Differentiation of Active Corneal Infections from Healed Scars Using Deep Learning

Mo Tiwari, BS,[1] Chris Piech, PhD,[1] Medina Baitemirova, BA,[2] Namperumalsamy V. Prajna, DO, DNB,[3] Muthiah Srinivasan, DO, DNB,[3] Prajna Lalitha, DO, DNB,[3] Natacha Villegas, MD,[4] Niranjan Balachandar, BS, MS,[1] Janice T. Chua, BS,[5] Travis Redd, MD, MPH,[6] Thomas M. Lietman, MD,[7] Sebastian Thrun, PhD,[1] Charles C. Lin, MD[4]

Purpose: To develop and evaluate an automated, portable algorithm to differentiate active corneal ulcers from healed scars using only external photographs.

Design: A convolutional neural network was trained and tested using photographs of corneal ulcers and scars.

Participants: De-identified photographs of corneal ulcers were obtained from the Steroids for Corneal Ulcers Trial (SCUT), Mycotic Ulcer Treatment Trial (MUTT), and Byers Eye Institute at Stanford University.

Methods: Photographs of corneal ulcers (n = 1313) and scars (n = 1132) from the SCUT and MUTT were used to train a convolutional neural network (CNN). The CNN was tested on 2 different patient populations from eye clinics in India (n = 200) and the Byers Eye Institute at Stanford University (n = 101). Accuracy was evaluated against gold standard clinical classifications. Feature importances for the trained model were visualized using gradient-weighted class activation mapping.

Main Outcome Measures: Accuracy of the CNN was assessed via F_1 score. The area under the receiver operating characteristic (ROC) curve (AUC) was used to measure the precision-recall trade-off.

Results: The CNN correctly classified 115 of 123 active ulcers and 65 of 77 scars in patients with corneal ulcer from India (F_1 score, 92.0% [95% confidence interval (CI), 88.2%−95.8%]; sensitivity, 93.5% [95% CI, 89.1%−97.9%]; specificity, 84.42% [95% CI, 79.42%−89.42%]; ROC: AUC, 0.9731). The CNN correctly classified 43 of 55 active ulcers and 42 of 46 scars in patients with corneal ulcers from Northern California (F_1 score, 84.3% [95% CI, 77.2%−91.4%]; sensitivity, 78.2% [95% CI, 67.3%−89.1%]; specificity, 91.3% [95% CI, 85.8%−96.8%]; ROC: AUC, 0.9474). The CNN visualizations correlated with clinically relevant features such as corneal infiltrate, hypopyon, and conjunctival injection.

Conclusions: The CNN classified corneal ulcers and scars with high accuracy and generalized to patient populations outside of its training data. The CNN focused on clinically relevant features when it made a diagnosis. The CNN demonstrated potential as an inexpensive diagnostic approach that may aid triage in communities with limited access to eye care. *Ophthalmology 2022;129:139-146 © 2021 by the American Academy of Ophthalmology*

Despite advances in our understanding of how to treat corneal infections, corneal blindness remains the fourth most common cause of blindness in the world, accounting for more than 5% of the total blind population.[1] Corneal opacities disproportionately affect developing countries, where they may constitute more than 10% of preventable visual impairment cases. For example, in India, corneal ulcers affect approximately 2 million people every year.[1] In such developing countries, delayed or inappropriate treatment resulting from misdiagnosis is a primary driving factor. Appropriate triage of these patients through accurate diagnosis of a corneal ulcer is the first step toward addressing this problem, but this requires a trained eye care provider, which often is a limited resource in these areas. India, for example, has an estimated 25 000 ophthalmologists for its 1.3 billion residents.[2] This ratio of 1:52 000 is well below typical public health guidelines.[3] In the absence of trained eye care providers,

differentiating active corneal ulcers from healed scars may be a more challenging, less intuitive task.

A sophisticated form of pattern recognition—identifying key discriminating features of ulcers and scars—lies at the heart of proper diagnosis. For example, ophthalmologists look for key features such as corneal infiltrate, conjunctival injection, and hypopyon when making a diagnosis. Recent advancements in the fields of machine learning and computer vision have enabled a wide variety of new technologies, particularly in pattern recognition and image classification. In particular, deep learning—a specific subfield of machine learning that uses large neural networks—recently has been used in a variety of image classification tasks in medicine. Deep learning has been used to classify different forms of skin cancer,[4] to detect heart attacks,[5] and to diagnose neurological stroke,[6] among other conditions. In ophthalmology, machine learning algorithms have been demonstrated successfully to detect age-related macular

degeneration,[7] to classify diabetic retinopathy severity,[8] and to diagnose glaucoma.[9] In this study, we trained a convolutional neural network (CNN), a particular type of deep learning model, to distinguish corneal ulcers from scars.

Methods

External photographs were acquired from the Steroid for Corneal Ulcer Trial (SCUT) and the Mycotic Ulcer Treatment Trial (MUTT).[10,11] Photographs from both studies were captured using a handheld Nikon D-series digital SLR camera with a 105-mm f/2.8D AF Micro Nikkor Autofocus Lens and a modified Nikon SB29s electronic flash or Nikon R1 Wireless Close-up Speedlight system. In the SCUT, 500 patients with bacterial corneal ulcers underwent photography at enrollment and the 3-week, 3-month, and 12-month follow-up visits.[10] In the MUTT, 323 patients with fungal corneal ulcers underwent photography at enrollment and the 3-week and 3-month follow-up visits.[11] Both the SCUT and MUTT were conducted at the Aravind Eye Hospital in Tamil Nadu, India. This study was approved by the institutional review board at Stanford University and adhered to the tenets of the Declaration of Helsinki. The requirement for informed consent was waived because of the retrospective nature of the study.

Inclusion criteria in both the SCUT and MUTT required an active, culture-proven bacterial or fungal corneal ulcer. As such, photographs obtained at enrollment were labeled as active corneal ulcers. When clinical review indicated resolution of the ulcer, defined as complete epithelialization, images from the 3-month follow-up visits were labeled as healed corneal scars. Most ulcers healed well before 3 months, and the mean elapsed time between clinical resolution to the 3-month follow-up visit was 76.9 days (standard deviation, 12.1 days). Because some cases, especially fungal ulcers, can remain active despite epithelialization, if ambiguity existed regarding clinical resolution based on chart records, such as worsening with cessation of antimicrobials, these patients and photographs were excluded.

For both SCUT and MUTT, multiple photographs were obtained at each study visit. Except for using different external flash devices, the SCUT and MUTT used the same cameras and aperture priority settings, centered and focused on the corneal apex. Because of differences in data management of the image collections, only 1 photograph from each study visit was available for SCUT patients, whereas an average of 4.8 and 4.2 photographs for the enrollment and 3-month visits, respectively, were available for MUTT. To maximize the number of photographs that could be used, only exact duplicates and out-of-focus images were excluded.

Patients from SCUT and MUTT were pooled into a single list and were assigned randomly to the training, validation, and test datasets according to a 60/20/20 split. The training, validation, and test datasets were stratified in this way (by patient) to ensure that no patient cross-contamination occurred between the test set and other sets. A total of 612, 200, and 200 patients were allocated to the training, validation, and test sets, respectively. For MUTT patients, multiple images from the same patient study visit could be used for the training set, whereas for patients stratified to the test set, only 1 image each from the enrollment and 3-month visit could be included. This test set, called test set 1, contained patients and images from SCUT and MUTT not seen by the CNN during training.

A separate dataset of 101 slit-lamp photographs was acquired from the Byers Eye Institute at Stanford University in Stanford, California (Canon 7D digital camera and Haag-Streit BX900) and comprised 55 active corneal ulcers and 46 corneal scars from 101 unique patients. These photographs were classified based on clinical interpretation by a cornea specialist (C.C.L.). None of the images in this dataset, called test set 2, were used for training the CNN and were used only for testing.

All images were normalized to the training dataset and rescaled to 224 × 224 pixels. The training images from the SCUT and MUTT were used to train a deep learning model to classify active corneal infections and corneal scars. We used the VGG-16 architecture for the CNN, a specific type of deep learning model.[12] The model architecture is depicted in Figure 1. The model was pretrained on the 2014 ImageNet Large Scale Visual Recognition Challenge, a database of approximately 1.28 million images of 1000 object categories, which currently is the standard form of pretraining in PyTorch, a deep learning software package.[13,14] The final 1000-node layer typically used for ImageNet was replaced with a 2-node layer for our task.[12] The pretrained model was fine-tuned using the training dataset from the SCUT and MUTT. The model was trained to perform a binary classification task (active ulcer vs. healed scar) using cross-entropy loss, via Adam optimization algorithm with an initial learning rate of 10^{-5} that decayed by a factor of 10 every 5 training epochs. Training was continued for 13 epochs, when the validation loss began to increase, which indicated overfitting. When available, multiple photographs from each patient were used during fine-tuning.

After fine-tuning was completed, the performance of the model was measured on test sets 1 and 2. A single photograph from each patient in each test set was used to measure final performance. Confidence intervals (CIs) for final test performance were computed using the bootstrap method.

A precision-recall curve was generated to assess the false-positive rate and false-negative rate trade-off as a function of model confidence. A receiver operating characteristic (ROC) curve to assess the true-positive rate and false-positive rate tradeoff was also calculated. A calibration curve was created by binning the patient predictions into separate confidence bins of being an active ulcer and then plotting the proportion of patients in each bin who were determined by the ophthalmologist to harbor active ulcers. The bins were selected to provide a sufficient level of granularity that may be relevant for clinical predictions, while still permitted at least 10 patients per bin.

As a metric of model performance, the F_1 score was chosen because it is a commonly used statistic in settings where test sets may be imbalanced. It is computed as the harmonic mean of the precision and recall. Recall is defined as the number of predicted and true-positive findings divided by the number of true-positive findings and is a measure of the model's ability to identify active ulcers accurately. Precision is the number of predicted and true-positive findings divided by the number of false-positive findings and is a measure of the model's ability to identify healed infections.

The deep learning model also was compared with a simple baseline model that classifies images as ulcers and scars based on the proportion of red pixels. The exact proportion of red pixels that served as the cutoff for classifying an image as an active ulcer is a hyperparameter that was tuned on the training dataset.

Gradient-weighted class activation mapping (Grad-CAM) was used to visualize the model and to determine salient regions of active ulcers. Gradient-weighted class activation mapping uses the gradients of the target class (in this case, active ulcer) and creates a localization map that highlights regions of importance that were used to make the active ulcer versus healed scar prediction.[15]

A cornea specialist (C.C.L.) manually annotated 40 random photographs with an active ulcer from test set 1 to highlight areas of clinical relevance. Pixels in each image then were marked manually as important according to the cornea specialist's annotations, as well as by thresholding the Grad-CAM heatmap. The
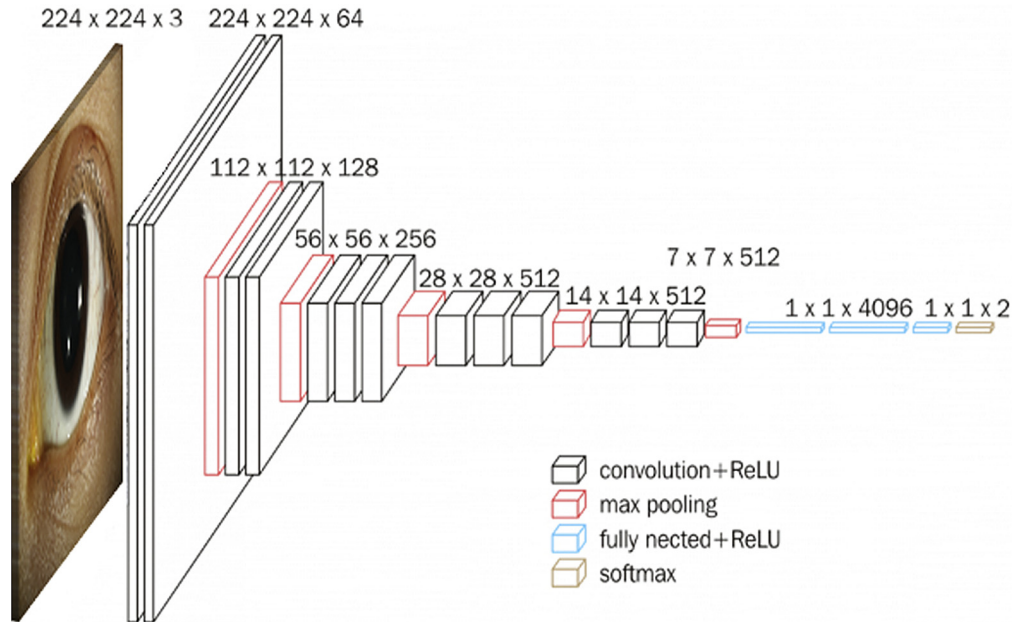
Figure 1. Architecture diagram showing convolutional neural network VGG-16. ReLU = rectified linear unit activation function.

similarity between Grad-CAM heatmaps and manually annotated photographs was calculated using the overlap coefficient, or Szymkiewicz−Simpson coefficient, defined as the number of pixels common to both importance maps divided by the number of pixels in the image with the smaller number of important pixels.[16] This segmentation analysis was not used to augment our deep learning classification model.

Results

On test set 1, comprising held out images from SCUT and MUTT that were reserved for testing, the model accurately classified 115 of 123 active corneal ulcers and 65 of 77 corneal scars, with an F_1 score of 92.0% (95% CI, 88.2%−95.8%). The sensitivity, or the ability of our model to identify active corneal ulcers correctly, was 93.5% (95% CI, 89.1%−97.9%). The specificity, or the ability of the model to identify true negative results, was 84.42% (95% CI, 79.42%−89.42%). Figure 2A shows the ROC curve, which plots the true-positive rate (TPR) against the false-positive rate (FPR) for the trained model (AUC, 0.9731). Figure 2B shows the precision-recall (PR) curve, which indicates the tradeoff between the FPR and false-negative rate as a function of decision threshold (AUC, 0.9811). More formally, the PR curve plots TPR / (TPR + FPR) against TPR / (TPR + false-negative rate). Figure 2C shows the calibration curve for test set 1.

On test set 2, comprising images from the Stanford Byers Eye Institute, the model correctly classified 43 of 55 active corneal ulcers and 42 of 46 corneal scars in patients, with an F_1 score of 84.3% (95% CI, 77.2%−91.4%), sensitivity of 78.2% (95% CI, 67.3%−89.1%), and specificity of 91.3% (95% CI, 85.8%−96.8%). Figure 2D−F shows the ROC (AUC, 0.9474), PR (AUC, 0.9524), and calibration curves for test set 2, respectively. The baseline classifier, which used the proportion of red pixels (defined as having a hue value between 0 and 10) in each achieved a lower accuracy of 76.0% (95% CI: 74.9%−77.1%) on the training set.

Figure 3 demonstrates a sample of heatmaps created by Grad-CAM. The heatmaps identify areas of the images that are relevant to the trained model's classification decisions. These areas align with clinical features associated with an active corneal ulcer, including conjunctival injection and corneal infiltrate (Fig 4).

Among the 40 randomly selected photographs with an active ulcer manually annotated by a cornea specialist to indicate areas of clinical relevance, some annotations overlapped significantly with areas highlighted by heatmaps (Fig 4), whereas in others, little or no overlap was found (Fig 5). The average overlap, or Szymkiewicz−Simpson coefficient, was 0.48. In cases of significant overlap (Szymkiewicz−Simpson coefficient, >0.7), the model accurately classified 10 of 12 ulcers (accuracy, 83%), whereas in cases of little overlap (Szymkiewicz−Simpson coefficient, <0.3), the model accurately classified only 9 of 16 ulcers (accuracy, 56%). As such, greater overlap is associated with better algorithm performance.

Discussion

In this work, we developed a deep learning model that can differentiate accurately between active corneal infections and healed scars. In particular, the model achieved a high level of accuracy in classifying corneal ulcers and scars when trained on photographs from the SCUT and MUTT. These results compare favorably with the performance of human experts in completing this task. A recent study from the University of Michigan evaluating cornea specialists' accuracy in diagnosing corneal ulcers from images captured from portable cameras reported a sensitivity of 82% to 94%.[17] Results from our model are similar, with sensitivities of 93.5% ± 4.4% and 78.2% ± 8.0% in test sets 1 and 2, respectively. These are promising results when taken in light of the proposed accuracy threshold of 80% sensitivity for telemedicine screening for conditions such as diabetic retinopathy.[18]

Crucial to the assessment of any machine learning model is its ability to "generalize," or perform well on data that were not used in the training dataset. Our model not only
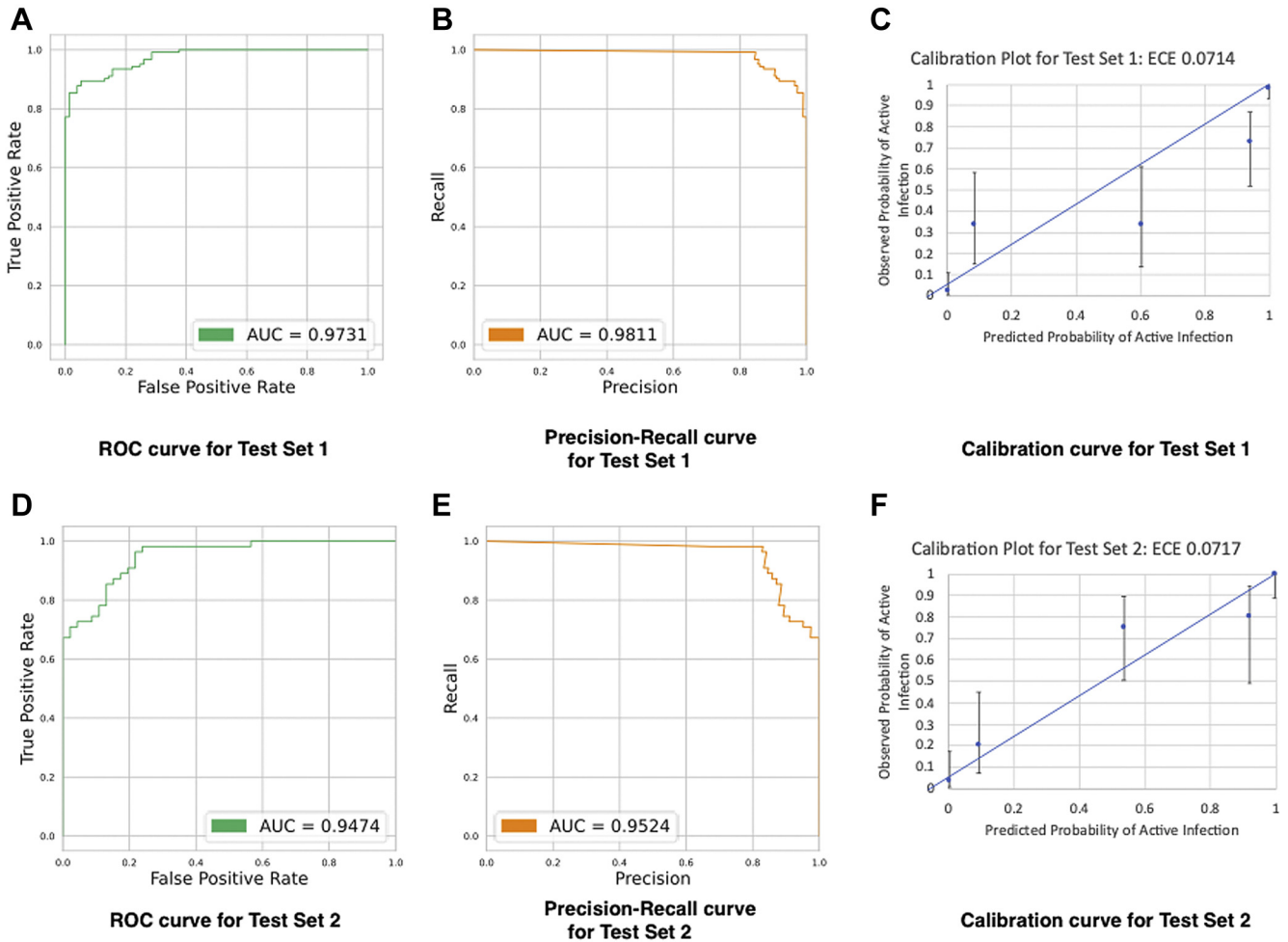
Figure 2. A–C, Test set 1 (**A**) receiver operating characteristic (ROC) curve, (**B**) precision-recall curve, and (**C**) calibration curve. D–F, Test set 2 (**D**) ROC curve, (**E**) precision-recall curve, and (**F**) calibration curve. AUC = area under the receiver operating characteristic curve; ECE = expected calibration error.

generalized to patients in test set 1 whose images it had never seen before, but also to a different patient population from the United States in test set 2. Notably, the Northern California population had significantly different patient demographics and microbiological epidemiologic features than the training data from the Indian clinics. The SCUT and MUTT enrolled patients predominantly from the Tamil Nadu state in South India, where a leading cause of corneal ulcers is agricultural trauma. In this tropical environment, fungal ulcers are more common than bacterial.[19] In contrast, in the coastal temperate climate of Northern California, most corneal ulcers are bacterial in origin and most often are caused by contact lens wear.[20] The generalizability of our model indicates that the algorithm identifies salient features of corneal ulcers and scars that are present despite epidemiologic differences.

In addition, the model generalizes well to a different imaging method, suggesting greater potential for a portable telemedicine implementation. The model was trained on images captured with a handheld camera, yet performed well on photographs captured with a slit-lamp–mounted camera. A robust evaluation of the performance of a model beyond simply overall accuracy is important as the

ramifications of misdiagnosis are significant. For example, misclassifying an active infection as a scar (false-negative result) can lead to corneal blindness, and classifying a scar as an active infection (false-positive result) can lead to unnecessary evaluations and treatment. The PR, ROC, and calibration curves in Figure 2 provide additional information to evaluate the model's performance. The high AUCs of the PR and ROC curves suggest that the FPR and false-negative rate are relatively low at reasonable decision thresholds. In practice, the decision threshold can be varied to balance the disparate costs associated with false-negative and false-positive results.

Furthermore, Figure 2 also contains calibration plots for the model's predictions on test Sets 1 and 2. The expected calibration error, a standard metric of calibration error, for both test sets 1 and 2 is approximately 7 percentage points, which is promising, although it may demonstrate mild overfitting in data-scarce regions. The 5 bins in the calibration plots were chosen to provide sufficient granularity in predicted confidence level to clinicians while ensuring more than 10 data points per bin.

The more sophisticated CNN outperformed the simple baseline classifier based solely on the number of red pixels
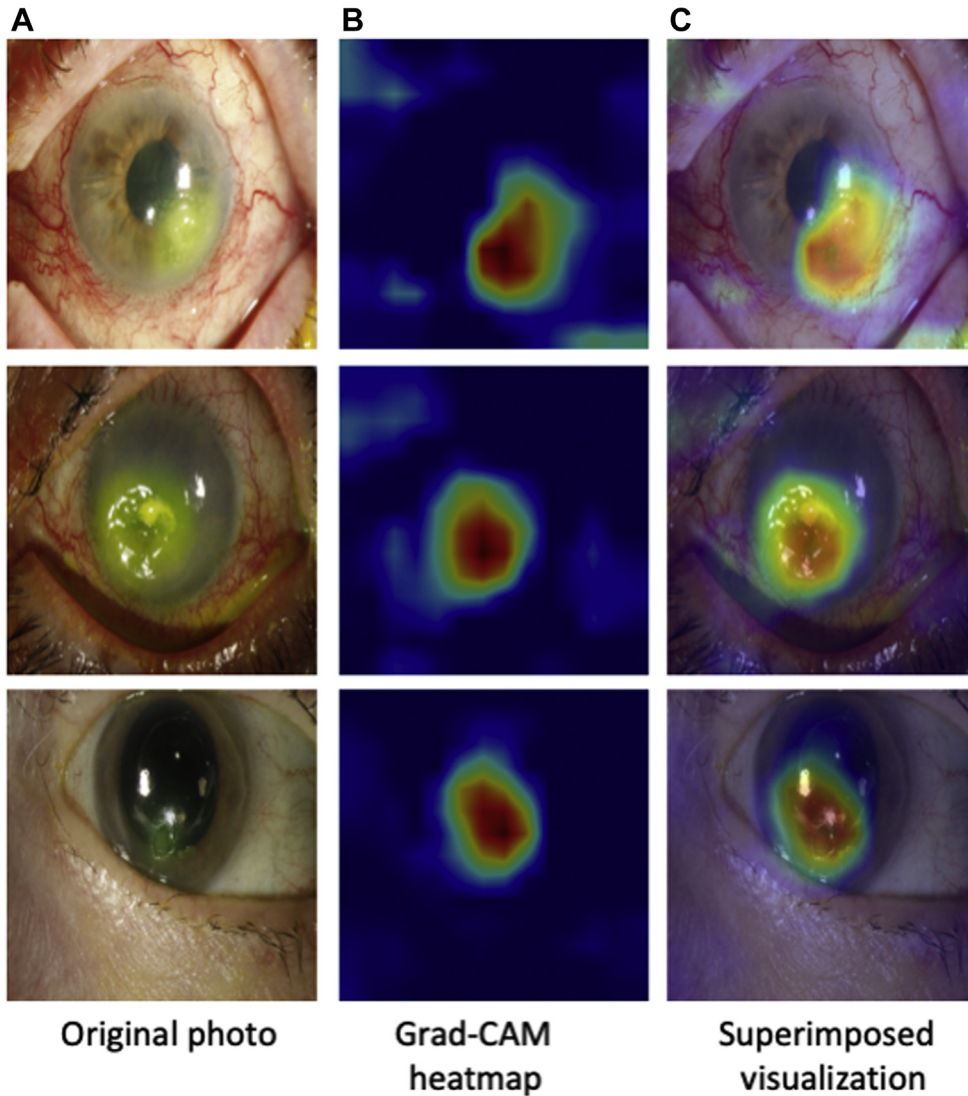
Figure 3. Examples of (**A**) original photographs, (**B**) gradient-weighted class activation mapping (Grad-CAM) heatmaps, and (**C**) superimposed visualization of the heatmap on the original photograph.
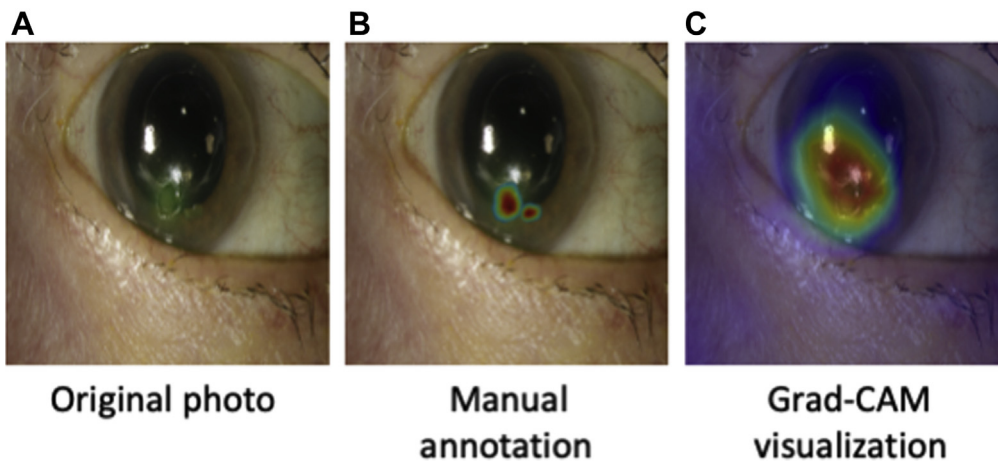

Figure 4. Comparison of (**A**) original photograph, (**B**) a manually annotated photograph of a corneal ulcer, and (**C**) its corresponding gradient-weighted class activation mapping (Grad-CAM) heatmap with good agreement.
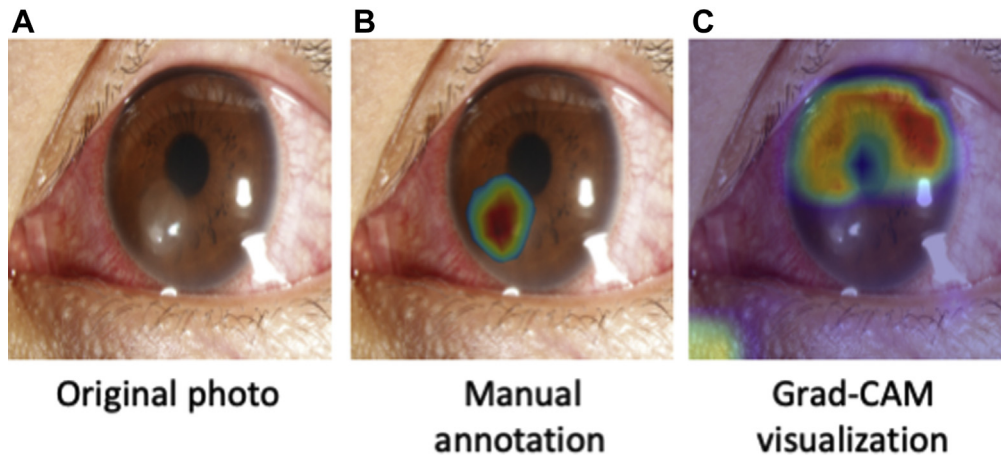
Figure 5. Comparison of (**A**) original photograph, (**B**) a manually annotated photograph of a corneal ulcer, and (**C**) its corresponding gradient-weighted class activation mapping (Grad-CAM) heatmap with poor overlap.

in an image. This suggests that the CNN identified features that are not accessible to this baseline classifier, such as the arrangement of redness (e.g., around the limbus). Furthermore, the Grad-CAM heatmaps (Figs 4 and 5) demonstrate that the model focuses on features beyond ciliary flush and conjunctival injection. Indeed, the heatmaps demonstrate the model's emphasis on clinically relevant features of an active infection, such as the presence of a hypopyon or purulent corneal infiltrate. Additionally, the generation of Grad-CAM visualizations does not require retraining or otherwise modifying the CNN, in contrast with some other CNN visualization techniques.[15] Such heatmaps may be presented to a medical provider for real-time human oversight of an automated diagnosis. Notably, the salient regions of an image in a classification prediction may not always be readily interpretable or even visible to the naked human eye. For example, features such as the amount of blood in sparsely or densely clustered blood vessels in an eye are harder to quantify visually. For these reasons, the regions highlighted by the model should be reviewed further by medical professionals. Discrepancies among model predictions and ophthalmologist predictions may allow for novel insights into clinical features of corneal ulcers. Furthermore, the human supervision provided by the ophthalmologist could be used to improve the model toward more fully automated artificial intelligence diagnosis.

The trained model is a step toward a portable system that can differentiate active corneal ulcers from healed infections, which may be helpful in areas with limited access to eye care. This model requires only low-resolution (224×224 pixels) photographs, which is well within the capabilities of most smartphones. The incorporation of this model into a smartphone-enabled application could provide a mobile diagnostic aid to assist triage efforts by a health care provider. Although much work lies ahead, including training providers to use the developed technology, potential exists to make a significant impact in developing countries that are disproportionately impacted by corneal blindness. In India, for example, only 1 ophthalmologist is available for every 52 000 residents, but almost 1 in every 3 people has a smartphone.[2,21] This model may help the triage process for primary care

providers, who generally have limited experience with eye care. Although patient history can provide important clues to providers regarding diagnosis, many conditions can mitigate symptoms of an active corneal ulcer, such as a neurotrophic cornea, treatment with topical steroids that mask inflammation, and an infection in its early stages. Patient-reported symptoms also vary considerably. For example, a scar can lead to debilitating photophobia for 1 patient and be relatively asymptomatic for another. Visualizing corneal pathologic features is crucial to accurate diagnosis, and our algorithm aims to support the diagnostic process through analysis of photographs. If a primary care provider flags a potential infection with this model, the patient then could be referred to an eye specialist. As telemedicine becomes more prevalent, especially in the era of coronavirus disease 2019, mobile diagnostic tools likely will become even more important.

Furthermore, the trained model could be used to help patients determine treatment efficacy. For example, patients could take serial photographs of the corneal ulcer to assess disease resolution. An ulcer that is not responding to treatment may be flagged as still active and may prompt additional consultation.

This study has a few limitations. First, a slight drop in generalization performance occurs when testing the model on test set 2, a patient population geographically and epidemiologically different from the training set. In practice, this could be remediated by continually retraining the model on new patients from each location in which the model is to be deployed. Second, all images in this study were obtained using high-resolution digital cameras. In these photographs, ambient lighting was relatively uniform, and out-of-focus images were excluded. Our model may not be as accurate with poorer quality photographs in which camera conditions such as lighting and focus are inconsistent. In practice, this may be ameliorated by publishing standardization guidelines for taking photographs to be classified by our model. Finally, our model was trained only on bacterial and fungal eye infections and may perform worse when tested on parasitic or viral infections. Future work could include additional data from such cases to improve our model.

In this work, we developed a deep learning model to classify corneal ulcers and scars that performs well on out-of-sample data. Our model offers a promising, inexpensive diagnostic approach with potential to aid triage in communities with limited access to eye care. In addition, the model also may be used as an adjunctive tool to help monitor clinical progression and to prompt re-evaluation in refractory cases.

Footnotes and Disclosures

Correspondence:

Charles C. Lin, MD, Byers Eye Institute, Stanford University, 2452 Watson Court, Palo Alto, CA 94303. E-mail: lincc@stanford.edu.

References

1. Austin A, Lietman T, Rose-Nussbaumer J. Update on the management of infectious keratitis. *Ophthalmology.* 2017;124(11):1678−1689.
2. Statista. India—number of ophthalmologists 2005−2020. https://www.statista.com/statistics/953172/india-number-of-ophthalmologists/; 2020. Accessed 22.12.20.
3. Healio. Access to eye care, uptake of services are issues in India. https://www.healio.com/news/ophthalmology/20120225/access-to-eye-care-uptake-of-services-are-issues-in-india. Accessed 22.12.20.
4. Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature.* 2017;542(7639):115−118. published erratum appears in *Nature.* 2017;546(7660):686).
5. Strodthoff N, Strodthoff C. Detecting and interpreting myocardial infarction using fully convolutional neural networks. *Physiol Meas.* 2019;40(1):015001.
6. Zech J, Pain M, Titano J, et al. Natural language-based machine learning models for the annotation of clinical radiology reports. *Radiology.* 2018;287(2):570−580.
7. Burlina PM, Joshi N, Pekala M, et al. Automated grading of age-related macular degeneration from color fundus images using deep convolutional neural networks. *JAMA Ophthalmol.* 2017;135(11):1170−1176.
8. Gulshan V, Peng L, Coram M, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA.* 2016;316(22):2402−2410.
9. Li Z, He Y, Keel S, et al. Efficacy of a deep learning system for detecting glaucomatous optic neuropathy based on color fundus photographs. *Ophthalmology.* 2018;125(8):1199−1206.
10. Srinivasan M, Mascarenhas J, Rajaraman R, et al. Corticosteroids for bacterial keratitis: the Steroids for Corneal Ulcers Trial (SCUT). *Arch Ophthalmol.* 2012;130(2):143−150.

11. Prajna NV, Krishnan T, Mascarenhas J, et al. The mycotic ulcer treatment trial: a randomized trial comparing natamycin vs voriconazole. *JAMA Ophthalmol.* 2013;131(4): 422−429.

12. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. Available at: http://arxiv.org/abs/1409.1556. Accessed 20.12.20.

13. Deng J, Dong W, Socher R, et al. ImageNet: a large-scale hierarchical image database. In: *2009 IEEE Conference on Computer Vision and Pattern Recognition.* IEEE; 2009:248−255.

14. Paszke A, Gross S, Massa F, et al. PyTorch: an imperative style, high-performance deep learning library. http://arxiv.org/abs/1912.01703. Accessed 22.12.20.

15. Selvaraju RR, Cogswell M, Das A, et al. Grad-CAM: visual explanations from deep networks via gradient-based Localization. In: 2017 IEEE International Conference on Computer Vision (ICCV). 2017:618−626.

16. Vijaymeena MK, Kavitha K. A survey on similarity measures in text mining. *Machine Learning and Applications: An International Journal.* 2016;3(1):19−28.

17. Woodward MA, Musch DC, Hood CT, et al. Tele-ophthalmic approach for detection of corneal diseases: accuracy and reliability. *Cornea.* 2017;36(10):1159−1165.

18. British Diabetic Association. Retinal Photography Screening for Diabetic Eye Disease: A British Diabetic Association Report. London: British Diabetic Association; 1997.

19. Lin CC, Lalitha P, Srinivasan M, et al. Seasonal trends of microbial keratitis in South India. *Cornea.* 2012;31(10):1123−1127.

20. Jeng BH, Gritz DC, Kumar AB, et al. Epidemiology of ulcerative keratitis in Northern California. *Arch Ophthalmol.* 2010;128(8):1022−1028.

21. Vaibhav A. Smartphone penetration in India 2014−2022. https://www.statista.com/statistics/257048/smartphone-user-penetration-in-india/. Accessed 22.12.20.

Pictures & Perspectives
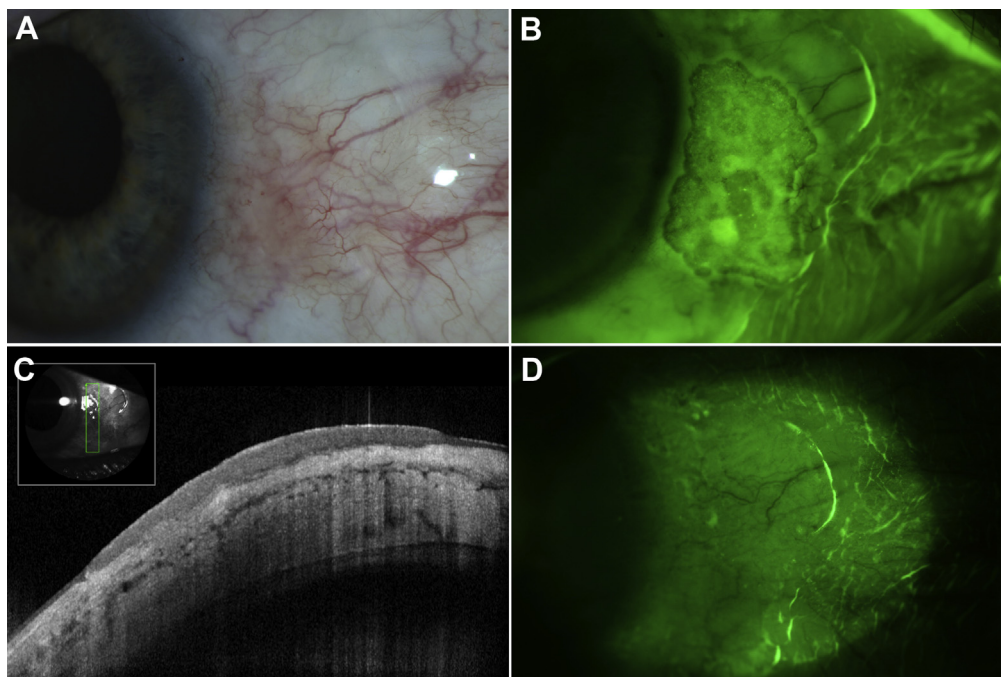
Visualizing Ocular Surface Squamous Neoplasia

A 66-year-old man presented with subtle conjunctival thickening (Fig **A**) 3 years after treatment for conjunctival intraepithelial neoplasia in the same location by primary surgical excision, cryotherapy, and adjuvant topical mitomycin C (MMC) 0.02%. Visualization of the ocular surface with a yellow barrier filter over cobalt blue light after instillation of fluorescein identified stippled staining of recurrent tumor, and the tumor margins were also clearly outlined (Fig **B**) corresponding to those found by anterior-segment OCT (Fig **C**). Six months after treatment with 2 cycles of topical MMC 0.02%, repeat fluorescein staining showed complete tumor regression (Fig **D**). (Magnified version of Fig **A**-**D** is available online at www.aaojournal.org).

EMILY M. WITSBERGER, MD
LAUREN A. DALVIN, MD
SANJAY V. PATEL, MD, FRCOPHTH

Department of Ophthalmology, Mayo Clinic, Rochester, Minnesota

Footnotes and Disclosures