

A Convolutional Neural Network Using Anterior Segment Photos for Infectious Keratitis Identification

Vannarut Satitpitakul^{1,2}, Apiwit Puangsrucharern³, Surachet Yuktiratna³, Yossapon Jaisarn⁴, Keeratika Sangsao⁴, Vilavun Puangsrucharern^{1,2}, Ngamjit Kasetsuwan^{1,2}, Usanee Reinprayoon^{1,2}, Thanachaporn Kittipibul^{1,2}

¹Center of Excellence for Cornea and Stem Cell Transplantation, Department of Ophthalmology, Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand; ²Excellence Center for Cornea and Stem Cell Transplantation, Department of Ophthalmology, King Chulalongkorn Memorial Hospital, Thai Red Cross Society, Bangkok, Thailand; ³IM Impower Company Limited, Bangkok, Thailand; ⁴Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand

Correspondence: Vannarut Satitpitakul, Department of Ophthalmology, King Chulalongkorn Memorial Hospital, 1873 Rama 4 Road, Pathumwan, Bangkok, 10330, Thailand, Tel +66-894959022, Email Vannarut.S@chula.ac.th

Purpose: To develop a comprehensively deep learning algorithm to differentiate between bacterial keratitis, fungal keratitis, non-infectious corneal lesions, and normal corneas.

Methods: This retrospective study collected slit-lamp photos of patients with bacterial keratitis, fungal keratitis, non-infectious corneal lesions, and normal cornea. Causative organisms of infectious keratitis were identified by either positive culture or clinical response to single treatment. Convolutional neural networks (ResNet50, DenseNet121, VGG19) and Ensemble with probability weighting were used to develop a deep learning algorithm. The performance including accuracy, precision, recall, F1 score, specificity and AUC has been reported.

Results: Total of 6478 photos from 2171 eyes, composed of 2400 bacterial keratitis, 1616 fungal keratitis, 1545 non-infectious corneal lesions, and 917 normal corneas were collected from hospital database. DenseNet121 demonstrated the best performance among three convolutional neural networks with the accuracy of 0.8 (95% CI 0.74–0.86). The ensemble technique showed higher performance than single algorithm with the accuracy of 0.83 (95% 0.78–0.88).

Conclusion: Convolutional neural networks with ensemble techniques provided the best performance in discriminating bacterial keratitis, fungal keratitis, non-infectious corneal lesions, and normal corneas. Our models can be used as a screening tool for non-ophthalmic health care providers and ophthalmologists for rapid provisional diagnosis of infectious keratitis.

Keywords: infectious keratitis, cornea ulcer, keratitis, conventional neural network, deep learning algorithm

Introduction

Infectious keratitis is considered a potentially sight threatening ocular condition and a leading cause of visual disability worldwide. A lack of eye care resources leads to delays in diagnosis and treatment of infectious keratitis, which lead to loss of ocular integrity, invasion of infectious organisms and poor long-term visual prognosis,^{1,2} especially in low- and middle-income countries.

Differentiation between infectious keratitis, other corneal lesions and normal corneas based on slit-lamp photography can be done by ophthalmologist. However, it can be challenging for non-health care providers or even for primary health care providers. After detecting infectious keratitis, it is also crucial to differentiate the etiology of infectious keratitis, as the treatment for each organism is varied. Bacteria and fungi are majorly responsible for infectious keratitis worldwide.³ Clinical presentation of bacterial and fungal keratitis can be similar including pain, decreased vision, photophobia, redness, corneal infiltration, and corneal ulceration. Patients' history and some clinical presentations may help

differentiating the etiology of infection. However, diagnosing the etiology of infectious keratitis is still challenging. Microbiological testing such as corneal scraping for culture, a gold standard test, could help identify causative microorganism. Unfortunately, microbiological culture requires one to four weeks, and the rates of positive culture have been shown to be only 31–71%.^{4–8} Furthermore, emerging techniques like metagenomics DNA sequencing may be valuable for samples with negative microbiological cultures. However, these methods are not widely available in most hospitals and come with high costs.⁹

Recently, artificial intelligence based on deep learning algorithms has been shown to automatically detect the distinct features of disease when providing a large dataset of labeled medical images.^{10,11} Several studies^{12–16} have used deep learning algorithms for identifying etiology of infectious keratitis. In this study, we would like to evaluate if we could corroborate the findings of previous studies and expand the ability of algorithms to differentiate between infectious keratitis and non-infectious eyes using the same algorithms. We elected to use a large data set to train convolutional neural networks for differentiating bacterial keratitis, fungal keratitis, non-infectious corneal lesions, and normal corneas.

Materials and Methods

This study was approved by Institutional Review Board, Faculty of Medicine, Chulalongkorn University and adhered to the tenets of the Declaration of Helsinki. The study received exemption from informed consent by the Institutional Review Board of our institution. The deidentified retrospective data from 2007 to 2022 has been approved for review by the King Chulalongkorn Memorial Hospital. All data were accessed in April 2023.

Dataset and Labelling

The diffused illumination slit-lamp images taken using the Topcon D series slit lamp and DC-4 digital camera attachment were collected. All the collected photos were taken by corneal fellows using the same setting with the same slit-lamp biomicroscope and stored in an in-patient clinic database. The device functioning is driven by the EZ Capture software. The resolution of each image was 2576×1934 pixels in JPG format. The EZ Capture dataset was divided into 4 categories including 1) bacterial keratitis, 2) fungal keratitis, 3) non-infectious corneal lesions and 4) normal corneas. One hundred fifty-two non-focus images were screened and excluded by an ophthalmologist (V.S). All the images from patients with infectious keratitis (bacterial and fungal keratitis) were included during the active phase of infection and were confirmed by positive microbiological results including culture, PCR, or pathological reports or complete response to a single treatment using either antibacterial or antifungal medications. Patients were excluded if mixed infections were reported by laboratory tests. The categories of non-infectious keratitis were taken from patients experiencing other corneal diseases with or without inflammation including peripheral ulcerative keratitis, corneal degeneration, corneal deposits, corneal tumors, bullous keratopathy, corneal scars, fibrous ingrowth and limbal stem cell deficiency. We had taken photos of all patients and all photos, which were eligible have been included in the study. Representative images in each category were shown in Figure 1.

The final image dataset in this study included 6478 diffuse illumination slit-lamp images including 2400 images of bacterial keratitis taken from 183 patients, 1616 images of fungal keratitis taken from 89 patients, 1545 images of non-infectious corneal lesions taken from 1545 patients and 917 images of normal corneas taken from 917 participants. There were 125 patients with bacterial keratitis (45.8%) and 74 patients with fungal keratitis (27.3%) diagnosed based on microbiological/pathological results. Fifty-eight (21.4%) and 15 (5.5%) were diagnosed as bacterial and fungal keratitis due to complete response to a single treatment, either antibacterial or antifungal medications, respectively.

To ensure the model is not biased towards recognizing patients rather than the categories, the data were randomly distributed into training, validating, and testing set based on patients with a targeted ratio of 70:10:20, respectively. The training set consisted of 1609 images of bacterial keratitis, 1274 images of fungal keratitis, 1089 images of non-infectious corneal lesions and 666 of normal corneas. The validating set consists of 196 images of bacterial keratitis, 120 images of fungal keratitis, 147 images of non-infectious corneal lesions and 70 of normal corneas and the testing set consisted of 595 images of bacterial keratitis, 222 images of fungal keratitis, 309 images of non-infectious corneal lesions and 181 of normal corneas.

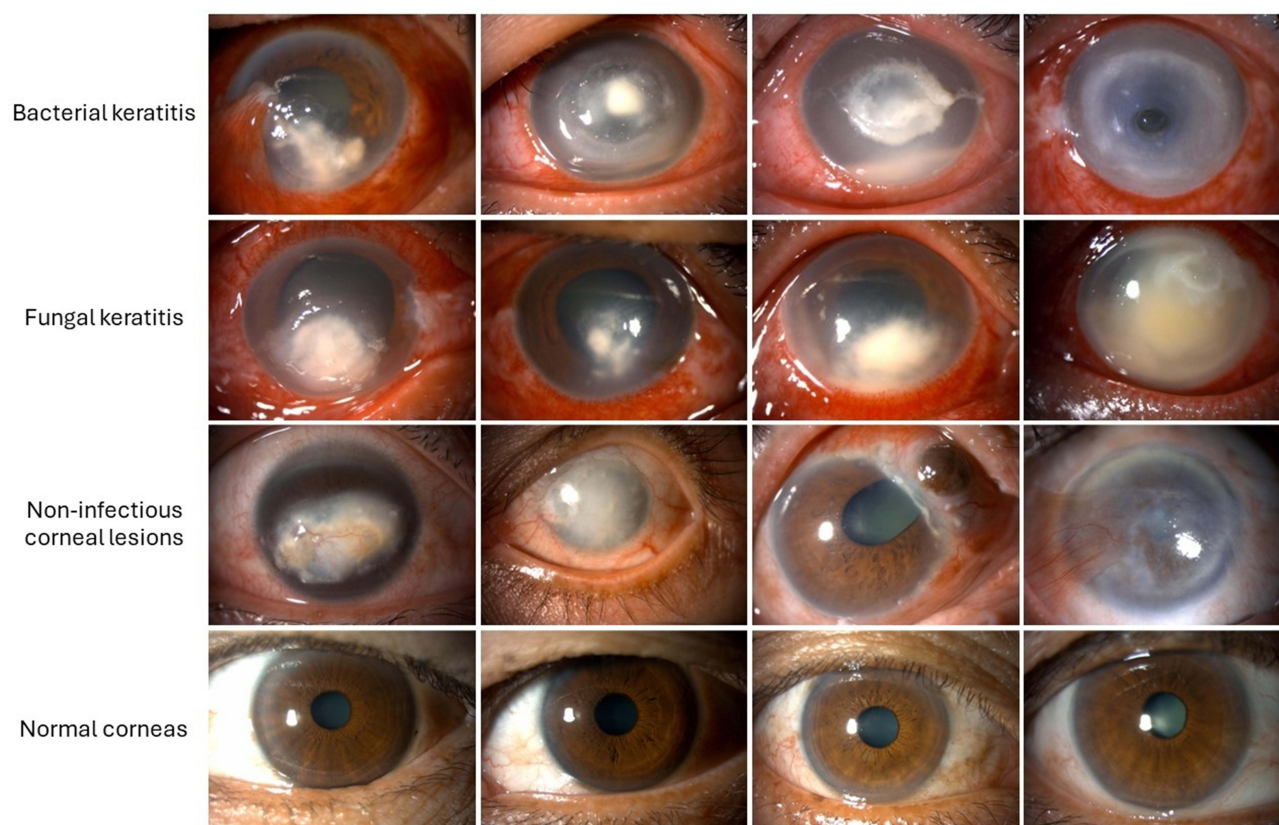


Figure 1 The representative slit-lamp images of four categories.

Model Selection and Development

In this study, we developed two deep learning models. First, we developed a deep learning model for automated classification of the ImageNet dataset into three categories including bacterial keratitis, fungal keratitis, and others (non-infectious corneal lesions and normal corneas). Second, we developed a deep learning model for automated classification of the ImageNet dataset into four categories including bacterial keratitis, fungal keratitis, non-infectious corneal lesions, and normal corneas.

We selected ResNet50, DenseNet121, and VGG19 as our neural network architectures to classify the ImageNet dataset. These models were chosen based on a thorough literature review, which revealed that they are among the most widely used and well-regarded architectures in the field. In addition, a probability weighting ensemble, which was designed to combine the predictions of multiple models was also included in the study. The probability weighting ensemble calculated the prediction by weighing the individual models' predictions based on their confidence levels. The weights assigned to each model in each class were determined based on their respective performance on the validation set.

The study framework encompassed the utilization of FastAI library, and extension of PyTorch, employing three neural network architectures (ResNet50, DenseNet121, and VGG19), initialized with pretrained parameters from the ImageNet dataset. The processing pipeline involves resizing the input images to $224 \times 224 \times 3$ pixels and employing a batch size of 64. The training was performed using the cross-entropy loss function and the Adam optimizer. The Adam optimizer is a gradient-based optimization algorithm. To assess the models' performance robustly, we adopt a 5-fold cross-validation strategy training each network for a maximum of 20 epochs. Early stopping was employed based on the minimum validation loss, with a patience of 5 epochs. This means that if the validation loss did not decrease for 5 consecutive epochs, the training process was stopped, and the model with the lowest validation loss was selected, with learning rates determined using FastAI's learning rate finder. The experiments are conducted on an NVIDIA GeForce RTX 4090 GPU equipped with 24GB of VRAM.

The accuracy, precision (positive predictive value), recall (sensitivity), F1 score, and specificity were shown as the indexes for the validity of the deep learning model for classification of bacterial keratitis, fungal keratitis, non-infectious

corneal lesion and normal corneas. F1 score is the harmonic weight average of precision and recall, and reflects a model's performance. F1 score (macro) is used as a main metric for measuring our model performance where score 1 indicates the best value and 0 indicates the worst value. Model interpretation using Gradient-weighted Class Activation Mapping (Grad-CAM) was applied to produce visual explanations, highlighting the areas of the image that had the greatest influence on the model's prediction.

Results

The Performance of Deep Learning Models to Differentiate Bacterial Keratitis, Fungal Keratitis, and Others

Among the three neural network architectures, the DenseNet 121 algorithm provided the best performance in classification bacterial keratitis, fungal keratitis, and others. The best algorithm achieved an accuracy of 0.74 (95% CI, 0.69–0.79), a precision of 0.61 (95% CI, 0.51–0.72), a recall of 0.61 (95% CI, 0.48–0.73), a F1 score of 0.59 (95% CI, 0.49–0.69), a specificity of 0.81 (95% CI, 0.75–0.87) and an AUC of 0.71 (95% CI, 0.65–0.77). Using the probability weighting ensemble technique, the models achieved higher performance with an accuracy of 0.77 (95% CI, 0.72–0.82), a precision of 0.65 (95% CI, 0.54–0.75), a recall of 0.63 (95% CI, 0.47–0.78), a F1 score of 0.60 (95% CI, 0.48–0.73), a specificity of 0.83 (95% CI, 0.75–0.90) and an AUC of 0.73 (95% CI, 0.66–0.79). The details on the performance of each algorithm are shown in Table 1.

The Performance of Deep Learning Models to Differentiate Bacterial Keratitis, Fungal Keratitis, Non-Infectious Corneal Lesions, and Normal Corneas

The DenseNet 121 algorithm still demonstrated the best performance in classifying bacterial keratitis, fungal keratitis, non-infectious corneal lesions, and normal corneas. The DenseNet 121 showed an accuracy of 0.80 (95% CI, 0.74–0.86),

Table 1 The Performance of Deep Learning Models to Differentiate Bacterial Keratitis, Fungal Keratitis, and Others

| Category | Model | Classification Performance (95% CI) | | | | | |
|----------|--------------|-------------------------------------|---------------------|---------------------|---------------------|---------------------|---------------------|
| | | Accuracy | Precision | Recall | F1 Score | Specificity | AUC |
| Bacteria | Resnet 50 | 0.65 (0.61–0.69) | 0.55 (0.38–0.71) | 0.59 (0.31–0.87) | 0.55 (0.44–0.65) | 0.69 (0.47–0.91) | 0.64 (0.59–0.69) |
| | DenseNet 121 | 0.67 (0.62–0.71) | 0.58 (0.47–0.70) | 0.45 (0.22–0.68) | 0.48 (0.35–0.61) | 0.8 (0.68–0.93) | 0.62 (0.57–0.68) |
| | VGG 19 | 0.64 (0.61–0.68) | 0.52 (0.40–0.65) | 0.77 (0.61–0.94) | 0.61 (0.59–0.64) | 0.58 (0.41–0.75) | 0.68 (0.67–0.68) |
| | Ensemble | 0.69 (0.66–0.72) | 0.59 (0.44–0.74) | 0.68 (0.46–0.90) | 0.61 (0.55–0.67) | 0.7 (0.52–0.88) | 0.69 (0.66–0.72) |
| Fungus | Resnet 50 | 0.74 (0.68–0.80) | 0.56 (0.32–0.80) | 0.23 (0.00–0.49) | 0.25 (0.04–0.47) | 0.92 (0.80–1.00) | 0.57 (0.50–0.65) |
| | DenseNet 121 | 0.7 (0.63–0.78) | 0.43 (0.30–0.57) | 0.55 (0.35–0.74) | 0.46 (0.37–0.55) | 0.76 (0.61–0.90) | 0.65 (0.61–0.68) |
| | VGG 19 | 0.75 (0.70–0.80) | 0.46 (0.26–0.66) | 0.28 (0.09–0.47) | 0.33 (0.14–0.52) | 0.9 (0.83–0.97) | 0.59 (0.53–0.66) |
| | Ensemble | 0.76 (0.71–0.80) | 0.52 (0.32–0.71) | 0.34 (0.08–0.61) | 0.37 (0.16–0.58) | 0.9 (0.79–1.00) | 0.62 (0.53–0.70) |

(Continued)

Table 1 (Continued).

| Category | Model | Classification Performance (95% CI) | | | | | |
|----------------|--------------|-------------------------------------|---------------------|---------------------|---------------------|---------------------|---------------------|
| | | Accuracy | Precision | Recall | F1 Score | Specificity | AUC |
| Others | Resnet 50 | 0.78 (0.68–0.89) | 0.73 (0.51–0.96) | 0.83 (0.61–1.00) | 0.75 (0.67–0.83) | 0.77 (0.51–1.00) | 0.8 (0.72–0.87) |
| | DenseNet 121 | 0.85 (0.82–0.89) | 0.82 (0.72–0.92) | 0.83 (0.68–0.98) | 0.81 (0.77–0.86) | 0.87 (0.77–0.97) | 0.85 (0.81–0.90) |
| | VGG 19 | 0.84 (0.79–0.89) | 0.88 (0.81–0.95) | 0.69 (0.51–0.87) | 0.76 (0.68–0.85) | 0.94 (0.88–0.99) | 0.81 (0.75–0.88) |
| | Ensemble | 0.87 (0.82–0.92) | 0.84 (0.70–0.98) | 0.85 (0.69–1.00) | 0.84 (0.78–0.89) | 0.89 (0.78–1.00) | 0.87 (0.82–0.93) |
| Average(Macro) | Resnet 50 | 0.72 (0.68–0.77) | 0.61 (0.51–0.71) | 0.55 (0.38–0.73) | 0.52 (0.39–0.65) | 0.79 (0.69–0.89) | 0.67 (0.61–0.73) |
| | DenseNet 121 | 0.74 (0.69–0.79) | 0.61 (0.51–0.72) | 0.61 (0.48–0.73) | 0.59 (0.49–0.69) | 0.81 (0.75–0.87) | 0.71 (0.65–0.77) |
| | VGG 19 | 0.74 (0.69–0.79) | 0.62 (0.50–0.74) | 0.58 (0.44–0.73) | 0.57 (0.45–0.68) | 0.8 (0.70–0.91) | 0.69 (0.64–0.75) |
| | Ensemble | 0.77 (0.72–0.82) | 0.65 (0.54–0.75) | 0.63 (0.47–0.78) | 0.6 (0.48–0.73) | 0.83 (0.75–0.90) | 0.73 (0.66–0.79) |

Notes: Accuracy = the number of correctly classified data over the total number of data; Precision = positive predictive value; Recall = sensitivity and F1 score = harmonic mean of precision and recall.

a precision of 0.64 (95% CI, 0.56–0.73), a recall of 0.66 (95% CI, 0.55–0.76), a F1 score of 0.63 (95% CI, 0.55–0.71), a specificity of 0.86 (95% CI, 0.81–0.91) and an AUC of 0.76 (95% CI, 0.69–0.82). The ensemble technique presented higher performance compared to each single algorithm. The ensemble technique achieved an accuracy of 0.83 (95% CI, 0.78–0.88), a precision of 0.70 (95% CI, 0.61–0.78), a recall of 0.68 (95% CI, 0.57–0.79), a F1 score of 0.68 (95% CI, 0.59–0.76), a specificity of 0.88 (95% CI, 0.82–0.93) and an AUC of 0.78 (95% CI, 0.72–0.84). The details on the performance of each algorithm are shown in [Table 2](#).

Table 2 The Performance of Deep Learning Models to Differentiate Bacterial Keratitis, Fungal Keratitis, Non-Infectious Corneal Lesions (NCL), and Normal Cornea

| Category | Model | Classification Performance (95% CI) | | | | | |
|----------|--------------|-------------------------------------|---------------------|---------------------|---------------------|---------------------|---------------------|
| | | Accuracy | Precision | Recall | F1 Score | Specificity | AUC |
| Bacteria | Resnet 50 | 0.59 (0.51–0.66) | 0.51 (0.41–0.60) | 0.75 (0.60–0.90) | 0.59 (0.55–0.63) | 0.49 (0.30–0.68) | 0.65 (0.59–0.70) |
| | DenseNet 121 | 0.68 (0.64–0.72) | 0.59 (0.52–0.66) | 0.43 (0.26–0.59) | 0.49 (0.39–0.58) | 0.82 (0.74–0.90) | 0.63 (0.58–0.67) |
| | VGG 19 | 0.66 (0.60–0.72) | 0.54 (0.43–0.64) | 0.72 (0.58–0.87) | 0.61 (0.54–0.68) | 0.62 (0.45–0.79) | 0.67 (0.64–0.71) |
| | Ensemble | 0.7 (0.66–0.74) | 0.59 (0.47–0.70) | 0.69 (0.52–0.86) | 0.62 (0.56–0.69) | 0.71 (0.56–0.87) | 0.7 (0.67–0.73) |

(Continued)

Table 2 (Continued).

| Category | Model | Classification Performance (95% CI) | | | | | |
|-----------------|--------------|-------------------------------------|---------------------|---------------------|---------------------|---------------------|---------------------|
| | | Accuracy | Precision | Recall | F1 Score | Specificity | AUC |
| Fungus | Resnet 50 | 0.66 (0.64–0.68) | 0.55 (0.42–0.67) | 0.42 (0.19–0.64) | 0.44 (0.31–0.57) | 0.79 (0.66–0.93) | 0.56 (0.49–0.64) |
| | DenseNet 121 | 0.7 (0.65–0.76) | 0.43 (0.27–0.59) | 0.6 (0.46–0.75) | 0.49 (0.36–0.62) | 0.74 (0.63–0.84) | 0.67 (0.62–0.72) |
| | VGG 19 | 0.75 (0.70–0.81) | 0.45 (0.32–0.59) | 0.38 (0.08–0.68) | 0.39 (0.16–0.63) | 0.86 (0.75–0.97) | 0.62 (0.53–0.72) |
| | Ensemble | 0.77 (0.74–0.81) | 0.52 (0.38–0.66) | 0.41 (0.24–0.58) | 0.45 (0.29–0.62) | 0.88 (0.83–0.93) | 0.64 (0.58–0.71) |
| NCL | Resnet 50 | 0.85 (0.83–0.87) | 0.55 (0.48–0.62) | 0.24 (0.11–0.38) | 0.33 (0.20–0.45) | 0.96 (0.95–0.98) | 0.74 (0.66–0.82) |
| | DenseNet 121 | 0.85 (0.81–0.88) | 0.74 (0.57–0.91) | 0.62 (0.44–0.80) | 0.66 (0.58–0.73) | 0.92 (0.84–1.00) | 0.77 (0.71–0.83) |
| | VGG 19 | 0.82 (0.77–0.86) | 0.66 (0.56–0.75) | 0.57 (0.37–0.77) | 0.59 (0.52–0.67) | 0.9 (0.82–0.97) | 0.73 (0.66–0.81) |
| | Ensemble | 0.86 (0.82–0.90) | 0.76 (0.61–0.92) | 0.7 (0.46–0.93) | 0.7 (0.60–0.80) | 0.92 (0.84–0.99) | 0.81 (0.72–0.90) |
| Normal | Resnet 50 | 0.95 (0.93–0.97) | 0.94 (0.92–0.97) | 0.48 (0.26–0.71) | 0.62 (0.40–0.83) | 1 (0.99–1.00) | 0.87 (0.77–0.98) |
| | DenseNet 121 | 0.96 (0.95–0.98) | 0.82 (0.74–0.91) | 0.97 (0.94–1.00) | 0.89 (0.84–0.93) | 0.97 (0.95–0.99) | 0.97 (0.95–0.98) |
| | VGG 19 | 0.9 (0.85–0.95) | 0.92 (0.79–1.00) | 0.47 (0.32–0.62) | 0.61 (0.48–0.74) | 0.99 (0.95–1.00) | 0.73 (0.65–0.80) |
| | Ensemble | 0.98 (0.97–0.98) | 0.92 (0.89–0.96) | 0.93 (0.89–0.97) | 0.93 (0.91–0.94) | 0.99 (0.98–1.00) | 0.96 (0.94–0.98) |
| Average (Macro) | Resnet 50 | 0.76 (0.69–0.83) | 0.64 (0.55–0.73) | 0.47 (0.36–0.58) | 0.49 (0.42–0.57) | 0.81 (0.71–0.92) | 0.71 (0.64–0.77) |
| | DenseNet 121 | 0.8 (0.74–0.86) | 0.64 (0.56–0.73) | 0.66 (0.55–0.76) | 0.63 (0.55–0.71) | 0.86 (0.81–0.91) | 0.76 (0.69–0.82) |
| | VGG 19 | 0.79 (0.74–0.84) | 0.65 (0.55–0.75) | 0.54 (0.45–0.64) | 0.56 (0.49–0.63) | 0.84 (0.77–0.92) | 0.69 (0.66–0.73) |
| | Ensemble | 0.83 (0.78–0.88) | 0.7 (0.61–0.78) | 0.68 (0.57–0.79) | 0.68 (0.59–0.76) | 0.88 (0.82–0.93) | 0.78 (0.72–0.84) |

Notes: Accuracy = the number of correctly classified data over the total number of data; Precision = positive predictive value; Recall = sensitivity and F1 score = harmonic mean of precision and recall.

Heatmaps

At the end of the convolutional neural network, we created a heatmap using Grad-CAM to visualize the regions of the corneal lesions that most influenced the model's prediction (Figure 2).

The heatmaps highlight the areas of the input image that had the highest contribution to the model's decision based on the activations of the final convolutional layer. Red regions indicate areas that strongly influenced the prediction, while blue regions signify areas with less influence. This visualization helps to better understand which parts of the image the model relied on for its classification.

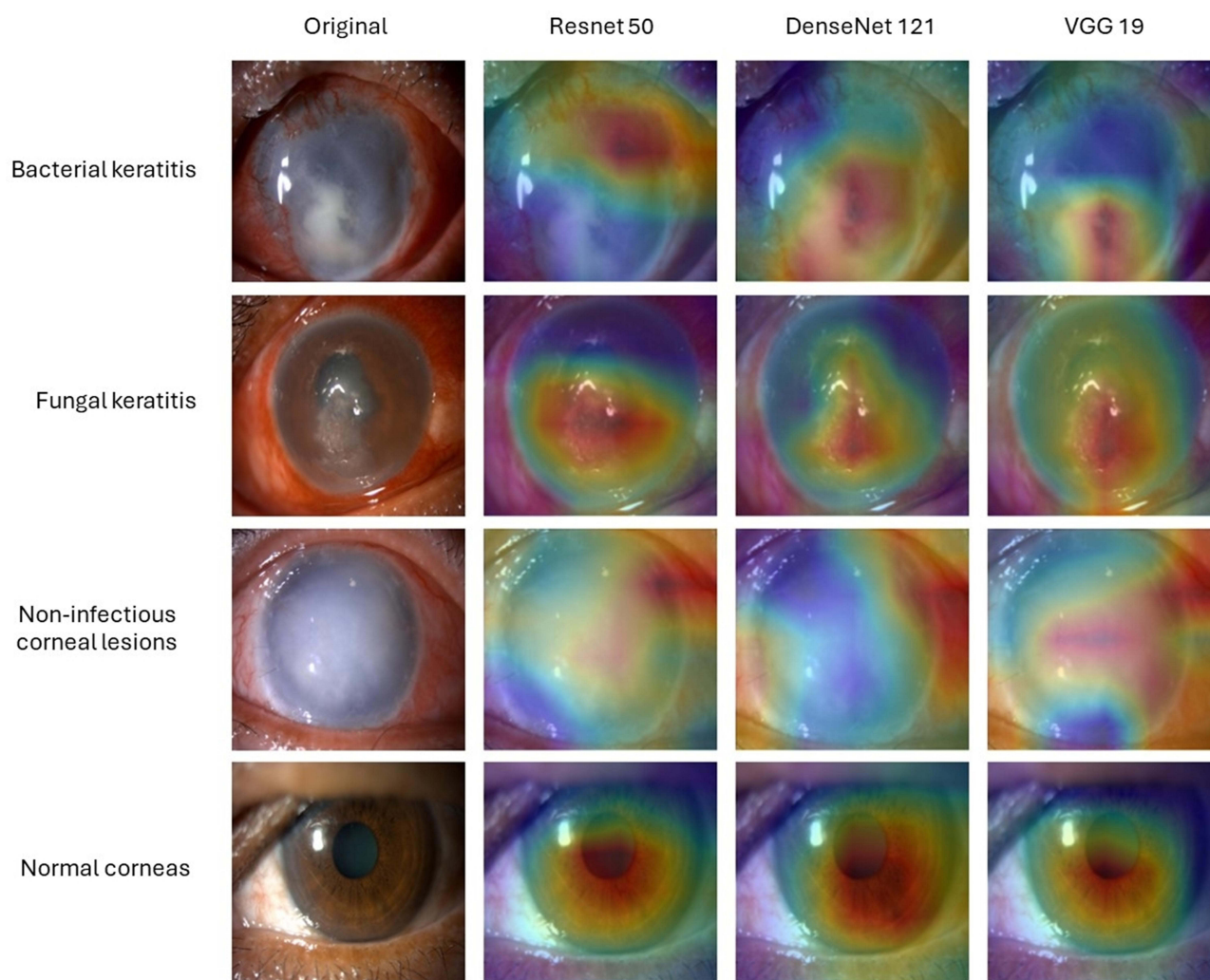


Figure 2 Heat maps.

Discussion

In this study, we evaluated the performance of deep learning system with additional of the ensemble techniques to classify bacterial keratitis, fungal keratitis, other corneal lesions, and normal corneas. Our findings showed that deep learning models could differentiate between infectious keratitis and non-infectious keratitis and also distinguish patients with bacterial and fungal keratitis using the same models. Separation of images with other corneal lesions and normal corneas before training the models helps obtaining better performance of the models. Among the three models, the DenseNet121 showed the best performance, with the accuracy of 80% followed by VGG19 79% and Resnet50 76%, respectively. The performance of the models increased with the probability weighting ensemble technique.

Rapid diagnosis of infectious keratitis is a crucial step to guide confirmative investigation and to initiate treatment.^{17,18} It also leads to better visual treatment outcomes. While differentiation between infectious and non-infectious corneas is challenging for non-ophthalmic health care workers, it can usually be done by ophthalmic health care workers. However, most ophthalmic health care workers are inexperienced in differentiation between bacterial and fungal keratitis, which is important for treatment initiation. The development of a deep learning system for differentiating infectious keratitis from other corneal lesions and normal corneas would benefit the health care system. In this study, we demonstrated the potential of developing a deep learning system for this purpose. However, the learning system still needs to improve both accuracy, precision and the system for clinical deployment in the future study.

In this study, we trained deep learning algorithms by replicating real-world conditions where patients with eye symptoms can have either infectious keratitis, non-infectious corneal lesions, or normal corneas with other eye problems. We included the group of normal corneas to train our algorithms to ensure that the algorithms can differentiate patients with eye symptoms without any corneal lesions. We found that the deep learning algorithms trained by combining images of other corneal lesions and normal corneas demonstrated the recall of 63% and the accuracy of 77% with ensemble technique, which is lower than the deep learning algorithm trained by separation between these two categories, which showed the recall of 68% and the accuracy of 83%. We believed that separation of the images of other corneal lesions and normal corneas replicated the real-world situation and lead to a more generalized learning process.

Providing a large data set, a deep learning algorithm can gain better performance compared with conventional statistical models and human performance.¹⁹ Our study used a large data set including 2400 images of bacterial keratitis, 1616 images of fungal keratitis, 1545 images of non-infectious corneal lesions and 917 images of normal corneas. While most of previous reports focused on either classification of infectious causes or differentiate infectious keratitis from others, we found that the accuracy of our model to classify bacterial and fungal keratitis and to differentiate between infectious, non-infectious corneas and normal corneas using the same model was 77–83% with ensemble technique. We believed that the accuracy of our model was comparable to the previous reports. The previous reports that aimed to classify causes of infectious keratitis demonstrated the accuracy of 51–90.7%,^{12,14,20,21} while the study of Li Z. et al aimed to differentiate keratitis, normal and others showed the accuracy of 99.8%.²² Soleimani M. et al demonstrated the model accuracy of 84% to classify fungal and bacterial keratitis and also developed the model to differentiate between keratitis caused by filamentous fungi and yeast. Previous reports^{12,20} also showed that the accuracy of the models was higher than that of ophthalmologists, with the accuracy ranged from 49% to 74%.

There are several limitations in our study. First, our models were trained using high resolution images from single-center data. The model performance may change when using different resolution images. However, our dataset has variable brightness, which should be suitable for real-world applications. Second, the amount of dataset in each classification did not reflect the real prevalence in each classification. This would directly affect the precision (positive predictive value) of the DL algorithm. However, lowering the number of cases in each class to the real prevalence, which is a lot higher rate of normal cornea cases, would lead to falsely high accuracy of the model. Third, some fungal keratitis cases have been reported to respond to antibiotics. Therefore, using a response to antibiotics for classification of bacterial keratitis may be misleading in some cases.²³ Lastly, due to bacterial and fungal keratitis were the most common cause of infectious keratitis in our center. Our deep learning algorithm was only trained using images of bacterial and fungal keratitis. Therefore, the algorithm cannot be applied to other etiologies of infectious keratitis such as viral keratitis and parasitic keratitis. A future study including the history of patients could be done to increase the performance of the model.

Conclusion

In conclusion, this is the deep learning algorithm developed to differentiate between bacterial keratitis, fungal keratitis, non-infectious corneal lesions, and normal corneas. This is the initial development of deep learning algorithms as a screening tool and a guide for initial treatment of infectious keratitis. Studies to improve accuracy, precision and the system for clinical deployment could be done in the future.

Funding

This research is funded by Thailand Science Research and Innovation Fund, Chulalongkorn University (CUFRB65_hea (48)_055_30_36). The funder has no role in study design, data collection, analysis, manuscript writing and publication.

Disclosure

The authors report no conflicts of interest in this work.

References

1. Cabrera-Aguas M, Khoo P, Watson S. Infectious keratitis: a review. *Clin Exp Ophthalmol*. 2022;50(5):543–562. doi:10.1111/ceo.14113
2. Bonzano C, Borroni D, Lancia A, et al. Doxycycline: from ocular rosacea to COVID-19 anosmia. New insight into the coronavirus outbreak. *Front Med Lausanne*. 2020;7:200. doi:10.3389/fmed.2020.00200
3. Flexman 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(12):e1221–e1234. doi:10.1016/S2214-109X(17)30393-5
4. Butler TKH, Spencer NA, Chan CCK, et al. Infective keratitis in older patients: a 4 year review, 1998–2002. *Br J Ophthalmol*. 2005;89(5):591–596. doi:10.1136/bjo.2004.049072
5. Keay L, Edwards K, Naduvilath T, et al. Microbial keratitis predisposing factors and morbidity. *Ophthalmology*. 2006;113(1):109–116. doi:10.1016/j.optha.2005.08.013
6. Gebauer A, McGhee CN, Crawford GJ. Severe microbial keratitis in temperate and tropical Western Australia. *Eye*. 1996;10(5):575–580. doi:10.1038/eye.1996.133
7. Leibovitch I, Lai TF, Senarath L, Selva D, Selva D. Infectious keratitis in south Australia: emerging resistance to cephazolin. *Eur J Ophthalmol*. 2005;15(1):23–26. doi:10.1177/112067210501500104
8. Reinprayoon U, Sitthanon S, Kasetsuwan N, Chongthaleong A. Bacteriological findings and antimicrobial susceptibility pattern of isolated pathogens from visual threatening ocular infections. *J Med Assoc Thai*. 2015;98(Suppl 1):S70–S76.
9. Borroni D, Bonzano C, Sanchez-Gonzalez J, et al. Shotgun metagenomic sequencing in culture negative microbial keratitis. *Eur J Ophthalmol*. 2023;33(4):1589–1595. doi:10.1177/11206721221149077
10. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015;521(7553):436–444. doi:10.1038/nature14539
11. Litjens G, Kooi T, Bejnordi BE, et al. A survey on deep learning in medical image analysis. *Med Image Anal*. 2017;42:60–88. doi:10.1016/j.media.2017.07.005
12. Kuo MT, Hsu BWY, Yin YK, et al. A deep learning approach in diagnosing fungal keratitis based on corneal photographs. *Sci Rep*. 2020;10(1):14424. doi:10.1038/s41598-020-71425-9
13. Hung N, Shih AKY, 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(7):1246. doi:10.3390/diagnostics11071246
14. Ghosh AK, Thammasudjarit R, Jongkhakornpong P, et al. Deep learning for discrimination between fungal keratitis and bacterial keratitis: deepKeratitis. *Cornea*. 2022;41(5):616–622. doi:10.1097/ICO.0000000000002830
15. Mayya V, Shevgoor SK, Kulkarni U, et al. Multi-scale convolutional neural network for accurate corneal segmentation in early detection of fungal keratitis. *J Fungi*. 2021;7(10):850. doi:10.3390/jof7100850
16. 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(2):100119. doi:10.1016/j.xops.2022.100119
17. Gupta N, Tandon R. Investigative modalities in infectious keratitis. *Indian J Ophthalmol*. 2008;56(3):209–213. doi:10.4103/0301-4738.40359
18. Labbe A, Khammari C, Dupas B, et al. Contribution of in vivo confocal microscopy to the diagnosis and management of infectious keratitis. *Ocul Surf*. 2009;7(1):41–52. doi:10.1016/S1542-0124(12)70291-4
19. Rajula HSR, Verlato G, Manchia M, et al. Comparison of conventional statistical methods with machine learning in medicine: diagnosis, drug development, and treatment. *Medicina*. 2020;56(9):455. doi:10.3390/medicina56090455
20. Xu Y, Kong M, Xie W, et al. Deep sequential feature learning in clinical image classification of infectious keratitis. *Engineering*. 2021;7(7):1002–1010. doi:10.1016/j.eng.2020.04.012
21. Saini JS, Jain AK, Kumar S, et al. Neural network approach to classify infective keratitis. *Curr Eye Res*. 2003;27(2):111–116. doi:10.1076/ceyr.27.2.111.15949
22. Li W, Yang Y, Zhang K, et al. Dense anatomical annotation of slit lamp images improves the performance of deep learning for the diagnosis of ophthalmic disorders. *Nat Biomed Eng*. 2020;4(8):767–777. doi:10.1038/s41551-020-0577-y
23. Matoba AY. Fungal keratitis responsive to moxifloxacin monotherapy. *Cornea*. 2012;31(10):1206–1209. doi:10.1097/ICO.0b013e31823f766c

Clinical Ophthalmology

Publish your work in this journal

Clinical Ophthalmology is an international, peer-reviewed journal covering all subspecialties within ophthalmology. Key topics include: Optometry; Visual science; Pharmacology and drug therapy in eye diseases; Basic Sciences; Primary and Secondary eye care; Patient Safety and Quality of Care Improvements. This journal is indexed on PubMed Central and CAS, and is the official journal of The Society of Clinical Ophthalmology (SCO). The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/clinical-ophthalmology-journal>

Dovepress
Taylor & Francis Group