An Outperforming Artificial Intelligence Model to Identify Referable Blepharoptosis for General Practitioners

Ju-Yi Hung, Ke-Wei Chen, Chandrashan Perera, Hsu-Kuang Chiu, Cherng-Ru Hsu, David Myung, An-Chun Luo, Chiou-Shann Fuh, Shu-Lang Liao, Andrea Lora Kossler

Submitted to: Journal of Medical Internet Research
on: October 24, 2021

Table of Contents

An Outperforming Artificial Intelligence Model to Identify Referable Blepharoptosis for General Practitioners

Ju-Yi Hung[1, 2, 3*] MD; Ke-Wei Chen[4*] MD, MS; Chandrashan Perera[2] MD; Hsu-Kuang Chiu[5] MS; Cherng-Ru Hsu[6] MD; David Myung[2] MD; An-Chun Luo[7] PhD; Chiou-Shann Fuh[1] PhD; Shu-Lang Liao[8, 9] MD; Andrea Lora Kossler[2] MD

[1]Computer Science and Information Engineering, National Taiwan University Taipei TW
[2]Ophthalmology, Byers Eye Institute, Stanford University School of Medicine Palo Alto US
[3]Ophthalmology, Taipei Medical University Hospital Taipei TW
[4]Biomedical Engineering, National Cheng Kung University Tainan TW
[5]Computer Science, Stanford University Stanford US
[6]Ophthalmology, Tri-Service General Hospital, National Defense Medical Center Taipei TW
[7]Industrial Technology Research Institute Hsinchu TW
[8]College of Medicine, National Taiwan University Taipei TW
[9]Ophthalmology, National Taiwan University Hospital, Taipei, Taiwan Taipei TW
[*]these authors contributed equally

Corresponding Author:
Shu-Lang Liao MD
Ophthalmology, National Taiwan University Hospital, Taipei, Taiwan
No. 7, Chung Shan S. Rd., (Zhongshan S. Rd.), Zongzheng Dist
Taipei
TW

Abstract

Background: Accurate identification and prompt referral for blepharoptosis can be challenging for general practitioners. An artificial intelligence-aided diagnostic tool could underpin decision-making.

Objective: To develop an AI model which accurately identifies referable blepharoptosis automatically and to compare the AI model's performance to a group of non-ophthalmic physicians.

Methods: Retrospective 1,000 single-eye images from tertiary oculoplastic clinics were labeled by three oculoplastic surgeons with ptosis, including true and pseudoptosis, versus healthy eyelid. The VGG (Visual Geometry Group)-16 model was trained for binary classification. The same dataset was used in testing three non-ophthalmic physicians. The Gradient-weighted Class Activation Mapping (Grad-CAM) was applied to visualize the AI model

Results: The VGG16-based AI model achieved a sensitivity of 92% and a specificity of 88%, compared with the non-ophthalmic physician group, who achieved a mean sensitivity of 72% [Range: 68% - 76%] and a mean specificity of 82.67% [Range: 72% - 88%]. The area under the curve (AUC) of the AI model was 0.987. The Grad-CAM results for ptosis predictions highlighted the area between the upper eyelid margin and central corneal light reflex.

Conclusions: The AI model shows better performance than the non-ophthalmic physician group in identifying referable blepharoptosis, including true and pseudoptosis, correctly. Therefore, artificial intelligence-aided tools have the potential to assist in the diagnosis and referral of blepharoptosis for general practitioners.

(JMIR Preprints 24/10/2021:34445)
DOI: https://doi.org/10.2196/preprints.34445

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?
✓ **Please make my preprint PDF available to anyone at any time (recommended).**
Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.

No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**

Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v

Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Original Manuscript

Original Study

An Outperforming Artificial Intelligence Model to Identify Referable Blepharoptosis for General Practitioners

Authors

Ju-Yi Hung[a, b, c, 1], Ke-Wei Chen[a, d, 1], Chandrashan Perera[a], Hsu-Kuang Chiu[e], Cherng-Ru Hsu [f], David Myung[a], An-Chun Luo[g], Chiou-Shann Fuh[b], Shu-Lang Liao[h,i, *], Andrea Lora Kossler [a, *]

Affiliations

[a] Ophthalmology, Byers Eye Institute, Stanford University School of Medicine, Palo Alto, California, United States

[b] Computer Science and Information Engineering, National Taiwan University, Taipei, Taiwan

[c] Ophthalmology, Taipei Medical University Hospital, Taipei, Taiwan

[d] Biomedical Engineering, National Cheng Kung University, Tainan, Taiwan

[e] Computer Science, Stanford University, Stanford, California, United States

[f] Ophthalmology, Tri-Service General Hospital, National Defense Medical Center, Taipei, Taiwan

[g] Industrial Technology Research Institute, Hsinchu, Taiwan

[h] Ophthalmology, National Taiwan University Hospital, Taipei, Taiwan

[i] College of Medicine, National Taiwan University, Taipei, Taiwan

[1]Two co-first authors contributed equally.

*Two co-corresponding authors.

Correspondence to:

1) Shu-Lang Liao, MD
   Department of Ophthalmology, National Taiwan University Hospital
   No. 7, Chung Shan South Road, Taipei City 100, Taiwan
   Email: liaosl89@ntu.edu.tw

2) Andrea Lora Kossler, MD
   Byers Eye Institute, Stanford University School of Medicine

2452 Watson Ct, Palo Alto, California 94303, USA
Email: akossler@stanford.edu

Abstract

Background

Accurate identification and prompt referral for blepharoptosis can be challenging for general practitioners. An artificial intelligence-aided diagnostic tool could underpin decision-making.

Purpose

To develop an AI model which accurately identifies referable blepharoptosis automatically and to compare the AI model's performance to a group of non-ophthalmic physicians.

Design

Original Study

Methods

Retrospective 1,000 single-eye images from tertiary oculoplastic clinics were labeled by three oculoplastic surgeons with ptosis, including true and pseudoptosis, versus healthy eyelid. The VGG (Visual Geometry Group)-16 model was trained for binary classification. The same dataset was used in testing three non-ophthalmic physicians. The Gradient-weighted Class Activation Mapping (Grad-CAM) was applied to visualize the AI model

Results

The VGG16-based AI model achieved a sensitivity of 92% and a specificity of 88%, compared with the non-ophthalmic physician group, who achieved a mean sensitivity of 72% [Range: 68% - 76%] and a mean specificity of 82.67% [Range: 72% - 88%]. The area under the curve (AUC) of the AI model was 0.987. The Grad-CAM results for ptosis predictions highlighted the area between the upper eyelid margin and central corneal light reflex.

Conclusion

The AI model shows better performance than the non-ophthalmic physician group in identifying referable blepharoptosis, including true and pseudoptosis, correctly. Therefore, artificial intelligence-aided tools have the potential to assist in the diagnosis and referral of blepharoptosis for general practitioners.

Keywords

artificial intelligence; blepharoptosis; general practitioners; computer-aided diagnosis (CAD)

Introduction

Blepharoptosis, also known as ptosis, is the drooping or inferior displacement of the upper eyelid. Ptosis can obstruct the visual axis and affect vision and can be a presenting sign of a serious medical disorder, such as ocular myasthenia [1], third cranial nerve palsy [2], or Horner syndrome [3]. It is important for general practitioners to accurately diagnosis ptosis to assist in decision-making for referral and work up when necessary. Ptosis is diagnosed by using a ruler and light source to measure the distance between the pupillary light reflex and the upper eyelid margin (margin reflex distance 1, or MRD1) with the eyes in the primary position [4]. With low repeatability and reproducibility in measuring the eyelid landmarks and the effect of learning curve [5, 6], accurately recognizing ptosis is challenging especially for non-ophthalmologists. Therefore, an automated tool for ptosis diagnosis may be useful for general practitioners.

Currently, artificial intelligence (AI) -aided diagnostic tools play a promising role in automatic detection of certain diseases, such as diabetic retinopathy [7] and skin cancer [8] from retinal fundus and skin images, respectively. Convolutional neural networks (CNNs)-based deep learning methods, a subset of machine learning techniques, has been the state of the art in artificial intelligence for years, leading to enhanced performance in various medical applications [9]. It requires less supervision and uses an end-to-end learning mechanism to map raw inputs, such as image pixels, to outputs without human-directed manipulation of the data [10]. The image-to-classification approach in one classifier replaces the multiple steps of previous image analysis methods [11].

In a previous study [12], a variety of CNN (convolutional neural network) architectures, such as VGG-16 [13], ResNet [14], DenseNet [15], diagnosed true ptosis without any inputs of eyelid measurements from a clinical facial photograph, achieving high accuracy from 83.3% to 88.6%. In this study, we further trained an AI model using VGG-16 architecture with larger and more diverse data to accurately diagnose blepharoptosis and compared the AI model's performance to a group of non-ophthalmic physicians. Our goal was to determine if our AI model could outperform physicians to support the need for an AI tool to diagnose blepharoptosis.

Methos

Image preparation

Original photographs, taken by hand-held digital camera (Canon DIGITAL IXUS 950 IS) at a tertiary oculoplastic clinic, from adult patients over 20-year-old were retrospectively collected in past 20 years for surgical evaluation. A total of 1,000 images were used in this study. IRB approval was granted for this study by Stanford University, and the research was conducted in accordance with National Taiwan University IRB protocol.

In order to crop a standardized image of a single eye, OpenFace [16], an open-source package, was utilized to identify major facial landmarks in each photograph. The cropped single-eye images were 400x600 pixels individually and were then resized to 200x300 pixels, matching the input size which was ready to be used in the CNN architectures.

Inclusion and Exclusion

The photographs involved only the periocular region of both eyes **[Figure 1].** The condition of healthy eyelid is illustrated in **Figure 1a**. The referable ptosis group included mild ptosis, severe ptosis, and pseudoptosis (dermatochalasis) a condition in which excess upper eyelid skin overhangs the eyelid margin **[Figure 1b]**. Upper eyelid retraction was excluded **[Figure 1c].** Poor quality images, including uncentered visual fixation, uneven curves of the upper eyelids, blurred upper eyelid margin due to dense eyelashes were excluded. A total of 1,000 images were evaluated, 218 images were removed, leaving 782 images used in this study.

The brow region was not included in the photographs; therefore, brow ptosis was not excluded. Exact measurements, such as margin to reflex distance 1 (MRD-1), MRD-2 [17, 18], levator function [19], or palpebral aperture [20] were not provided. The condition of the fellow eye and the history of the patients were withheld.

Annotations for the ground truth

Two labelers, both oculoplastic surgeons, achieved 82% consensus rate in discussion meetings. The major reasons for the disagreements were decisions between healthy eyelids and mild ptosis. To lessen the spectrum bias, a third senior oculoplastic surgeon, as an arbiter, yielded the decisional answer for these disagreements, which included 24 images. **Figure 2** shows the voting system, with 593 images (account for 75%) in referable ptosis group and 189 (25%) images in healthy group.

Data allocation for training, validation and testing

A total of 50 images, including 25 healthy eyelids and 25 ptotic eyelids were randomly selected into testing datasets. The same testing datasets were used in testing the AI model and the physician group. The rest of the photographs were then divided into training and validation datasets with the ratio of 8:2 **[Table 1]**.

Table 1. The number of images in the training, validation and testing sets.

	Training	Validation (for Training)	Testing
Referable ptosis group	455	113	25
Healthy group	132	32	25

Model Architecture and Training

VGG16 was used as the base structure. [21, 22]. We replace VGG16's last few layers by a global max pooling layer followed by fully connected layers and a sigmoid function for our binary classification problem. In order to reduce the memory usage, the size of the input images was adjusted to 200x300 pixels. The details of our model architecture can be seen in **Table 2**.

Table 2. Structure of the model

Input Size	Layer	Output Size	Number of Feature Maps	Kernel Size	Stride	Activation
-	Image	200x300x3	-	-	-	-
200x300x3	Convolution	200x300x64	64	3x3	1	ReLU
200x300x64	Convolution	200x300x64	64	3x3	1	ReLU
200x300x64	Max pooling	100x150x64	64	-	2	-
100x150x64	Convolution	100x150x128	128	3x3	1	ReLU
100x150x128	Convolution	100x150x128	128	3x3	1	ReLU
100x150x128	Max pooling	50x75x128	128	-	2	-
50x75x128	Convolution	50x75x256	256	3x3	1	ReLU
50x75x256	Convolution	50x75x256	256	3x3	1	ReLU
50x75x256	Global max pooling	1x256	-	-	-	-
1x256	Fully connected	1x512	-	-	-	ReLU
1x512	Fully connected	1	-	-	-	Sigmoid

Python module scikit-learn was used to compute the ROC curve and the confusion matrix.

Transfer Learning and Data Augmentation

Transfer learning was performed by importing weights trained on ImageNet [23]. Tensorflow 2.0 with Keras is used as our training framework. For the learning rate optimization, Adam optimizer was applied [24]. Data augmentation was also used to prevent overfitting. The transformation of photograph included:

- Images flipped horizontally
- Random image rotation up to 15 degrees
- Random zoom in or out with the range of 90% to 120%
- Adjusted brightness/contrast by 50%
- Images shifted horizontally or vertically by 10%

Testing in non-ophthalmic physician group

Three specialists, one from emergency medicine, neurology, and family medicine were tested, on behalf of the non-ophthalmic physician group. The clinical experience of each of the three physicians was over five years. The same testing set, including 25 healthy eyelids and 25 ptotic eyelids, was given to the group for distinguishing ptotic eyelids from healthy eyelids. No other information, such as MRD-1 measurements, the condition of the fellow eye, or patient histories, were provided. Moreover, no further training on blepharoptosis diagnosis was given. The decision making relied on each physician's personal background knowledge.

Visualize the CNN model - Grad-CAM.

Gradient-weighted Class Activation Mapping (Grad-CAM) [25] was applied to visualize the AI model. There are generally two steps in Grad-CAM. First, we calculated the neuron importance weights α_k^c:

$$\alpha_k^c = \overbrace{\frac{1}{Z}\sum_i \sum_j}^{\text{Sum up all pixel elements and do averaging}} \underbrace{\frac{\partial y^c}{\partial A_{ij}^k}}_{\text{Gradient from single pixel of feature map}} \quad \textbf{(1)}$$

where c represented the class. In our study, there was only one class. Thus, there was only one number for c. A^k represented the k-th feature map. In our study, there were 256 feature maps in the last convolutional layer. i and j represented the pixel indices in the width and height dimensions of the feature map. In our study, the number were between 1 to 50 and 1 to 75 respectively. Z was the total number of pixels in the feature map, which was 50 times 75 in our study. $\frac{\partial y^c}{\partial A_{ij}^k}$ represented the gradient with respect to the single pixel *(i, j)* of the feature map A^k.

Second, we performed linear combination of all feature maps, and each feature map had a corresponding neuron importance weight. The ReLU function was also applied to filter out pixels with negative gradients. The result was then a matrix $L_{Grad-CAM}^c \in R^{50 \times 75}$

$$L_{Grad-CAM}^c = ReLU\left(\sum_k \alpha_k^c A^k\right) \quad \textbf{(2)}$$

The result was then normalized to have values between 0 to 1 to show the relative importance of each location.

Results

There were 45 correct predictions, including 22 healthy and 23 ptosis answers, by the CNN model from a total of 50 testing images. The accuracy of the AI model was 90% with a sensitivity of 92% and a specificity of 88%. Three false positives and two false negatives were found.

The confusion matrix with 0.5 threshold setting is shown in **Figure 3**. The Receiver Operating Characteristic curve (ROC curve) is presented in **Figure 4**. The Area Under the Curve (AUC) is 0.987. The mean accuracy of the non-ophthalmic physician group is 77.33% [Range: 70%-82%] with a mean sensitivity of 72% [Range: 68%-76%] and a mean specificity of 82.67% [Range: 72%-88%] **[Figure 5].**

Grads-CAM results

The result of Grads-CAM showed that the weight in the background is around 0~0.2. In the ptotic eyelids, the area between upper eyelid margin and central cornea light reflex showed the highest weight (around 0.5~1.0). **[Figure 6]**

Discussion

It is important for general practitioners to promptly refer eyelid ptosis, including pseudoptosis, to ophthalmologic specialists for further evaluation, work up and treatment. Pseudoptosis is a heterogeneous group of disorders where the upper eyelid can droop in the absence of pathology of the upper eyelid muscles [26]. Dermatochalasis is likely the most common eyelid condition that causes confusion when evaluating the patient with apparent ptosis. Excess upper eyelid skin may overhang the eyelashes and obstruct visualization of the eyelid margin, giving the impression of a low-lying eyelid **[Figure 2].** In a previous, proof of concept study, we demonstrated that an AI model could detect true ptosis from healthy eyelids [12]. In this study, we evaluated true ptosis and pseudoptosis versus health eyelids, applied a larger dataset of 782 images and compared the AI model performance to physicians. Our results demonstrate that the AI model achieved an accuracy of 90% with 92% specificity and 88% sensitivity. Additionally, the AI model performed well even with including pseudoptosis cases, the condition that could better mimic the real clinical situation in primary care.

A non-ophthalmic group of 3 physicians, including experts in family medicine, neurology, and emergency medicine, were chosen as a comparator group. The family medicine doctor represented general practitioners who are commonly the first line in seeing and diagnosing age related and systemic causes of ptosis. The neurologist was selected due to specialized training in diagnosing ptosis, particularly related to neurologic or myogenic causes. Finally, the emergency medicine doctor was selected due to expertise in diagnosing acute causes of ptosis such as Horner syndrome, third nerve palsy [27] or trauma. Hence, in our non-ophthalmic group, the expertise of three physicians from family medicine, neurology, and emergency medicine was claimed to be better at identifying blepharoptosis than general practitioners. The result shows that the mean accuracy in the non-ophthalmic physician group was 77.33% [Range: 70% - 82%] with a mean sensitivity of 72% [Range: 68% - 76%] and a mean specificity of 82.67% [Range: 72% - 88%]. While the AI model achieved an accuracy of 90% with a sensitivity of 92% and a specificity of 88%. These results suggest that an AI-aided diagnostic tool may help in detection of blepharoptosis to assist decision-making of referral for general practitioners.

CNNs (Convolutional neural networks) have achieved great success for image classification. For example, the current largest image classification dataset classification challenge ImageNet, all models with top performance used CNN architectures. The general trend is that the deeper the model, the higher discernment you can get. Some of the model structures can be very deep like RestNet-152, which has 152 CNN layers. In a previous smaller scale study [12], where less than 500 eyelid pictures were exploited, a variety of CNN models had showed high performance over an accuracy of 80%. This study shows that for ptosis classification, most common model such as VGG, ResNet, AlexNet, SqueezeNet and DenseNet all have similar performance. Therefore, among those models, we choose a relatively simple model VGG-16, a computing resource-efficient model, as our base model.

Gradient-weighted Class Activation Mapping (Grad-CAM) is a visual explanation of the AI model, which is applicable to a wide variety of CNN model-families [25]. To aid understanding of AI model predictions, a heat map that identifies the areas of the input image that contributed most to the AI model's classification using a technique called class activation mappings is produced. In addition, to visualize the reasonable AI predictions, Grad-CAM explanations also helped identify dataset biases in images. For example, a preoperative marking around the eye or a postoperative suture on the eyelid may provide misleading clues to the AI model, rather than the eyelid information about blepharoptosis. The results of Grad-CAM [**Figure 6**] demonstrated a hotspot area (0.5-1.0 in weights) between upper eyelid margin and central corneal light reflex, which is compatible with the MRD-1 concept clinically. The cold zone (0-0.2 in weights) in the background successfully excludes the dataset biases, providing stronger faithfulness. With larger and more diverse data utilization in the future, more precise results to understand the AI predications can be expected.

Limitations to this study include that the data resource was only from Asian ethnicities, setting limitation in both model training and testing process. Future studies will analyze external photographs from diverse ethnicities to further train the AI model and expand the application for all users. Additionally, only adults were included in this study, setting limitation for pediatric care. Future applications of AI assisted ptosis diagnostic tools will focus on congenital ptosis, since up to one-third of congenital ptosis patients are at risk for amblyopia [28]. Accurate diagnosis of ptosis based on external photographs would prove especially helpful in the pediatric population, for ophthalmologists and general practitioners alike, as they eyelid exam can be challenging in uncooperative or crying children, patients with developmental delay and babies. AI assisted detection of congenital ptosis could have a huge impact on preventing and treating amblyopia promptly. External validation with outsourced images, including mobile phone photographs, to confirm the strength and weakness of this AI model, also deserves further investigation.

Conclusion

The AI model using convolutional neural networks achieves better performance than the non-ophthalmic physician group and shows potential value to be used in assisting referral of blepharoptosis, including true and pseudoptosis.

Acknowledgements

We thank Taiwan National Center for High-performance Computing (NCHC) for providing computational and storage resources, departmental core grants from the National Eye Institute (P30 EY026877) and Research to Prevent Blindness (RPB) to the Byers Eye Institute at Stanford, and Karen Chang for assisting in the research.

References

1. Al-Haidar M, Benatar M, Kaminski HJ. Ocular Myasthenia. Neurol Clin. 2018 May;36(2):241-51. PMID: 29655447. doi: 10.1016/j.ncl.2018.01.003.
2. Bagheri A, Borhani M, Salehirad S, Yazdani S, Tavakoli M. Blepharoptosis Associated With Third Cranial Nerve Palsy. Ophthalmic Plast Reconstr Surg. 2015 Sep-Oct;31(5):357-60. PMID: 25376747. doi: 10.1097/iop.0000000000000336.

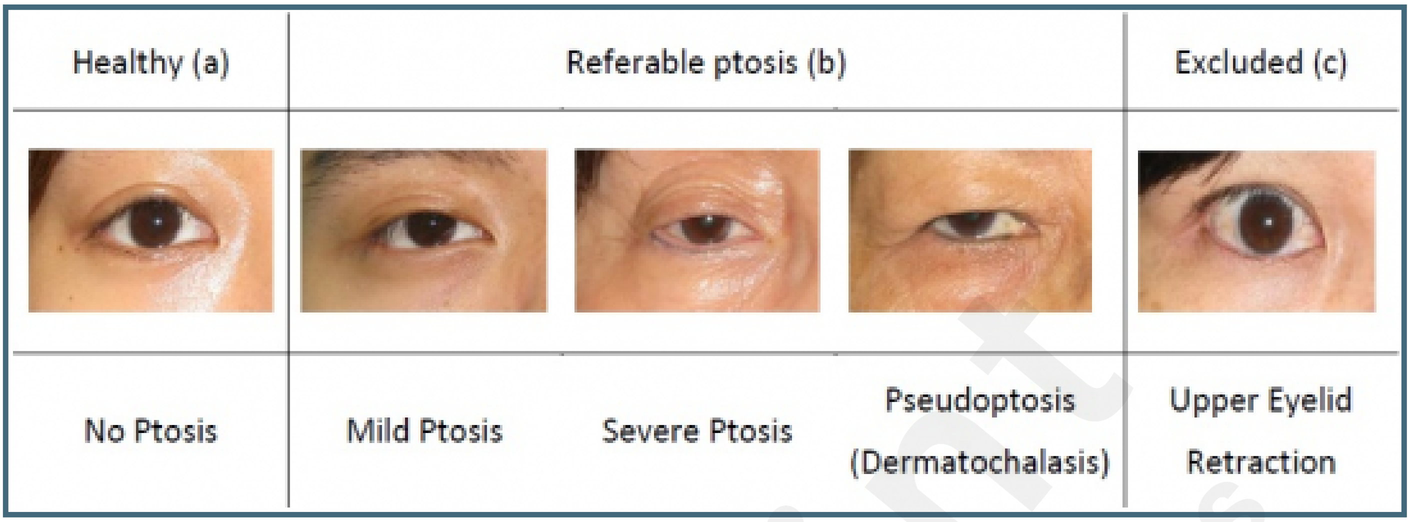
3. Martin TJ. Horner Syndrome: A Clinical Review. ACS Chem Neurosci. 2018 Feb 21;9(2):177-86. PMID: 29260849. doi: 10.1021/acschemneuro.7b00405.
4. Putterman AM. Margin reflex distance (MRD) 1, 2, and 3. Ophthalmic Plastic & Reconstructive Surgery. 2012;28(4):308-11.
5. Boboridis K, Assi A, Indar A, Bunce C, Tyers A. Repeatability and reproducibility of upper eyelid measurements. British journal of ophthalmology. 2001;85(1):99-101. PMID: 11133723. doi: 10.1136/bjo.85.1.99.
6. Nemet AY. Accuracy of marginal reflex distance measurements in eyelid surgery. Journal of Craniofacial Surgery. 2015;26(7):e569-e71.
7. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, et al. Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. JAMA. 2016;316(22):2402-10. PMID: 27898976. doi: 10.1001/jama.2016.17216.
8. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature. 2017 2017/02/01;542(7639):115-8. PMID: 28117445. doi: 10.1038/nature21056.
9. Shen D, Wu G, Suk H-I. Deep Learning in Medical Image Analysis. Annual Review of Biomedical Engineering. 2017;19(1):221-48. PMID: 28301734. doi: 10.1146/annurev-bioeng-071516-044442.
10. Wang F, Casalino LP, Khullar D. Deep learning in medicine—promise, progress, and challenges. JAMA internal medicine. 2019;179(3):293-4. PMID: 30556825. doi: 10.1016/j.ncl.2018.01.003.
11. Zeiler MD, Fergus R, editors. Visualizing and Understanding Convolutional Networks. 2014; Cham: Springer International Publishing.
12. Hung J-Y, Perera C, Chen K-W, Myung D, Chiu H-K, Fuh C-S, et al. A deep learning approach to identify blepharoptosis by convolutional neural networks. International Journal of Medical Informatics. 2021 2021/04/01/;148:104402. PMID: 33609928. doi: 10.1016/j.ijmedinf.2021.104402.
13. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:14091556. 2014.
14. He K, Zhang X, Ren S, Sun J, editors. Deep residual learning for image recognition. Proceedings of the IEEE conference on computer vision and pattern recognition; 2016.
15. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ, editors. Densely connected convolutional networks. IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2017.
16. Amos B, Ludwiczuk B, Satyanarayanan M. Openface: A general-purpose face recognition library with mobile applications. CMU School of Computer Science. 2016;6(2).
17. Bodnar ZM, Neimkin M, Holds JB. Automated Ptosis Measurements From Facial Photographs. JAMA Ophthalmol. 2016 Feb;134(2):146-50. PMID: 26605967. doi: 10.1001/jamaophthalmol.2015.4614.
18. Lou L, Yang L, Ye X, Zhu Y, Wang S, Sun L, et al. A Novel Approach for Automated Eyelid Measurements in Blepharoptosis Using Digital Image Analysis. Curr Eye Res. 2019 Oct;44(10):1075-9. PMID: 31148484. doi: 10.1080/02713683.2019.1619779.
19. Lai H-T, Weng S-F, Chang C-H, Huang S-H, Lee S-S, Chang K-P, et al. Analysis of Levator Function and Ptosis Severity in Involutional Blepharoptosis. Annals of Plastic Surgery. 2017;78(3):S58-S60. PMID: 28195895. doi: 10.1097/sap.0000000000001007.
20. Thomas PB, Gunasekera CD, Kang S, Baltrusaitis T. An Artificial Intelligence Approach to the Assessment of Abnormal Lid Position. Plastic and Reconstructive Surgery Global Open. 2020;8(10). PMID: 33173665. doi: 10.1097/GOX.0000000000003089.
21. Karen Simonyan AZ. Very Deep Convolutional Networks for Large-Scale Image

Recognition. ICLR2015.
22. Da K. A method for stochastic optimization. arXiv preprint arXiv:14126980. 2014.
23. Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, et al. Imagenet large scale visual recognition challenge. International journal of computer vision. 2015;115(3):211-52.
24. Kingma DaB, Jimmy. Adam: A method for stochastic optimization. arXiv preprint. 2014;arXiv:1412.6980.
25. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D, editors. Grad-cam: Visual explanations from deep networks via gradient-based localization. IEEE International Conference on Computer Vision (ICCV); 2017.
26. Cohen A.J. WDA. Pseudoptosis. Evaluation and Management of Blepharoptosis: Springer, New York, NY; 2011.
27. Huff JS, Austin EW. Neuro-Ophthalmology in Emergency Medicine. Emerg Med Clin North Am. 2016 Nov;34(4):967-86. PMID: 27741997. doi: 10.1016/j.emc.2016.06.016.
28. Zhang JY, Zhu XW, Ding X, Lin M, Li J. Prevalence of amblyopia in congenital blepharoptosis: a systematic review and Meta-analysis. Int J Ophthalmol. 2019;12(7):1187-93. PMID: 31341812. doi: 10.18240/ijo.2019.07.21.

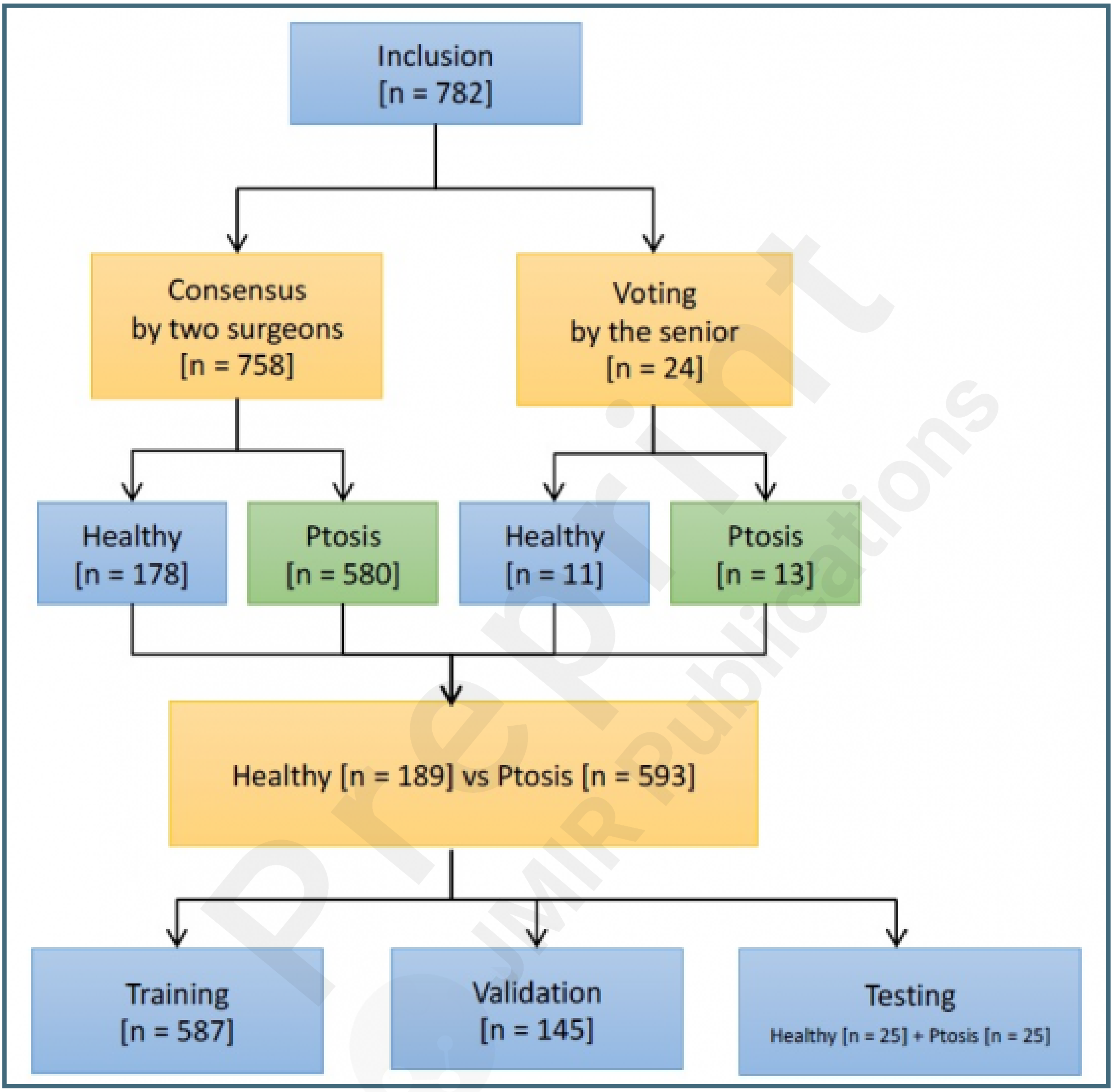
Supplementary Files

Figures

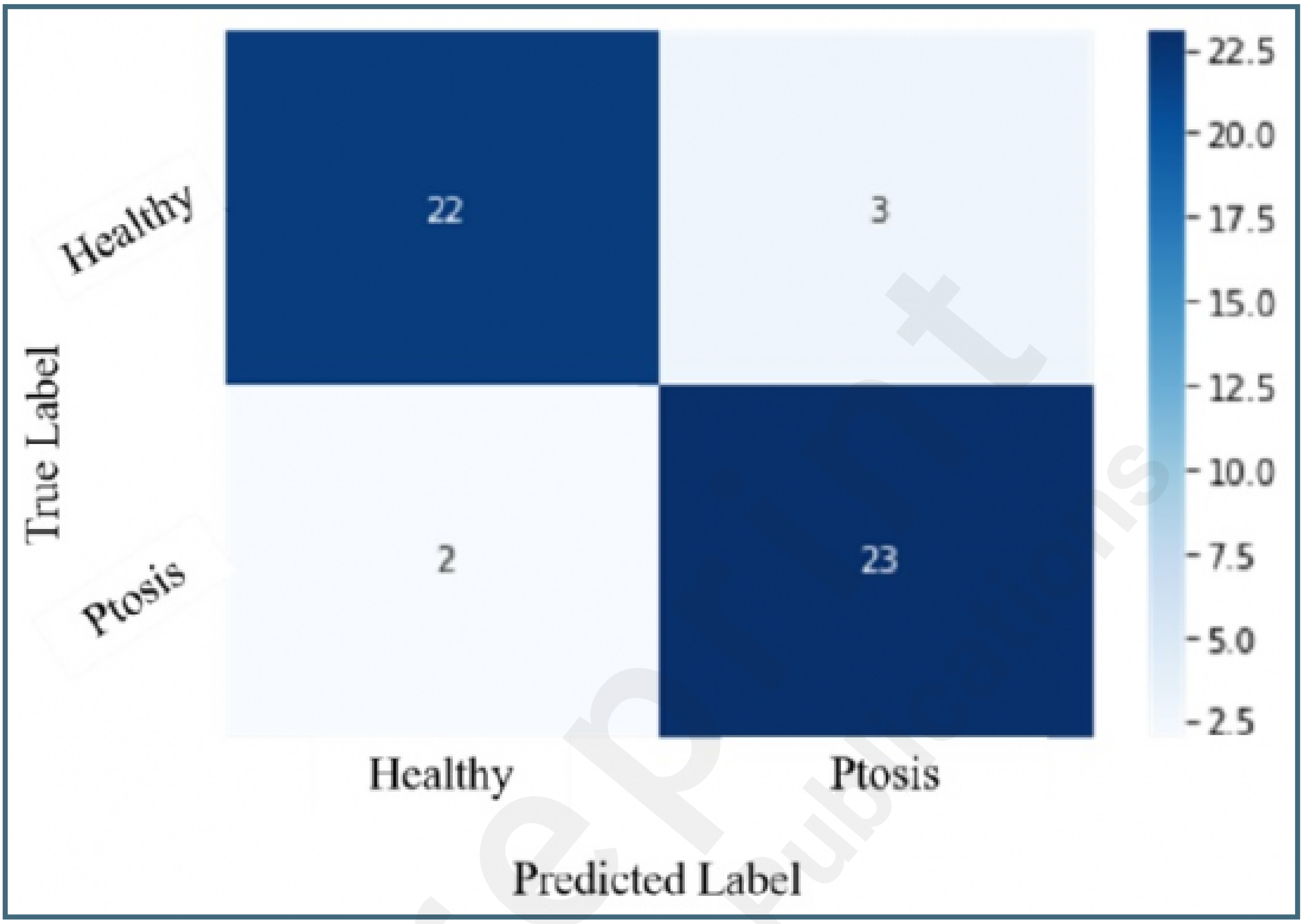
Examples of images used in this study.



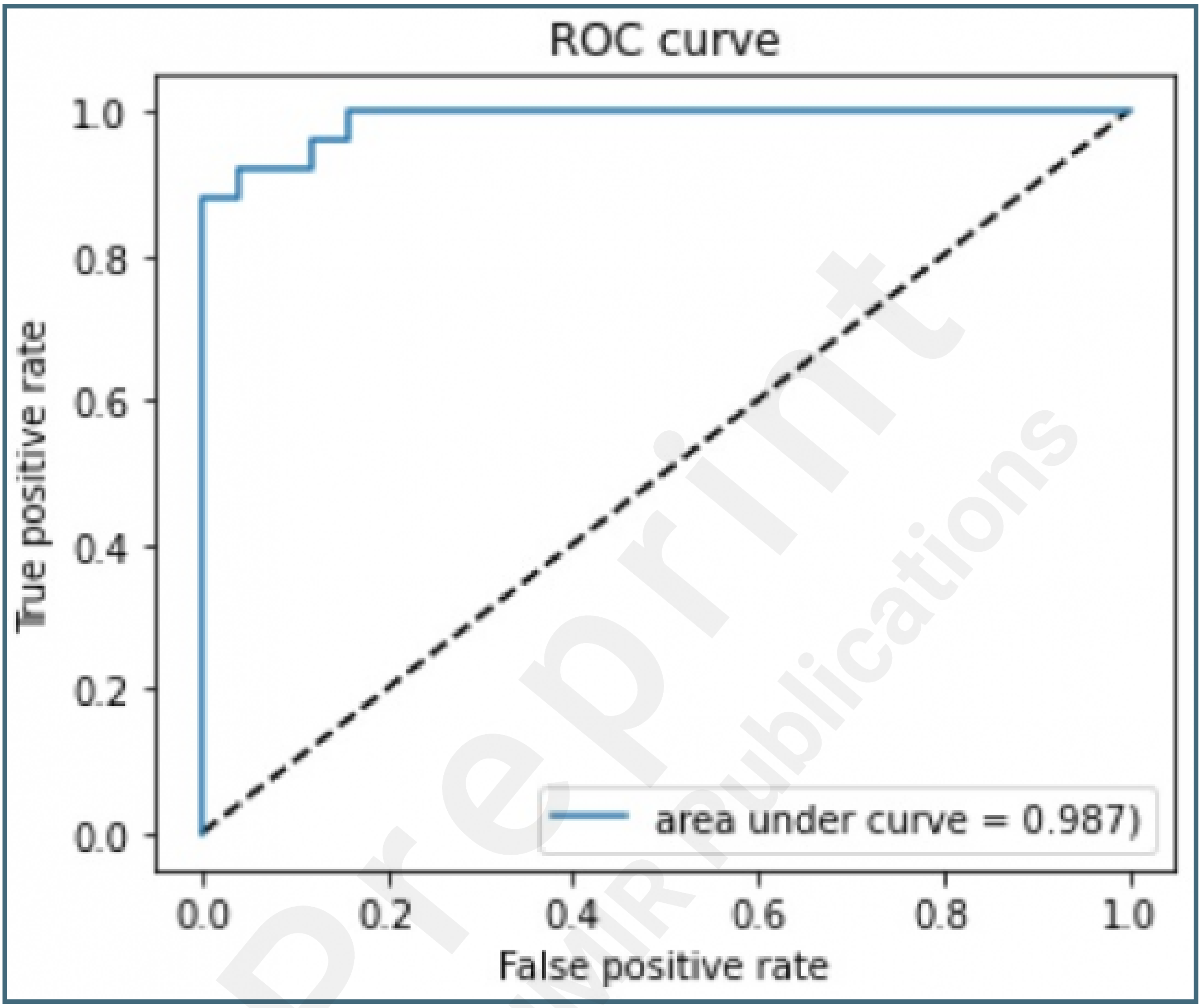
Flowchart for data labeling.



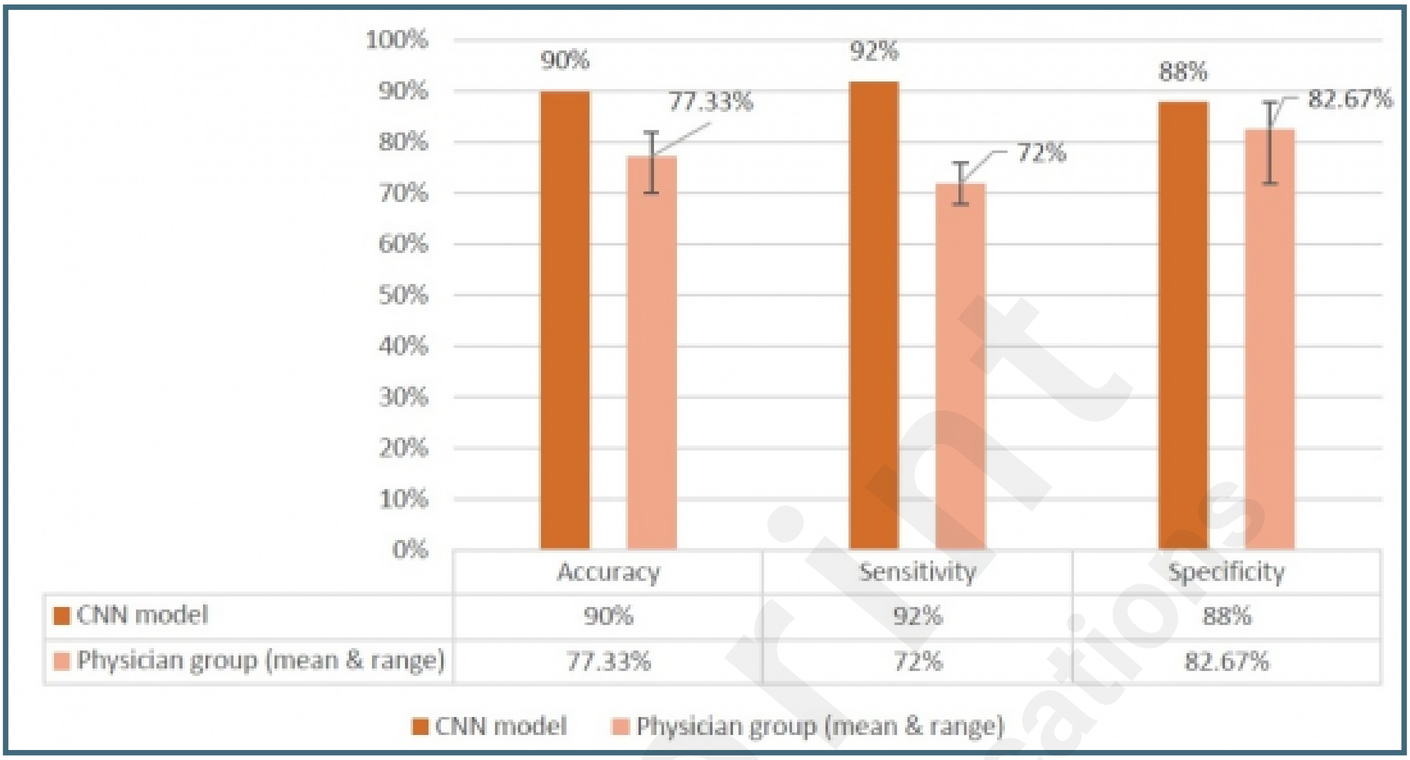
Confusion matrix for model evaluation. The threshold is 0.5.



ROC curve for the model evaluation.



Performance comparison between model and physicians.



Original images and Grads-CAM results of the AI model predictions.