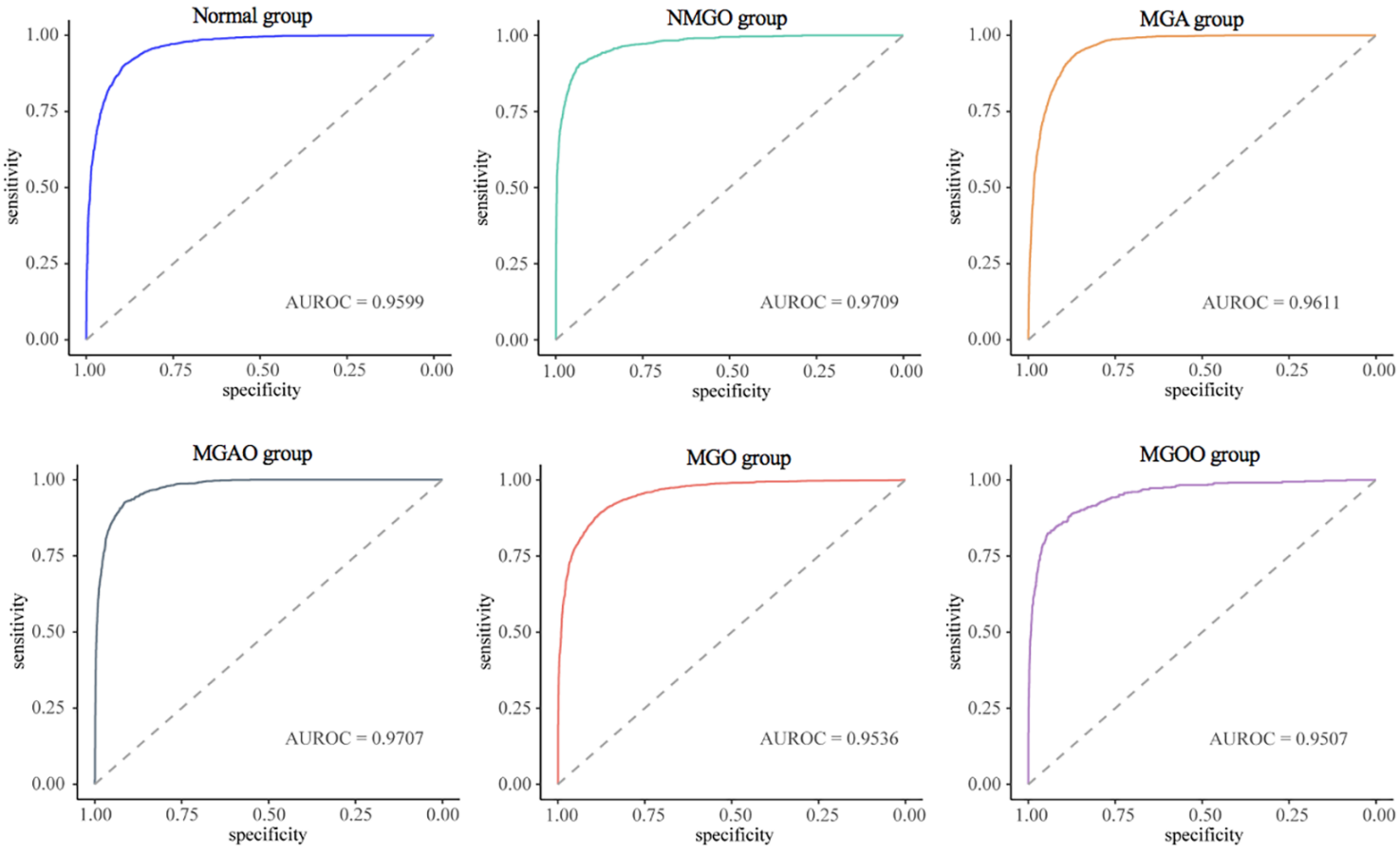

Figure 4

The ROC curve of the six classification external validation sets. NMGO: Normal with meibomian gland opening; MGA: Meibomian gland atrophy; MGO: Meibomian gland obstruction; MGAO: Meibomian gland atrophy with obstruction; MGOO: Meibomian gland obstruction with opening; AUROC: Area under the receiver operating characteristic curve.