



# Deep Learning Model and its Application for the Diagnosis of Exudative Pharyngitis

Seo Yi Chng<sup>1</sup>, Paul Jie Wen Tern<sup>2</sup>, Matthew Rui Xian Kan<sup>3</sup>, Lionel Tim-Ee Cheng<sup>4</sup>

<sup>1</sup>Department of Paediatrics, National University of Singapore, Singapore

<sup>2</sup>Department of Cardiology, National Heart Centre, Singapore

<sup>3</sup>NUS High School of Math and Science, Singapore

<sup>4</sup>Department of Diagnostic Radiology, Singapore General Hospital, Singapore

**Objectives:** Telemedicine is firmly established in the healthcare landscape of many countries. Acute respiratory infections are the most common reason for telemedicine consultations. A throat examination is important for diagnosing bacterial pharyngitis, but this is challenging for doctors during a telemedicine consultation. A solution could be for patients to upload images of their throat to a web application. This study aimed to develop a deep learning model for the automated diagnosis of exudative pharyngitis. Thereafter, the model will be deployed online. **Methods:** We used 343 throat images (139 with exudative pharyngitis and 204 without pharyngitis) in the study. ImageDataGenerator was used to augment the training data. The convolutional neural network models of MobileNetV3, ResNet50, and EfficientNetB0 were implemented to train the dataset, with hyperparameter tuning. **Results:** All three models were trained successfully; with successive epochs, the loss and training loss decreased, and accuracy and training accuracy increased. The EfficientNetB0 model achieved the highest accuracy (95.5%), compared to MobileNetV3 (82.1%) and ResNet50 (88.1%). The EfficientNetB0 model also achieved high precision (1.00), recall (0.89) and F1-score (0.94). **Conclusions:** We trained a deep learning model based on EfficientNetB0 that can diagnose exudative pharyngitis. Our model was able to achieve the highest accuracy, at 95.5%, out of all previous studies that used machine learning for the diagnosis of exudative pharyngitis. We have deployed the model on a web application that can be used to augment the doctor's diagnosis of exudative pharyngitis.

**Keywords:** Artificial Intelligence, Deep Learning, Diagnosis, Pharyngitis, Telemedicine

**Submitted:** September 27, 2023

**Revised:** December 16, 2023

**Accepted:** January 15, 2024

## Corresponding Author

Seo Yi Chng

Department of Paediatrics, National University of Singapore, 5 Lower Kent Ridge Road, Singapore 119074, Singapore. Tel: +65-96233755, E-mail: drchngsy@gmail.com (<https://orcid.org/0000-0001-8216-4022>)

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

© 2024 The Korean Society of Medical Informatics

## I. Introduction

Telemedicine refers to the assessment of health, diagnosis, or treatment by a medical practitioner over a distance through technology. The coronavirus disease 2019 pandemic exponentially increased the number of telemedicine consultations [1]. Post-pandemic, telemedicine continued as a modality of care [2], and telemedicine is now firmly established in the healthcare landscape in many countries including Europe and the United States [3]. There are more than 800 telemedicine service providers (clinics, hospitals) in Singapore listed on the Ministry of Health website, in a move by the Ministry to shift care to the home and community and decrease unnecessary visits to clinics and hospitals [4].

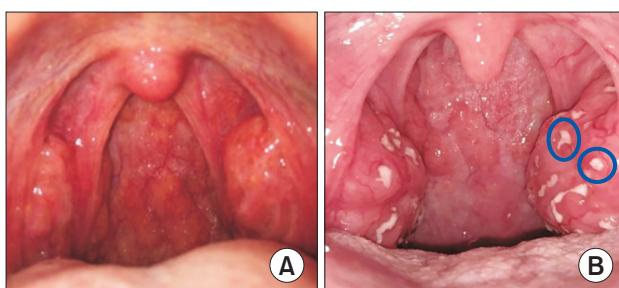
Telemedicine consultations offer numerous advantages over traditional face-to-face clinic consultations [3]. A recent review of more than 2,230 primary studies has clearly demonstrated the benefits of telemedicine in screening for, diagnosing, treating, managing, and providing long-term follow-up for chronic diseases [3].

Acute respiratory infections are a common reason for telemedicine consultations. A throat examination is important in patients with acute respiratory infections to exclude bacterial tonsillitis and pharyngitis, in particular, group A beta-haemolytic streptococcal infection [5]. If untreated, streptococcal pharyngitis can progress to life-threatening complications, such as peritonsillar abscess, retropharyngeal abscess, rheumatic fever, and rheumatic heart disease [5]. Figure 1 illustrates the appearance of a normal throat and a throat with exudative pharyngitis.

However, it is challenging for doctors to perform a proper throat examination during a telemedicine consultation [6]. A possible solution is for patients to take photographs of their throat with a smartphone and upload these images to a mobile or web application.

There is currently no mobile or web application for the automated diagnosis of pharyngitis. Such an application will have potential benefits for both medical doctors (if an automated diagnosis is available to doctors only and augments the doctors' diagnoses during telemedicine consultations) and patients (if automated diagnosis is available to patients, and this empowers patients to have knowledge of their own medical condition).

The aim of this project was to develop a deep learning model for the automated diagnosis of exudative pharyngitis. Thereafter, the model will be deployed on a web application with a graphical user interface.



**Figure 1.** (A) A normal pharynx. (B) A diseased pharynx with exudative pharyngitis, as can be seen by the white exudates on the tonsils. Some of the exudates have been circled for clearer illustration.

## II. Methods

### 1. Dataset

This study used an open-source dataset [7] of smartphone-based throat images. There were 361 images (146 images with exudative pharyngitis and 215 images without pharyngitis).

Throat images that included the uvula and both tonsillar beds were kept, while 5.0% of images captured only one side of the throat and were discarded. The final cleaned dataset contained 343 images (139 with pharyngitis and 204 without pharyngitis). Eighty percent of the images were randomly selected for training and 20% were kept for validation.

### 2. Image Augmentation

As the dataset was small, data augmentation techniques were applied to the training images for improved classification performance. Transformation techniques of horizontal flipping, dimming (by up to 25%) or brightening (by up to 25%), zooming out (by up to 25%) and zooming in (by up to 25%) were applied randomly to the training images in real-time using ImageDataGenerator in Keras.

### 3. Classification Models

Three convolutional neural network (CNN) models—MobileNetV3, ResNet50, and EfficientNetB0—were used in this image classification problem.

#### 1) MobileNetV3

Mobile models are built on increasingly efficient building blocks. MobileNetV1 [8] utilises depth-wise separable convolutions in place of traditional convolutional layers. Depth-wise separable convolutions separate spatial filtering from feature generation by employing two separate layers: lightweight depth-wise convolution for spatial filtering and heavier  $1 \times 1$  point-wise convolutions for feature generation. MobileNetV2 [9] implemented a linear bottleneck and inverted residual structure (such that the residual connections are between the bottleneck areas), to create even more efficient layer structures.

MobileNetV3 [10] is tuned to mobile phone CPUs through a combination of hardware-aware network architecture search complemented by the NetAdapt algorithm and then subsequently improved through novel architecture advances. MobileNetV3-Large achieves 3.2% higher accuracy on ImageNet with 15% less latency than MobileNetV2 [10]. Hence, MobileNetV3 is an efficient neural network for classification tasks on mobile devices.

## 2) ResNet50

ResNet50 [11] is a residual learning framework that is substantially deeper than previously used neural networks. The ResNet50 layers are reformatted to learn residual functions referenced to layer inputs instead of learning unreference functions [11]. These residual networks are easier to optimise and more accurate, due to their markedly increased depth [11].

On the ImageNet dataset, ResNet had a depth of up to 152 layers (eight times deeper than VGG nets) and yet was less complex [11]. An ensemble of residual nets outperformed all previous ensemble models including VGG v5 on the ImageNet dataset [11]. The ResNet ensemble placed first on the ILSVRC2015 (ImageNet Large Scale Visual Recognition Competition 2015) classification task. Thus, ResNet is a deep CNN that is accurate for image classification.

## 3) EfficientNetB0

Tan and Le [12] systematically studied model scaling and concluded that balancing the depth, width, and resolution of a CNN can result in superior performance. Their team designed a new scaling method that uniformly scales depth, width, and resolution with the aid of a compound coefficient and demonstrated its effectiveness in scaling up MobileNet and ResNet.

Tan and Le [12] further used a neural architecture search to design and scale a new baseline network to obtain a family of EfficientNet models. EfficientNetB0 is the baseline model in the EfficientNet family and achieved slightly higher accuracy on ImageNet at 77.1% than ResNet50 at 76.0%. EfficientNetB7 was able to achieve state-of-the-art accuracy (84.3%) on ImageNet, while remaining small and fast on inference compared to other networks. Thus, EfficientNet is able to achieve high accuracy in image classification. In addition, EfficientNet has the benefit of transferring well.

## 4. Classification

The images were pre-processed by each model. The shapes of the images and labels were ( $m$ , 224, 224, 3) and ( $m$ , 2), respectively, where  $m$  is the number of images. MobileNetV3, ResNet50, and EfficientNetB0 were trained from scratch. The fully connected layer at the top of the network was excluded. A dropout (0.4) layer was added to prevent overfitting. A final dense layer with softmax activation was then added. Adam optimiser was used for a faster result. The batch size was set at 64, and each model was trained for 200 epochs. WandB (weights and biases) was used to track the performance of the models.

## 5. Hyperparameter Tuning

The learning rate was tuned for each model, as the default learning rate (0.001) resulted in over-fitting for all three models. The following optimum learning rates were obtained for each model: MobileNetV3, 0.0003; ResNet50, 0.0001; EfficientNetB0, 0.00005.

## 6. Evaluation Metrics

The performance of the MobileNetV3, ResNet50, and EfficientNetB0 models in classifying exudative pharyngitis versus healthy throats was evaluated using standard machine learning metrics: accuracy, area under the receiver operating characteristic curve (AUC), precision, recall, and F1-score. Accuracy and AUC scores could be used as the data were balanced, with the proportion of the minority class in this study being more than 40%.

## 7. Development of a Graphical User Interface

Flask was used as the framework to develop an intuitive graphical user interface (GUI) for users to upload their throat images to a web application for automated diagnosis of exudative pharyngitis using our deep learning model.

The summary of the methodology used in this paper is illustrated in Figure 2.

## III. Results

### 1. Successful Training of All Three Models

Figure 3 illustrates the successful training of all three models.

### 2. Evaluation of Test Performance

Table 1 shows the accuracy, AUC score, precision, recall, and F1-score for the three models. The EfficientNetB0 model achieved the highest accuracy (95.5%), compared to the accuracy of 82.1% for MobileNetV3 and 88.1% for ResNet50. The EfficientNetB0 model also achieved the highest AUC (0.975), compared to the AUCs of 0.876 for MobileNetV3 and 0.954 for ResNet50. The EfficientNetB0 model was also able to achieve high precision (1.00), recall (0.89), and F1-score (0.94).

Figures 4 and 5 illustrate the receiver operating characteristic curve and the confusion matrix, respectively, for the EfficientNetB0 model.

### 3. Intuitive GUI

As illustrated in Figures 6 and 7, we developed an intuitive GUI for doctors and patients to upload throat images. The

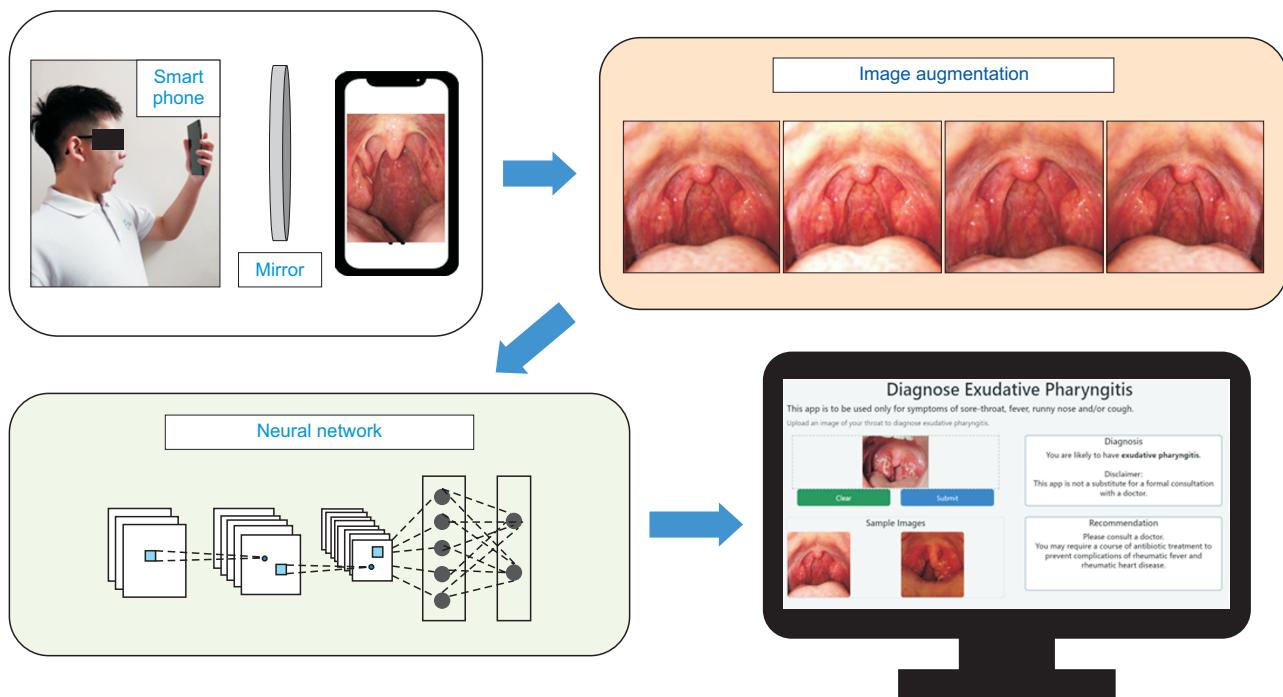


Figure 2. The methodology used in this paper for the automated diagnosis of exudative pharyngitis.

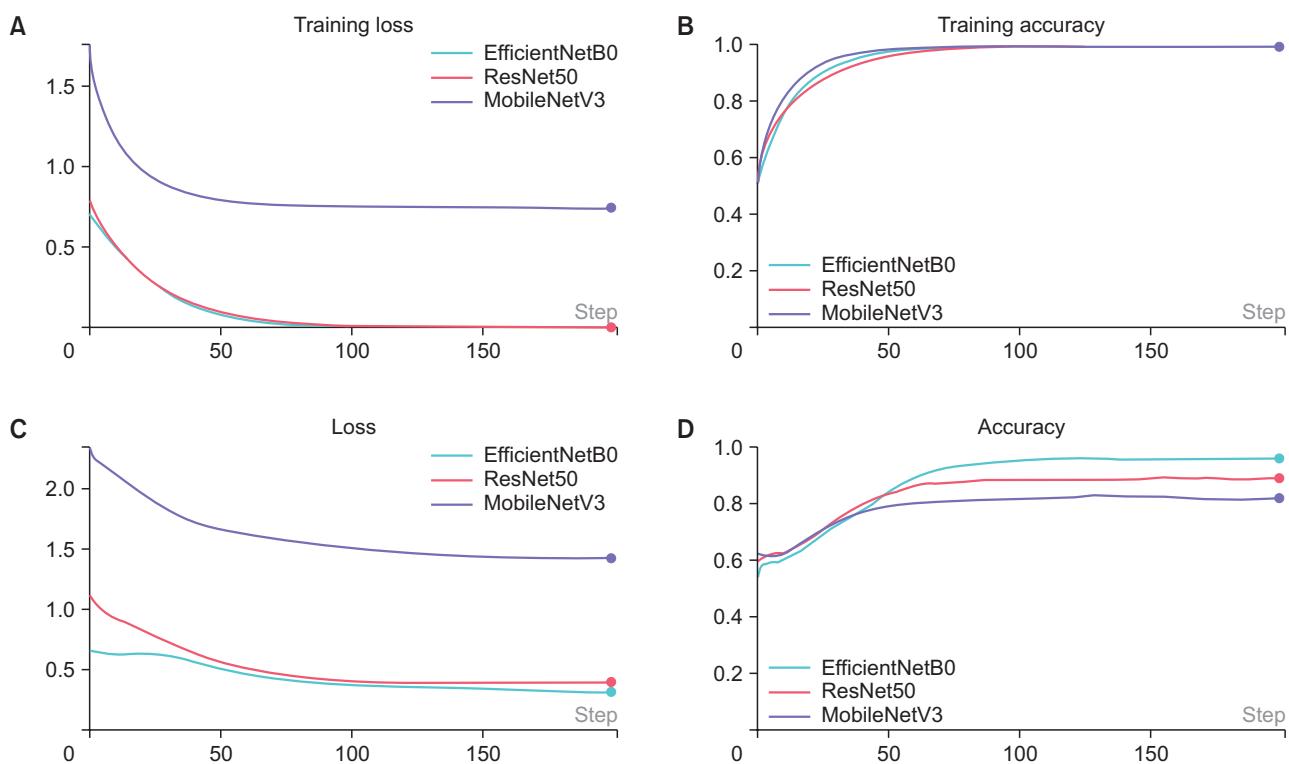


Figure 3. The training results for three models: (A) training loss, (B) training accuracy, (C) loss, and (D) accuracy. For all three models, the training loss decreased, and the training accuracy increased with successive epochs of training. The loss decreased and the accuracy increased with training.

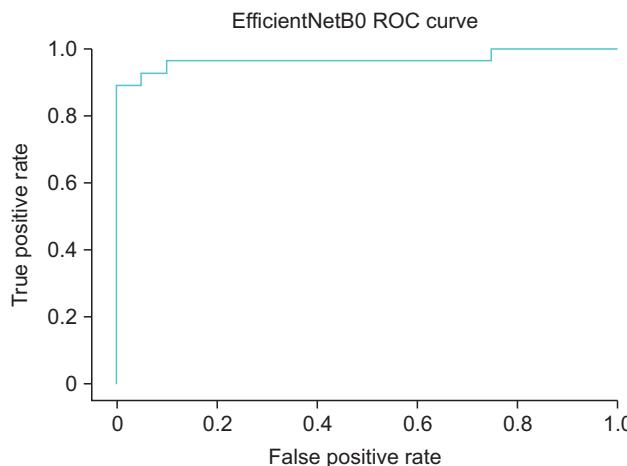
developed web application is deployed at <https://exudative-pharyngitis-v2.herokuapp.com> for the automated or augmented diagnosis of exudative pharyngitis using our Effi-

cientNetB0 model.

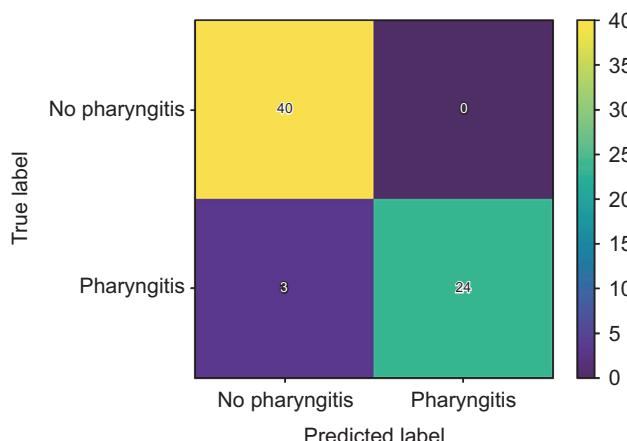
**Table 1.** Accuracy, AUC, precision, recall, and F1-scores for the three models

	Accuracy	AUC	Precision	Recall	F1-score
ResNet50	0.881	0.954	0.91	0.78	0.84
MobileNetV3	0.821	0.876	0.89	0.63	0.74
EfficientNetB0	<b>0.955</b>	<b>0.975</b>	<b>1.00</b>	<b>0.89</b>	<b>0.94</b>

The bolded results indicate the best performance for each metric amongst the 3 models. AUC: area under the receiver operating characteristic curve.



**Figure 4.** Receiver operating characteristic (ROC) curve for the EfficientNetB0 model. The area under the ROC curve was **0.975**.

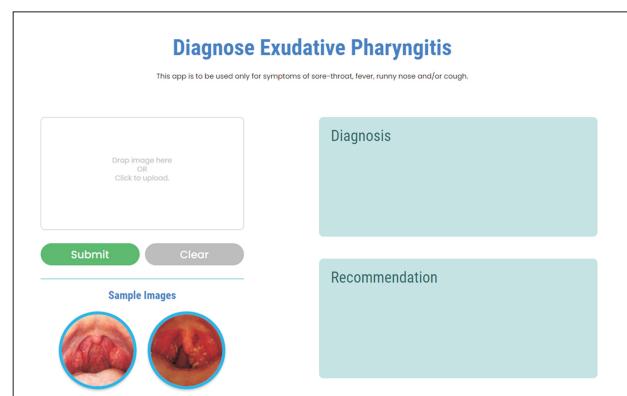


**Figure 5.** Confusion matrix for the EfficientNetB0 model. The EfficientNetB0 model achieved high precision (1.00), recall (0.89), and F1-score (0.94).

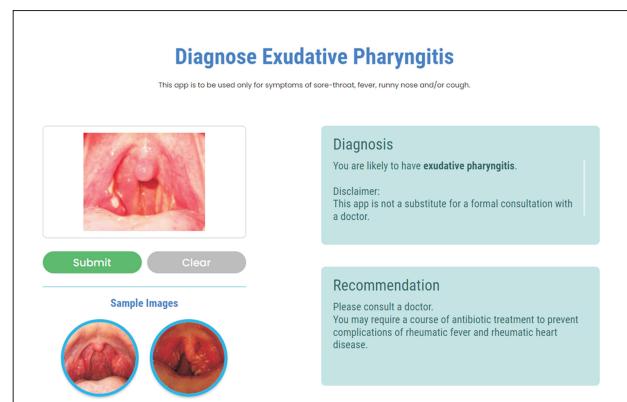
## IV. Discussion

### 1. Implications

We successfully trained an EfficientNetB0 model that can automatically diagnose exudative pharyngitis from non-pharyngitis. Compared to the two previous studies in the literature that also used machine learning for the automated



**Figure 6.** Screenshot of our web application for users (doctors or patients) to upload throat images for the automated diagnosis of exudative pharyngitis.



**Figure 7.** Screenshot of our web application that displays the correct diagnosis after a user has uploaded an image of his throat. In this case, the patient requires antibiotics to prevent complications such as rheumatic fever and rheumatic heart disease.

diagnosis of pharyngitis, this study achieved slightly superior accuracy, at 95.5%, compared to an accuracy of 95.3% by Yoo et al. [13] using an augmented ResNet50 model and an accuracy of 93.75% by Askarian et al. [14] using a k-nearest neighbour algorithm. However, the AUC score of 0.975 in our study is slightly lower than the AUC score of 0.992 achieved by Yoo et al. [13]. We also note that Yoo et al. [13] and Askarian et al. [14] did not deploy their models on

a web or mobile application for clinical use.

The precision in this study was very high (1.00) and the recall was high (0.89). The two previous studies by Yoo et al. [13] and Askarian et al. [14] did not examine precision and recall; thus, we were not able to compare our models' performance on these metrics with that of previous work. In our study where the precision was very high, at 1.00, the implication is that the model classified all non-pharyngitis patients correctly, thus avoiding the unnecessary use of antibiotics in patients who do not require them. The recall in this study was also high, at 0.89. This implies that the model can detect almost 9 in 10 patients who have exudative pharyngitis. These patients will be able to receive antibiotic treatment before the disease progresses to life-threatening complications.

We have deployed our model on a web application available at <https://exudative-pharyngitis-v2.herokuapp.com> that can be used by patients to upload their own throat image for automated diagnosis of exudative pharyngitis. The application is to be used by patients with symptoms of acute respiratory infection (fever, sore throat, running nose and/or cough). The application can also be used by doctors to upload their patient's throat image for augmented diagnosis of exudative pharyngitis. The application can be used as a second reader to aid the doctor in diagnosis. If the application is used to augment doctors' diagnosis (in conjunction with a doctor reading the image uploaded by the patient), the accuracy and recall will likely be much higher, to the benefit of the patient.

## 2. Future Work

With Institutional Review Board approval and patients' consent, the throat images uploaded can be used as additional training and testing data to further improve the model performance.

## 3. Conclusion

We trained a deep learning model based on EfficientNetB0 that can diagnose exudative pharyngitis. Our model was able to achieve the highest accuracy (95.5%) of all previous studies that used machine learning for the automated diagnosis of exudative pharyngitis. The model also achieved a very high precision of 1.00 and high recall of 0.89. As a proof of concept, we have deployed the model on a web application that can be used to augment the doctor's diagnosis of exudative pharyngitis during telemedicine consultations.

## Conflict of Interest

The following potential conflict of interest is declared: Dr. Seo Yi Chng is a director at Krsyma Medical AI Pte Ltd.

## ORCID

Seo Yi Chng (<https://orcid.org/0000-0001-8216-4022>)

Paul Jie Wen Tern (<https://orcid.org/0000-0001-5382-9881>)

Matthew Rui Xian Kan (<https://orcid.org/0000-0002-6688-3454>)

Lionel Tim-Ee Cheng (<https://orcid.org/0000-0002-1068-7868>)

## References

- Cortez C, Mansour O, Qato DM, Stafford RS, Alexander GC. Changes in short-term, long-term, and preventive care delivery in US office-based and telemedicine visits during the COVID-19 pandemic. *JAMA Health Forum* 2021;2(7):e211529. <https://doi.org/10.1001/jamahealthforum.2021.1529>
- Mandal S, Wiesenfeld BM, Mann D, Lawrence K, Chunara R, Testa P, et al. Evidence for telemedicine's ongoing transformation of health care delivery since the onset of COVID-19: retrospective observational study. *JMIR Form Res* 2022;6(10):e38661. <https://doi.org/10.2196/38661>
- Saigi-Rubio F, Borges do Nascimento IJ, Robles N, Ivanovska K, Katz C, Azzopardi-Muscat N, et al. The current status of telemedicine technology use across the World Health Organization European Region: an overview of systematic reviews. *J Med Internet Res* 2022;24(10):e40877. <https://doi.org/10.2196/40877>
- Ministry of Health. Licensing of Telemedicine Services under the Healthcare Services Act (HCSA) [Internet]. Singapore: Ministry of Health; 2023 [cited at 2024 Jan 24]. Available from: <https://www.moh.gov.sg/licensing-and-regulation/telemedicine>.
- Mustafa Z, Ghaffari M. Diagnostic methods, clinical guidelines, and antibiotic treatment for Group A Streptococcal Pharyngitis: a narrative review. *Front Cell Infect Microbiol* 2020;10:563627. <https://doi.org/10.3389/fcimb.2020.563627>
- Akhtar M, Van Heukelom PG, Ahmed A, Tranter RD, White E, Shekem N, et al. Telemedicine physical examination utilizing a consumer device demonstrates poor concordance with in-person physical examination in emergency department patients with sore throat: a prospective blinded study. *Telemed J E Health* 2018;24(10):

- 790-6. <https://doi.org/10.1089/tmj.2017.0240>
7. Yoo TK. Toward automated severe pharyngitis detection with smartphone camera using deep learning networks [Internet]. Amsterdam, Netherlands: Mendeley Data, Elsevier; 2020 [cited at 2024 Jan 24]. Available from: <https://doi.org/10.17632/8ynyhnj2kz.2>.
  8. Howard AG, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T, et al. Mobilenets: efficient convolutional neural networks for mobile vision applications [Internet]. Ithaca (NY): arXiv.org; 2017 [cited at 2024 Jan 24]. Available from: <https://doi.org/10.48550/arXiv.1704.04861>.
  9. Sandler M, Howard A, Zhu M, Zhmoginov A, Chen LC. MobileNetV2: inverted residuals and linear bottlenecks [Internet]. Ithaca (NY): arXiv.org; 2019 [cited at 2024 Jan 24]. Available from: <https://doi.org/10.48550/arXiv.1801.04381>.
  10. Howard A, Sandler M, Chen B, Wang W, Chen LC, Tan M, et al. Searching for MobileNetV3. Proceedings of IEEE/CVF International Conference on Computer Vision (ICCV); 2019 Oct 27–Nov 2; Seoul, South Korea. p. 1314-24. <https://doi.org/10.1109/ICCV.2019.00140>
  11. He KM, Zhang XY, Ren SQ, Sun J. Deep residual learning for image recognition. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2016 Jun 27–30; Las Vegas, NV, USA. p. 770-8. <https://doi.org/10.1109/CVPR.2016.90>
  12. Tan M, Le QV. EfficientNet: rethinking model scaling for convolutional neural networks [Internet]. Ithaca (NY): arXiv.org; 2020 [cited at 2024 Jan 24]. Available from: <https://doi.org/10.48550/arXiv.1905.11946>.
  13. Yoo TK, Choi JY, Jang Y, Oh E, Ryu IH. Toward automated severe pharyngitis detection with smartphone camera using deep learning networks. Comput Biol Med 2020;125:103980. <https://doi.org/10.1016/j.combiomed.2020.103980>
  14. Askarian B, Yoo SC, Chong JW. Novel image processing method for detecting strep throat (streptococcal pharyngitis) using smartphone. Sensors (Basel) 2019;19(15): 3307. <https://doi.org/10.3390/s19153307>