AMERICAN ACADEMY OF OPHTHALMOLOGY®

Assessing the Impact of Image Quality on Deep Learning Classification of Infectious Keratitis

Adam Hanif, MD,[1] N. Venkatesh Prajna, MD,[2] Prajna Lalitha, MD,[2] Erin NaPier, BA,[3] Maria Parker, MD,[1] Peter Steinkamp, MS,[1] Jeremy D. Keenan, MD, MPH,[4] J. Peter Campbell, MD, MPH,[1] Xubo Song, PhD,[5] Travis K. Redd, MD, MPH[1]

Objective: To investigate the impact of corneal photograph quality on convolutional neural network (CNN) predictions.

Design: A CNN trained to classify bacterial and fungal keratitis was evaluated using photographs of ulcers labeled according to 5 corneal image quality parameters: eccentric gaze direction, abnormal eyelid position, over/under-exposure, inadequate focus, and malpositioned light reflection.

Participants: All eligible subjects with culture and stain-proven bacterial and/or fungal ulcers presenting to Aravind Eye Hospital in Madurai, India, between January 1, 2021 and December 31, 2021.

Methods: Convolutional neural network classification performance was compared for each quality parameter, and gradient class activation heatmaps were generated to visualize regions of highest influence on CNN predictions.

Main Outcome Measures: Area under the receiver operating characteristic and precision recall curves were calculated to quantify model performance. Bootstrapped confidence intervals were used for statistical comparisons. Logistic loss was calculated to measure individual prediction accuracy.

Results: Individual presence of either light reflection or eyelids obscuring the corneal surface was associated with significantly higher CNN performance. No other quality parameter significantly influenced CNN performance. Qualitative review of gradient class activation heatmaps generally revealed the infiltrate as having the highest diagnostic relevance.

Conclusions: The CNN demonstrated expert-level performance regardless of image quality. Future studies may investigate use of smartphone cameras and image sets with greater variance in image quality to further explore the influence of these parameters on model performance.

Financial Disclosure(s): Proprietary or commercial disclosure may be found after the references.

Corneal opacification is responsible for approximately 3.5% of blindness and is the fifth leading cause of blindness in the world.[1,2] Infectious keratitis (or "corneal ulceration") is the most common cause of corneal opacification, with low- and middle-income countries demonstrating a particularly high burden of disease.[1,3,4] Prompt diagnosis of an ulcer's microbial etiology can facilitate more rapid initiation of treatment and improve outcomes. Cultures of corneal scrapings are currently considered the diagnostic gold standard for infectious keratitis. However, there are significant limitations to this approach as cultures take multiple days to grow, require infrastructure and expertise to interpret, and frequently result in false negatives.

The advent of deep learning (DL) for image classification in ophthalmology has introduced convolutional neural networks (CNNs) capable of expert-level diagnostic performance in use cases including diabetic retinopathy, age-related macular degeneration, and glaucoma.[5–7] In the context of corneal disease, DL has increasingly been investigated as a solution for automated diagnosis of bacterial and fungal ulcers. However, preliminary investigations of this method utilized images obtained from slit lamp-mounted cameras which may difficult to operate, less plentiful, and more expensive than handheld digital cameras in resource-poor settings that lack ophthalmic professionals.[8–13] Recently, we demonstrated superior-to-human performance of a CNN trained to classify bacterial and fungal corneal ulcers from images obtained with a handheld camera, an imaging method with greater potential for a portable telemedicine implementation.[14] These findings are promising, though little has been studied regarding the impact of external photograph quality on CNN predictions. In developing regions where such models are likely to have the greatest public health impact, images might conceivably be captured with handheld devices by untrained photographers. The resultant variability in photographic quality may influence CNN performance and thus should be quantified before

https://doi.org/10.1016/j.xops.2023.100331
ISSN 2666-9145/23

such technology could be implemented. While the impact of image quality on diagnostic DL models has been increasingly investigated in the context of ophthalmology, relatively little is known regarding the specific features of external clinical photographs determining quality and their relative influence on model accuracy.[15–17] In this study, we explored the impact of image quality parameters including gaze direction, eyelid position, exposure, focus, and light reflections on a deep CNN's performance in classifying bacterial and fungal corneal ulcers.

Methods

Participant Recruitment and Image Collection

Corneal photographs were prospectively obtained from all consenting patients with a clinical diagnosis of infectious keratitis presenting to Aravind Eye Hospital in Madurai, India from January 1, 2021 to December 31, 2021. Upon initial presentation, each patient underwent corneal photography with a handheld Nikon D-series digital single-lens reflex camera according to a standardized lighting and photography protocol. Image collection was repeated until $\geq$ 1 high-quality photograph was obtained for each ulcer, as determined by the trained photographer. A single trained photographer performed all image acquisition and was masked to the type of corneal infection. The photographer was instructed to ensure that the image was in focus on the corneal infiltrate (the area of active infection within the cornea) and to displace the light reflection from the camera flash when needed to avoid obscuring the infiltrate. In cases of large infiltrates where overlap of the light reflection and infiltrate was unavoidable, the photographer was instructed to place the flash in the center of the infiltrate rather than the border, based on prior research suggesting the morphology of the border of corneal infiltrates may be a particularly high-yield feature in distinguishing bacterial and fungal keratitis.[18]

In all cases, scrapings were obtained from the ulcer to perform smears (Gram stain and potassium hydroxide prep) and cultures (bacterial cultures on blood agar and fungal cultures on potato dextrose agar). Because of the known limitations in the individual sensitivity and specificity of corneal cultures and smears, we elected to use a composite of these results to establish the gold standard label of the microbiologic etiology of infection in each case.[19] Specifically, an ulcer was classified as "bacterial" if both the smear (Gram stain) and culture results indicated the presence of bacteria, and "fungal" if both the smear (potassium hydroxide prep) and culture results indicated the presence of filamentous fungus. Only cases with a clear identification of fungal and/or bacterial etiology according to these criteria were included in the study. In some cases, polymicrobial infections were observed, as indicated by presence of bacteria and filamentous fungus confirmed by the criteria mentioned above. A deep CNN (MobileNetV2) was trained to perform multilabel classification of bacterial and fungal keratitis using the collected corneal photographs using a similar approach to our previously published model.[14] A detailed description of model development and evaluation is provided elsewhere (manuscript in preparation), and all relevant code is publicly accessible (github.com/tkredd2/MobileNet_corneal_ulcers).

Image Quality Grading

After excluding participants with images in the training and validation sets used to develop the MobileNet model, images were randomly selected to undergo image quality grading according to 5 prespecified parameters: gaze direction (primary or eccentric), eyelid position (open or obscuring either $\geq$ 1 limbal clock hour or $\geq$ 10% of the infiltrate), degree of light exposure (normal exposure [where details are adequately visible], underexposed [where images looked too dark and details are somewhat lost], or overexposed [where images look too bright and details are somewhat blurry]), whether or not the image was in focus on the corneal infiltrate, and the presence or absence of light reflections causing significant obscuration of the corneal infiltrate (Figure 1). These parameters were identified as the most likely features that may influence the ability to perform image-based identification of bacterial and fungal keratitis based on the clinical expertise of several cornea specialists and based on subjective evaluation of heatmaps generated from a previously-published DL model trained to differentiate bacterial and fungal keratitis from corneal photographs.[14] Based on these evaluations, we expected eccentric gaze, eyelid obscuration of the cornea, over- or underexposure, out-of-focus images, and/or light reflection obscuration of the corneal infiltrate to negatively impact the computer vision model's performance. Two professional image graders at the Casey Reading Center (M.P. and P.S.) recorded their interpretation of each of the 5 parameters for each image. Each grader underwent 2 separate hour-long training sessions with a cornea specialist (T.K.R.) to practice identifying the quality parameters of interest. Grading was performed independently by each expert, and images were presented in random order. Graders were masked to each other's interpretations of the images, as well as the microbiologic status of the ulcer represented by the image. In cases where the two graders demonstrated disagreement on any image quality parameter, the image in question was submitted to a third grader (T.R.) for adjudication, resulting in a final consensus image quality assessment.

Considering the fact that other quality parameters not included in the above analysis may be influential, we separately evaluated the model's performance on another test set composed of all images obtained from each participant (excluding those included in the training set for the MobileNet model). This was compared to model performance on the test set comprised of the "best quality image" for each participant, subjectively determined by a single image grader (T.R.) without specification of particular image quality parameters, but instead determining which image this cornea specialist felt represented the highest quality depiction of the corneal ulcer.

Model Performance Evaluation

Because of the class imbalance (i.e., the larger proportion of fungal compared to bacterial corneal ulcers) in this dataset, we selected 2 primary metrics to quantify model performance: the area under the receiver operating characteristic curve (AUROC) and area under the precision recall curve (AUPRC) (Figure 2). The ROC is somewhat easier to interpret but is more susceptible to class imbalance, whereas the PRC curve is less intuitive but more robust to imbalanced classes. Because this CNN performed multilabel classification of bacterial and fungal ulcers (i.e., class predictions were not mutually exclusive, making it possible for the model to predict that an ulcer would be both bacterial and fungal), we calculated microaveraged metrics across the bacterial and fungal predictions to obtain single overall values for the AUROC and AUPRC.[20] Bivariate comparisons of the model's performance were then performed for each of the 5 quality parameters, using bootstrapping at the subject level with 10 000 replications to generate 95% confidence intervals (CIs). Multivariate comparisons were also performed based on these bivariate results. For each output prediction of either fungal (P[fungal]) or bacterial (P[bacterial]) etiology, the logistic (i.e., cross entropy) loss was calculated to provide a subject-level measure of model performance.[21] Gradient class activation maps were used to qualitatively assess which image regions exerted

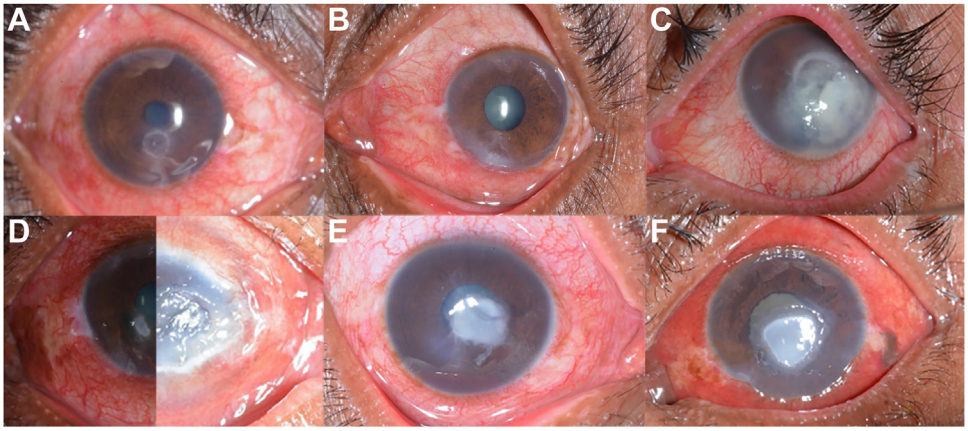

Figure 1. Representative images demonstrating photographs without (**A**) and with (**B**–**F**) the designated image quality parameters, including eccentric gaze (**B**), corneal lid obscuration (**C**), under- (**D,** left) and overexposure (**D,** right), infiltrate features out of focus (**E**), and reflection obscuring infiltrate features (**F**).

the greatest influence on CNN predictions.[22] Statistical analyses were performed in Python 3 using the Scikit-learn library and in R software version 4.0.5 (R Foundation for Statistical Computing).

This study adhered to the tenets of the Declaration of Helsinki and was approved by the institutional review board at Oregon Health & Science University and the Institutional Ethics Committee and Aravind Eye Hospital in Madurai, India. Informed consent was obtained from all participants.

Results

Eight hundred ninety-seven images from 332 participants underwent assessment and labeling by the expert graders. Approximately half these participants demonstrated culture- and smear-confirmed bacterial or fungal keratitis, resulting in a final image set of 438 images from 164 participants included in this study. Three hundred seventy-four images were from fungal corneal ulcers (85%), and the remaining 64 images were from bacterial corneal ulcers. Table 1 exhibits the distribution of quality parameters within the dataset. For cases in which the independent graders disagreed (gaze = 108, eyelid obscuration = 26, exposure = 107, focus = 103, reflection = 38), adjudication by a third grader was performed to give a final classification. Convolutional neural network performance, as measured by AUROC, was found to be 0.85 for fungal cases and 0.79 for bacterial cases. No significant difference in performance on each ulcer class was found by this metric. For images in which the corneal reflection was labeled as significantly obscuring ulcer

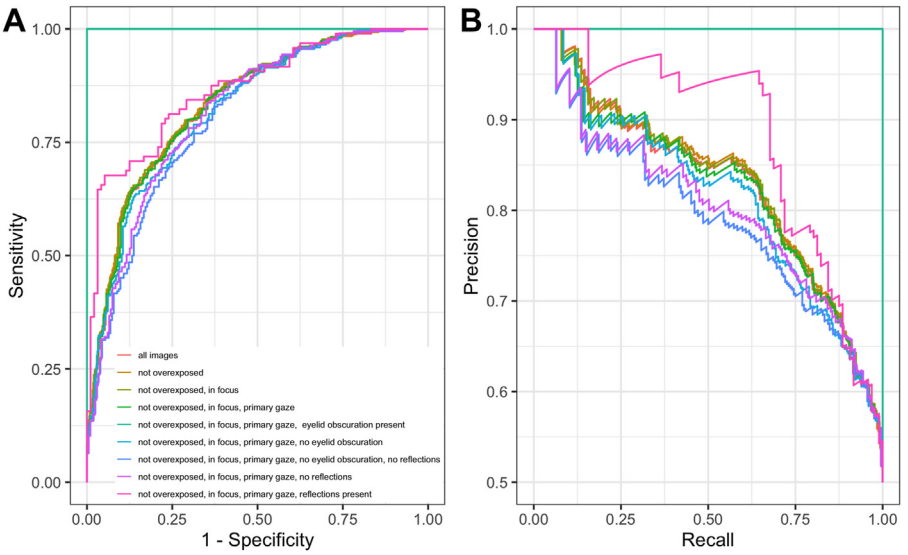

Figure 2. Receiver operating characteristic curve (ROC) (**A**) and precision recall curve (PRC) (**B**) curves of model performance on all images in the test set, as well as images with varied combinations of image quality parameters. Convolutional neural network performance on the entire test set demonstrated a microaveraged area under the ROC of 0.83 (95% confidence interval [CI], 0.80–0.87) and area under the PRC of 0.83 (95% CI, 0.78–0.87) for fungal and bacterial classification.

Table 1. CNN Multiclass Predictions on Image Subsets with Either Presence or Absence of Each Quality Parameter

		Bacterial	Fungal	AUPRC (95% CI)	AUROC (95% CI)
Reflection	Present	11	145	0.91 (0.86–0.94)	0.90 (0.85–0.93)
	Absent	53	229	0.79 (0.73–0.84)	0.80 (0.74–0.85)
	Overexposed	4	16	0.77 (0.59–0.93)	0.77 (0.44–0.94)
Exposure	Normal	52	276	0.83 (0.78–0.87)	0.83 (0.78–0.87)
	Underexposed	8	82	0.83 (0.71–0.91)	0.86 (0.76–0.92)
Focus	In focus	50	320	0.83 (0.78–0.87)	0.83 (0.79–0.87)
	Out of focus	14	54	0.83 (0.71–0.90)	0.83 (0.72–0.91)
Gaze	Eccentric	5	50	0.89 (0.73–0.95)	0.87 (0.77–0.93)
	Primary	59	324	0.82 (0.78–0.86)	0.83 (0.79–0.87)
Lid obscuration	Yes	5	28	0.94 (0.86–0.98)	0.93 (0.84–0.98)
	No	59	346	0.83 (0.78–0.86)	0.83 (0.78–0.86)

AUPRC = area under the precision recall curve; AUROC = area under the receiver operating characteristic curve; CI = confidence interval; CNN = convolutional neural network.

features, the CNN performed significantly better than on those without significant reflection interference (significant reflection: AUROC = 0.90 [95% CI, 0.85–0.93], AUPRC = 0.91 [95% CI, 0.86–0.94]; insignificant reflection: AUROC = 0.80 [95% CI, 0.74–0.85], AUPRC = 0.79 [95% CI, 0.73–0.84]). For images in which the eyelid significantly obscured the view of the cornea, the CNN performed significantly better than on those without significant lid obscuration (significant lid obscuration: AUPRC = 0.94 [95% CI, 0.86–0.98], insignificant lid obscuration: AUPRC = 0.83 [95% CI, 0.78–0.86]). There were no significant differences in the model's performance with respect to gaze, focus, or image exposure (Table 1). These associations persisted on multivariate analysis (Figure 3). For all 17 images in which lid obscuration was noted as present, but were otherwise not over/under-exposed, in focus, and in primary gaze, the model correctly identified the microbial etiology of infection. Model performance was also excellent on images identified as having significant light reflections obscuring the infiltrate, but with no eccentric gaze, defocus, or over/underexposure (AUROC = 0.86 [95% CI, 0.77–0.92]; AUPRC = 0.87 [95% CI, 0.78–0.93]).

When evaluating quality subjectively by evaluating the CNN on the "best quality" image from each participant (n = 278) versus all images for each participant (n = 733), we found no difference in the model's performance. Specifically, the CNN achieved an AUROC of 0.88 (95% CI, 0.84–0.92) and AUPRC of 0.89 (95% CI, 0.85–0.92) on the "best quality" image set, compared to an AUROC of 0.89 (95% CI, 0.86–0.91) and an AUPRC of 0.89 (95% CI, 0.86–0.91) on the multiple-image-per-participant set.

Gradient Class Activation Heatmaps

Gradient class activation heatmaps were generated for 2 sets of images. The first set of heatmaps were generated for the 10 images with the highest model performance (determined by lowest logistic loss) which were labeled as having significant light reflections obscuring the infiltrate (Figure 4). The model appeared to consistently identify the corneal infiltrate as the primary region of interest in these 10 cases. A second set of heatmaps was generated for all 17 images of corneal ulcers labeled as having significant obscuration of the cornea by the eyelid, but with no other image quality concerns (i.e., primary gaze, in focus, and with good image exposure) (Figure 5). Eleven of the 17 heatmaps demonstrated a similar region of activation focused on the corneal infiltrate, but the remaining 6 (35%) demonstrated more diffuse activation which focused more on the ocular surface and eyelids, with 3 (17.6%) such cases appearing to exclude the cornea entirely.

Discussion

To our knowledge, this represents the first study evaluating the influence of variable quality in corneal photographs on CNN classification of bacterial and fungal ulcers. Overall, the CNN achieved similar performance in this task compared to prior studies, with accuracy at least as good as expert clinicians.[14,23] In this set of images, neither eccentric gaze, out-of-focus images, eyelid obscuration of the cornea, over/underexposure, or obscuration of the corneal infiltrate by light reflections significantly decreased the model's performance in identifying the underlying microbial cause of infection. However, eyelid obscuration of the cornea and light reflection causing significant obscuration of infiltrate features were actually noted to significantly improve AUROC and AUPRC values. Even in cases of suboptimal image quality, the CNN achieved high performance on this task, which has promising implications for its translation to more portable imaging devices, including smartphones.

While no individual quality parameter negatively impacted model performance, reflection and lid obscuration, when present, counterintuitively improved performance. Though the specific cause of these observations remains unclear, it is possible that the enhanced CNN performance reflects model detection of microbe-level differences in presentation and infiltrate phenotype that are associated with these quality parameters. We considered the possibility that because some strains of fungal ulcers are more indolent at onset, they may ultimately present in the clinic with greater severity and discomfort, posing a greater challenge to

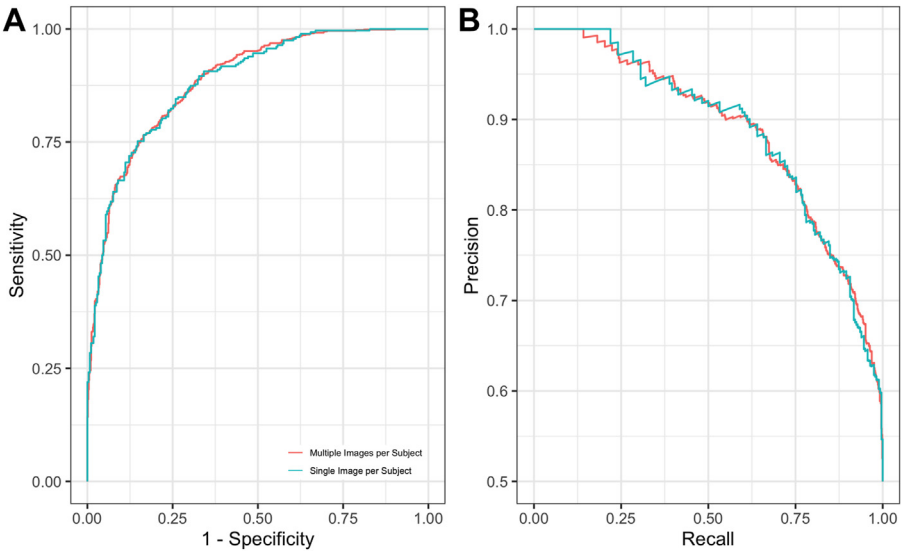

Figure 3. Receiver operating characteristic curve (ROC) (**A**) and precision recall curve (PRC) (**B**) curves of model performance test sets representing all images subjectively determined to be the highest quality for each participant ("Single Image per Participant", area under the ROC [AUROC] = 0.88 [95% confidence interval {CI}, 0.84–0.92], area under the PRC [AUPRC] = 0.89 [95% CI, 0.85–0.92]), and separately all images associated with all participants ("Multiple Images per Participant", AUROC = 0.89 [95% CI, 0.86–0.91], AUPRC = 0.89 [95% CI, 0.86–0.91]). No significant differences were found between performances on these test sets by either performance measurement.

sustained lid opening during corneal photography.[24] Similarly, a more severe ulcer with a larger infiltrate is more likely to be partially obscured by the light reflection of the camera flash, thus, this quality metric may have simply served as a surrogate measure of ulcer severity rather than providing a useful feature in itself. Alternatively, it is possible that microbe-specific variations in the surface texture of a corneal infiltrate are demonstrated by the specular information provided by the reflection of the camera flash. Specifically, filamentous fungal corneal infiltrates often have a dry, rough texture, whereas bacterial ulcers often cause more stromal necrosis with a more "soupy" appearance.[25]

We additionally considered explanations for why other quality parameters demonstrated no influence on model performance. Regarding exposure, it is possible that although varying levels of lighting may alter human perception and interpretation of these images, sufficient detail was retained in this image set to be detected and informative to the model. With regard to focus, we considered the possibility that the loss of fine, "high-frequency" image detail in out-of-focus cases may be irrelevant to the algorithm, which might instead rely on "low-frequency" spatial and general intensity features to make predictions.

Of note, heatmaps of these images labeled as having significant obscuration from light reflections appeared to demonstrate consistently high activation within the region of the corneal infiltrate. A similar pattern of greatest activation in the area of the infiltrate was evident in 11 of 17 images labeled as having significant lid obscuration but with good focus, exposure, and gaze. However, the remaining 6 images revealed a more diffuse pattern of activation, entirely missing the corneal surface in 3 such cases. Nonetheless, in all 17 images the CNN correctly predicted the underlying etiology of infection. It remains unclear how features of the ocular surface and lids involved in these diffuse patterns of activation may contribute to the prediction as meaningfully as the infiltrate. It is possible that in these cases the CNN takes features external to the cornea into account like conjunctival injection, which may be less prominent in fungal keratitis which may have a "quieter" presentation. The CNN may also be detecting features yet unrecognizable by humans, as has been increasingly reported of DL systems capable of identifying nonophthalmic clinical traits from retinal images including cardiovascular disease risk, Alzheimer's disease, and gender.[26–28] Alternatively, it is also possible that in these instances the process by which the heatmap is generated is flawed, despite an accurate classification. As with all interpretability methods, heatmap explanations should be interpreted in the context of their clinical application and with an appropriate degree of scrutiny. Additional studies with more examples of obscuration from the eyelids and light reflections, as well as measurement of ulcer severity as a covariate, will be required to differentiate these possibilities, but regardless, neither appeared to negatively affect model performance in this study.

Images were captured using handheld cameras, which may be the preferred option in resource-poor settings where the alternative slit-lamp-mounted cameras may be expensive, harder to use by nonophthalmic professionals, and less portable, despite producing photographs of similar quality.[11–13] While the use of handheld cameras inevitably introduces greater variation in angle, distance, and lighting, we found model performance remained high despite variations in gaze, lid position, exposure, focus, and light obscuration. This reinforces the notion that these portable

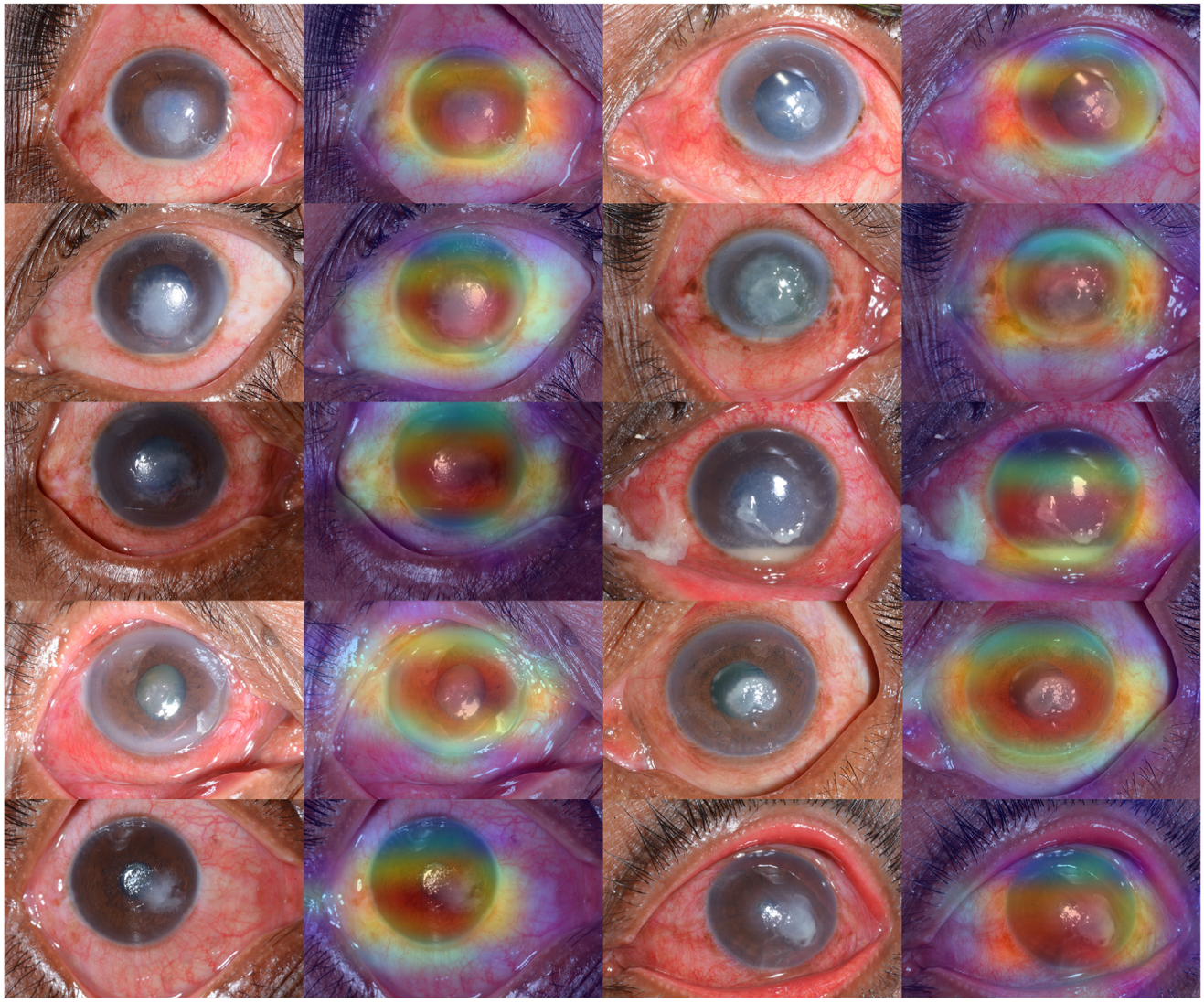

Figure 4. Ten corneal photographs and their corresponding gradient class activation heatmaps that were labeled as having significant reflection obscuring the corneal infiltrate and the lowest log loss amongst other images with the same label. Heatmaps indicate regions of greatest contribution to the model's predictions or "interest" according to colors on the continuous visible spectrum (red = highest relevance to model prediction). By the authors' qualitative review, the cornea and associated infiltrate were identified as having the greatest relevance to the model prediction.

imaging modalities may provide adequate information for DL models to make accurate predictions regarding the underlying cause of infection. In this way, DL could offer a viable solution for improving diagnosis of infectious keratitis where human assessment has been shown to be less accurate, and the gold-standard of cultures from corneal scrapings often involves a several-day lag time, shows high false-negative rates, and requires lab infrastructure which may be unavailable in developing countries.[19,29]

These findings should be interpreted in the context of several limitations. While the image set in this study is comparable or larger than related studies using computer vision for infectious keratitis, higher volume may still be required to capture enough photographs of varying combinations of quality parameter labels to identify a conclusive impact on model performance.[9,10,30] In addition to volume, greater variation in photograph quality may additionally provide a more robust analysis of the influence of our chosen quality parameters on model performance. In this study all images were captured by trained ophthalmic photographers instructed to obtain $\geq$ 1 high-quality photograph per ulcer, according to their judgment. These photographers are excellent, hence the relatively low number of poor-quality photos despite the large sample size. We will perform similar analyses in future studies with a dataset of images captured on smartphone cameras by amateur photographers. An additional limitation of the study involves the use of photographs with labels for microbial etiology derived from confirmation by both smear and culture results. While useful in ensuring validity of labels, the confirmation of microbial etiology is not always obtainable from all ulcers. Because the training sets did not include cases that were "ambiguous" in this regard, the results are somewhat

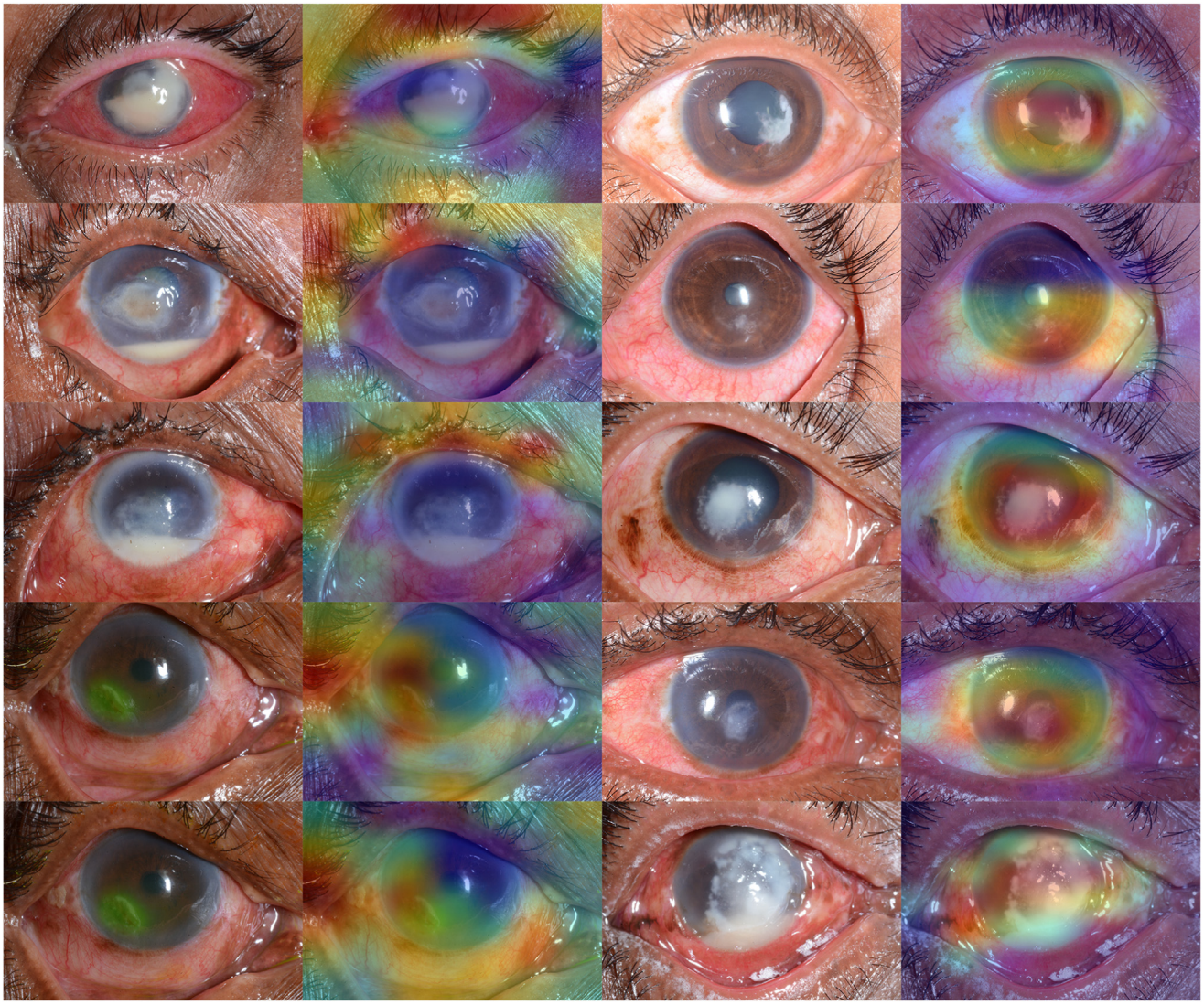

Figure 5. Ten corneal photographs and their corresponding gradient class activation heatmaps that were labeled as having significant eyelid obscuration of the cornea, but otherwise in focus, not overexposed, and in primary gaze. Heatmaps indicate regions of greatest contribution to the model's predictions or "interest" according to colors on the continuous visible spectrum (red = highest relevance to model prediction). The left column of pairs demonstrates a diffuse pattern of "interest" which appears to focus on the ocular surface and eyelid tissue, excluding the cornea in some cases. The right column of pairs demonstrates a distribution around the central cornea and associated infiltrate. The model correctly identified the microbial etiology in all cases.

less generalizable. Lastly, the quality parameters included for labeling in this study were chosen according to a subjective consensus by the authors on the image features thought to be most impactful on adequate interpretation of a corneal infiltrate. We acknowledge that other image features not labeled and evaluated in this study may influence CNN interpretation of microbial etiology in infectious keratitis. These may include effect of distance on relative size of cornea to the image frame, lack of flash during image capture, and presence of other artifacts like finger holding upper lid and patient nose.

From these observations, we conclude that a CNN trained to identify bacterial and fungal corneal ulcers in a multilabel classification task exhibits high predictive accuracy, even in the presence of variable image quality. Presence of features like eyelid obscuration of the cornea or obscuration of the infiltrate by corneal light reflection may be informative to model predictions. Future investigation of image quality with datasets of greater volume and quality variability may further establish these trends and elucidate features of relevance to computer vision models for the diagnostic evaluation of infectious keratitis.

Footnotes and Disclosures

Originally received: January 21, 2023.
Final revision: April 13, 2023.
Accepted: May 8, 2023.
Available online: May 16, 2023. Manuscript no. XOPS-D-23-00014.

[1] Casey Eye Institute, Oregon Health & Science University, Portland, Oregon.

[2] Aravind Eye Hospital, Madurai, Tamil Nadu, India.

[3] John A. Burns School of Medicine, University of Hawai'i, Honolulu, Hawaii.

[4] Francis I. Proctor Foundation, University of California, San Francisco, San Francisco, California.

[5] Department of Medical Informatics and Clinical Epidemiology and Program of Computer Science and Electrical Engineering, Oregon Health & Science University, Portland, Oregon.

Disclosures:

All authors have completed and submitted the ICMJE disclosures form.

The authors made the following disclosures: J.P.C: Financial support − Genentech; Equity owner − Boston AI Labs, Inc. (this potential conflict of interest has been reviewed and managed by OHSU) Founder − Siloam Vision, LLC.

This work was supported by grants K12EY027720, P30EY10572, U10EY015114, and U10EY018573 from the National Institutes of Health (Bethesda, MD), funding from the Collins Medical Trust (Portland, OR), and by unrestricted departmental funding from Research to Prevent Blindness (New York, NY). The sponsors or funding organizations had no role in the design or conduct of this research.

HUMAN SUBJECTS: Human Subjects were used in this study. This study adhered to the tenets of the Declaration of Helsinki and was approved by the institutional review board at Oregon Health and Science University and the IEC and Aravind Eye Hospital in Madurai, India. Informed consent was obtained from all participants. No animal subjects were used in this study.

Author Contributions:

Conception and design: Hanif, Prajna, Lalitha,Keenan, Campbell, Song, Redd

Analysis and interpretation: Hanif, Redd

Data collection: Hanif, Prajna, Lalitha, Parker, Steinkamp, Redd

Obtained funding: Redd

Overall responsibility: Hanif, NaPier, Redd

Abbreviations and Acronyms:
AUPRC = area under the precision recall curve; **AUROC** = area under the receiver operating curve; **CI** = confidence interval; **CNN** = convolutional neural network; **DL** = deep learning.

Keywords:
Artificial intelligence, Deep learning, External photograph, Image quality, Infectious keratitis.

Correspondence:
Travis K. Redd, MD, MPH, 515 SW Campus Drive, Portland, OR 97239. E-mail: redd@ohsu.edu.

References

1. Flaxman SR, Bourne RRA, Resnikoff S, et al. Global causes of blindness and distance vision impairment 1990-2020: a systematic review and meta-analysis. *Lancet Glob Health*. 2017;5:e1221−e1234.
2. Cabrera-Aguas M, Khoo P, Watson SL. Infectious keratitis: a review. *Clin Exp Ophthalmol*. 2022;50:543−562.
3. Stapleton F. The epidemiology of infectious keratitis. *Ocul Surf*. 2021:S1542-0124(21)00089-6.
4. Ung L, Bispo PJM, Shanbhag SS, et al. The persistent dilemma of microbial keratitis: global burden, diagnosis, and antimicrobial resistance. *Surv Ophthalmol*. 2019;64:255−271.
5. Asaoka R, Murata H, Iwase A, Araie M. Detecting Preperimetric glaucoma with standard automated Perimetry using a deep learning classifier. *Ophthalmology*. 2016;123:1974−1980.
6. Brown JM, Campbell JP, Beers A, et al. Automated diagnosis of Plus disease in retinopathy of Prematurity using deep convolutional neural networks. *JAMA Ophthalmol*. 2018;136:803−810.
7. Lee CS, Baughman DM, Lee AY. Deep learning is effective for the classification of OCT images of normal versus Age-related Macular Degeneration. *Ophthalmol Retina*. 2017;1:322−327.
8. Kuo M-T, Hsu BW-Y, Yin Y-K, et al. A deep learning approach in diagnosing fungal keratitis based on corneal photographs. *Sci Rep*. 2020;10:14424.
9. Hung N, Shih AK-Y, Lin C, et al. Using slit-lamp images for deep learning-based identification of bacterial and fungal keratitis: model development and validation with different convolutional neural networks. *Diagnostics*. 2021;11:1246.
10. Ghosh AK, Thammasudjarit R, Jongkhajornpong P, et al. Deep learning for discrimination between fungal keratitis and bacterial keratitis: DeepKeratitis. *Cornea*. 2022;41:616−622.
11. Sim PY, La CJ, Than J, Ho J. National survey of the management of eye emergencies in the accident and emergency department by foundation doctors: has anything changed over the past 15 years? *Eye*. 2020;34:1094−1099.
12. Tauber J, Ayoub S, Shah P, et al. Assessing the Demand for Teleophthalmology in Florida emergency departments. *Telemed J E Health*. 2020;26:1500−1506.
13. Jan Bond C, Hao Chi H, Nor Fariza N, Elias H. Diy - smartphone slit-lamp adaptor. *J Mob Technol Med*. 2014;3:16−22.
14. Redd TK, Prajna NV, Srinivasan M, et al. Image-based differentiation of bacterial and fungal keratitis using deep convolutional neural networks. *Ophthalmol Sci*. 2022;2:100119.
15. Nadeem MW, Goh HG, Hussain M, et al. Deep learning for diabetic retinopathy analysis: a review, research challenges, and future directions. *Sensors*. 2022;22:6780.
16. Li Z, Jiang J, Qiang W, et al. Comparison of deep learning systems and cornea specialists in detecting corneal diseases from low-quality images. *iScience*. 2021;24:103317.
17. Joseph N, Kolluru C, Benetz BAM, et al. Quantitative and qualitative evaluation of deep learning automatic segmentations of corneal endothelial cell images of reduced image quality obtained following cornea transplant. *J Med Imaging*. 2020;7:014503.
18. Dalmon C, Porco TC, Lietman TM, et al. The clinical differentiation of bacterial and fungal keratitis: a photographic survey. *Invest Ophthalmol Vis Sci*. 2012;53:1787−1791.

19. McLeod SD, Kolahdouz-Isfahani A, Rostamian K, et al. The role of smears, cultures, and antibiotic sensitivity testing in the management of suspected infectious keratitis. *Ophthalmology*. 1996;103:23−28.
20. Grandini M, Bagli E, Visani G. Metrics for multi-class classification: an Overview. *arXiv [statML]*. 2020:2008.05756.
21. Vovk V. The fundamental nature of the log loss function. In: Beklemishev LD, Blass A, Dershowitz N, et al., eds. *Fields of Logic and Computation II: Essays Dedicated to Yuri Gurevich on the Occasion of His 75th Birthday*. Cham: Springer International Publishing; 2015:307−318.
22. Selvaraju RR, Cogswell M, Das A, et al. Grad-CAM: visual explanations from deep networks via Gradient-based localization, *arXiv [csCV]*, 2016;1610.02391:618−626.
23. Redd TK, Prajna NV, Srinivasan M, et al. Expert performance in visual differentiation of bacterial and fungal keratitis. *Ophthalmology*. 2022;129:227−230.
24. Srinivasan M. Fungal keratitis. *Curr Opin Ophthalmol*. 2004;15:321−327.
25. Thomas PA, Leck AK, Myatt M. Characteristic clinical features as an aid to the diagnosis of suppurative keratitis caused by filamentous fungi. *Br J Ophthalmol*. 2005;89:1554−1558.
26. Korot E, Pontikos N, Liu X, et al. Predicting sex from retinal fundus photographs using automated deep learning. *Sci Rep*. 2021;11:10286.
27. Poplin R, Varadarajan AV, Blumer K, et al. Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning. *Nat Biomed Eng*. 2018;2:158−164.
28. Wisely CE, Wang D, Henao R, et al. Convolutional neural network to identify symptomatic Alzheimer's disease using multimodal retinal imaging. *Br J Ophthalmol*. 2022;106:388−395.
29. Varaprasathan G, Miller K, Lietman T, et al. Trends in the etiology of infectious corneal ulcers at the F. I. Proctor Foundation. *Cornea*. 2004;23:360−364.
30. Xu Y, Kong M, Xie W, et al. Deep Sequential feature learning in clinical image classification of infectious keratitis. *Engineering*. 2021;7:1002−1010.