Journal of Information System and Technology Management JISTM
EISSN: 0128-1666
Volume 8 Issue 32 (September 2023) PP. 113-128
DOI: 10.35631/JISTM.832008

JOURNAL OF INFORMATION SYSTEM AND TECHNOLOGY MANAGEMENT (JISTM)
www.jistm.com

UTILIZING DEEP LEARNING ON WEB APPLICATION FOR COVID-19 PNEUMONIA DETECTION

Nheelam Mitthera a/p Nandagobalan[1], Jun Kit Chaw[2*]

[1] Faculty of Computing And Information Technology, Tunku Abdul Rahman University of Management and Technology, Malaysia
Email: nheelamn@gmail.com
[2] Institute of Visual Informatics, Universiti Kebangsaan Malaysia, Malaysia
Email: chawjk@ukm.edu.my
[*] Corresponding Author

Article Info:

Article history:
Received date: 25.07.2023
Revised date: 15.08.2023
Accepted date: 20.09.2023
Published date: 27.09.2023

To cite this document:

Nandagobalan, N. M., & Chaw, J. K. (2023). Identify And Localize Covid-19 Abnormalities on Chest Radiographs. *Journal of Information System and Technology Management, 8* (32), 113-128.

DOI: 10.35631/JISTM.832008

Abstract:

Quarantining Coronavirus (COVID-19) patients is a critical measure to prevent the spread of the disease, thus it is important to perform the correct diagnosis quickly. Among other diagnosis tools, radiograph scans exist as one of the common screening methods for detecting COVID-19. This study aims to reduce the healthcare workers' burden by mitigating the impacts of COVID-19 pneumonia by developing an image classification system to assess radiographs in an accurate and timely manner. The proposed work also assists junior radiologists in familiarizing themselves with reading radiographs confidently using localization features through Explainable Artificial Intelligence. This project focuses on the detection of COVID-19 pneumonia in a multi-classification environment by pre-trained convolutional neural network models such as Inception-V3, VGG-16, and VGG-19. The overall system's idea is to first perform transfer learning based on the binary classification of Normal and Pneumonia. The knowledge is then transferred to the task-specific models to classify COVID-19, bacterial, viral pneumonia, and normal. Out of all the models, VGG-19, the multiclass CNN model outperformed others by achieving a 99.27% accuracy in detecting COVID-19 in X-rays. The model also has the highest overall accuracy of 85.59% compared to multiclass Inception-V3, VGG-16, VGG-19, and the ensemble model. It was then chosen to be deployed as the Web Application where users are able to pass an input of radiograph to get an output with the label classifying whether the patient is infected with COVID-19 pneumonia, viral pneumonia, bacterial pneumonia, or normal together with highlights on the infected areas of lungs.

Keywords:

Convolutional Neural Networks, COVID-19, chest X-ray, Deep Learning, X-ray Classification

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 8 Issue 32 (September 2023) PP. 113-128
DOI: 10.35631/JISTM.832008

Introduction

COVID-19 is a contagious disease caused by the SARS-CoV-2 virus. It has since spread globally, leading to a declaration of outbreak by the World Health Organization (WHO) and the implementation of measures like social distancing and quarantine to prevent its spread. One critical measure to control the spread is to diagnose infected individuals quickly and accurately.

Reverse transcription polymerase chain reaction (RT-PCR) is the standard reference test for detecting viruses in the lungs, including COVID-19, but it has limited supplies, requires proper laboratory facilities, and can have high false-negative rates (Fang et al., 2020). Computed tomography (CT) is another tool used for diagnosing COVID-19 pneumonia. However, CXR is more widely used due to its cost-effectiveness, availability, and convenience (Wang, Li, Li, Zhang, & Wang, 2021). Chest X-rays (CXRs) can be used to diagnose COVID-19 and differentiate it from other types of pneumonia, but accurate diagnosis can be challenging as it requires radiologists to have sufficient knowledge, skills, and time to read radiographs following existing medical guidelines.

With the increasing severity of COVID-19, public health is at risk and requires immediate assistance from the healthcare industry to control the rising hospital. Hence, healthcare workers must be able to diagnose rapidly and act promptly to give the correct treatment, especially when those who have taken a CXR can have a more severe case.

Our work developed a web application that utilized VGG-16 model to detect various forms of pneumonia through CXR images. The proposed model used deep learning for multilabel classification and localization through visual explanation. It focused on opacity, and the cloudy grey areas in radiographs, which indicates the presence of COVID-19 or other forms of pneumonia. The multiclass classification has 4 labels: Normal, Covid-19, Viral Pneumonia, and Bacterial Pneumonia. The prototype is intended for moderate to severe COVID-19 cases and can aid medical undergraduates, newly trained doctors, interns, general practitioners, and junior radiologists in learning to read radiographs. Senior healthcare professionals can also use the model for confirmatory purposes. Besides that, this work used interpretability saliency maps as an attention mechanism to concentrate feature representations on specific image regions that indicate the class to which they pertain. This study represents a proof-of-concept work inspired by the research of Rajaraman et al, as outlined in their paper titled ‘Iteratively Pruned Deep Learning Ensembles for COVID-19 Detection in Chest X-Rays’ (Rajaraman et al., 2020).

Literature Review

This section presents the literature review of this study by dividing it into three parts. The first part of the literature review will study the existing methods in the market that are used to diagnose COVID-19 traditionally while the second part of this study breaks down the different architectures of Convolutional Neural Network precisely in relation to previous literature work by researchers based on the topic of diagnosing COVID-19 in chest X-rays. A brief description of the Gradient-weighted Class Activation Mapping technique is included in the third part of the literature review.

Diagnose Methods for COVID-19

Diagnosing COVID-19 is essential in order to contain the spread of the virus and provide appropriate medical care to infected individuals. There are several methods available for

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 8 Issue 32 (September 2023) PP. 113-128
DOI: 10.35631/JISTM.832008

diagnosing the presence of COVID-19 pneumonia, each with its own advantages and limitations. The methods are categorized into invasive and non-invasive methods. Invasive methods are diagnostic techniques that involve penetrating the skin or body cavity to obtain a sample for analysis. These methods usually require trained professionals to perform the procedure and can be uncomfortable or painful for the patient. Non-invasive methods on the other hand do not require penetration and are generally easier and less uncomfortable for the patient. They involve collecting respiratory secretions or analysing imaging studies of the lungs.

One of the most widely used invasive methods for diagnosing COVID-19 is reverse transcription–polymerase chain reaction (RT-PCR), which is considered the gold standard for COVID-19 diagnosis (Mahendiratta et al., 2020). RT-PCR is a molecular test that detects the Ribonucleic acid (RNA) of the SARS-CoV-2 virus, if present, in a specimen sourced from nasopharyngeal, oropharyngeal or saliva swabs (Sethuraman, Jeremiah, & Ryo, 2020), which involves inserting a long swab into the back of the throat or nose to collect a sample of mucus. The RT-PCR test takes about 4 to 24 hours to display its results and is highly available in markets. However, it is slightly low in specificity, which can sometimes result in false positive results, typically to asymptomatic patients with low viral load or due to technical errors (Kucirka, Lauer, Laeyendecker, Boon, & Lessler, 2020).

Another widely used diagnostic method is the Antigen Rapid Test Kit (RTK-Ag), which is a more rapid and economical alternative to RT-PCR (Albert et al., 2021). RTK-Ag tests are commonly used in at-risk environments like hospitals, schools, and universities where personnel are frequently exposed to the virus (Augustine et al., 2020). The test results are displayed in less than 30 minutes and the test is done by collecting a swabbed nasal, throat or saliva sample. While the collection of nasal and throat sample is invasive, a saliva sample is considered to be non-invasive. However, RTK-Ag tests have several limitations, such as low specificity and high sensitivity, which results in a 27.9% false negative rate and no false-positive results (Cerutti et al., 2020). It is suggested that RTK-Ag tests should be followed by a RT-PCR analysis for a more precise diagnosis and confirmation.

Chest CT scans is a non-invasive medical imaging technique used for detecting COVID-19 pneumonia. CT scans can show the severity of the disease over time by capturing its varying appearance (Harahwa et al., 2020). However, CT scans may not reveal lung opacities, which are the baseline for disease detection, during the first two days of illness. The use of CT scans is also prone to accidental transmission of the disease if proper personal protective equipment is not used (Raptis et al., 2020).

CXR are a commonly used radiology modality for initial assessments of patients suspected of having COVID-19 (Devi, A., A., & H., 2021). It is widely used as a first-line investigation tool in countries like China and Italy due to its speed and infection control capabilities (Kermali, Khalsa, Pillai, Ismail, & Harky, 2020). It also has the advantage of being a non-invasive diagnose method. In addition to the benefits of less discomfort and ease of performance, non-invasive methods also reduce the risk of complications. Non-invasive methods do not carry these risks. It is also safer for healthcare workers because invasive methods do not involve exposure to potentially infectious bodily fluids, which can help to reduce the risk of infection for healthcare workers performing the tests.

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 8 Issue 32 (September 2023) PP. 113-128
DOI: 10.35631/JISTM.832008

Convolutional Neural Network in Diagnosing COVID-19

Many researchers have used Convolutional Neural Networks (CNN) as a base to develop solutions for identifying COVID-19 in chest radiographs (Musleh & Maghari, 2020). CNN is a deep learning algorithm that primarily focuses on the domain of computer vision applications such as object detection, object recognition, face recognition and many more. Although CNN is similar to Artificial Neural Network (ANN), it varies in terms of connectivity. In ANN, all the neurons from the previous layer are connected to the neurons in the subsequent layer whereas only the last layer of CNN is fully connected. ANN is not suitable to process images in large size because it is prone to overfitting due to its full connectivity (Singla, 2020).

One such CNN architecture is the VGG-16, developed by Karen Simonyan and Andrew Zisserman in 2014, with 16 layers and a large number of parameters, making it a large and complex model (Author: Tinsy John Perumanoor, n.d.). Renukadevi et al. implemented the VGG-16 model to classify patients with and without COVID-19 using CXR datasets of different sizes (Renukadevi, Saraswathi, Vigneshwaran, Pradeep, & Arulsanthosh, 2021). The images were preprocessed by rescaling them to 224x224 pixels and standardizing the intensity values to range between 0 and 1. The VGG-16 model achieved an accuracy of 99.81% compared to the LeNet model, which only had an accuracy of 98.72%. LeNet is another CNN network made up of 5 layers with hyperparameters that tune themselves automatically. However, both models have the risk of overfitting due to the small sample size used in the study.

The Xception CNN, introduced by Google in 2017, is a 71 layer deep model and an advanced version of the Inception model, with linear stack of distinguishable depth convolutions (Jabber, Lingampalli, Basha, & Krishna, 2020). Santoso and Purnomo proposed a further advanced Xception model, called FCovNet, for COVID-19 detection, with extra dense layers and batch normalization to prevent overfitting (Santoso & Purnomo, 2020). FCovNet achieved an accuracy score of 99.1% on the validation set, surpassing other models such as Resnet50, Inception V3, and Xception. FCovNet also had a higher recall compared to other models, but the added layers and batch normalization resulted in higher computational and training cost.

In a study by Narin et al., a model based on the ResNet-50 architecture was proposed to minimize the rate of false positive and negative cases and reduce training time (Narin, 2020). The model applied 5 convolution blocks and used a Support Vector Machine (SVM) for multi-class classification, with a quadratic SVM function achieving the highest accuracy of 95.28% in patient condition classification. MobileNet, a low-latency CNN architecture developed by Google in 2017, was used by Jabber et al. in a comparison study of different CNN architectures for COVID-19 detection. The study found that MobileNet achieved the highest accuracy score of 98.9% among the four evaluated models (Jabber et al., 2020).

Table 1 presents a comparison table as a summary for the studies mentioned in this section.

Table 1 : Task-Specific Models Results

Study Author	CNN Architecture	Key Details	Accuracy	Remarks
Renukadevi et al., 2021	VGG-16	Used VGG-16 for	99.81%	Achieved high accuracy with

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 8 Issue 32 (September 2023) PP. 113-128
DOI: 10.35631/JISTM.832008

		COVID-19 classification with image preprocessing details.		preprocessing but noted the risk of overfitting.
Jabber, Lingampalli, Basha, & Krishna, 2020	Xception	Introduced the Xception CNN architecture and an advanced version called FCovNet.	99.1%	FCovNet outperformed other models but with increased computational cost.
Narin, 2020	ResNet-50	Proposed a ResNet-50-based model with SVM for multi-class classification.	95.28%	Focused on reducing false positives/negatives and training time.
Jabber et al., 2020	MobileNet	Used MobileNet for COVID-19 detection and compared it to other CNN architectures.	98.9%	MobileNet was the top-performing model in the comparison study.

Gradient-weighted Class Activation Mapping (GRAD-CAM)

The Gradient-weighted Class Activation Mapping (GRAD-CAM) technique has been used in various medical image analysis studies to understand the features of certain diseases and improve their detection accuracy. This technique has also been utilized in the detection and diagnosis of COVID-19.

One study by Panwar et al., used the GRAD-CAM technique to visualize the regions of interest (ROIs) in COVID-19 X-ray images. The study found that the technique was effective in identifying the regions of the image that the model relied on to make its prediction, which could be useful for radiologists in understanding the disease and its progression (Panwar et al., 2020).

Similarly, a study by Umair et al., GRAD-CAM was used to visualize the regions of the X-ray images that were most important in the CNNs' decisions. By analysing the Grad-CAM visualizations, the authors were able to identify the specific regions of the lungs that were affected by COVID-19 pneumonia, such as the periphery of the lungs and the lower lobes. This information can be helpful in understanding the pathology of the disease and in developing

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 8 Issue 32 (September 2023) PP. 113-128
DOI: 10.35631/JISTM.832008
targeted treatments. It is stated the use of Grad-CAM in combination with transfer learning and CNNs can be an effective approach for detecting COVID-19 pneumonia in X-ray images (Umair et al., 2021).

Overall, the GRAD-CAM technique has been chosen to visualize COVID-19 pneumonia in X-ray and CT images due to its ability to generate a heatmap that highlights the most relevant features and regions of the image for classification. This can help medical professionals in understanding the features of the disease and improving its detection accuracy.

In addition, information systems play an important role in the medical field, enabling healthcare providers to manage patient data, make informed decisions, and improve patient care. An area where information systems have been useful is in the development of Artificial Intelligence (AI)-powered diagnostic tools. Information systems have enabled the integration and analysis of large datasets, improving the accuracy and efficiency of these AI-based diagnostic tools (Jin et al., 2020). By solely training the Machine Learning (ML) model, practitioners may not have a clear understanding of how to effectively utilize my COVID detection tool. This proposed method of a covid detection system via web application, however, enhances interpretability by highlighting infected areas in chest x-rays and provides a practical application platform for practitioners to both use and upload images.

Methodology

The methodology of the proposed system is illustrated in **Figure 1**. The overall system's idea is to first perform modality-specific transfer learning based on binary classification of Normal and Pneumonia. The knowledge is then transferred to the task-specific models to classify COVID-19, bacterial, viral pneumonia and normal. The algorithms involved, Inception-V3, VGG16 and VGG19, are then combined together as an ensemble model. The best performing model is deployed as a web application using Streamlit and hosted to a public server provided by Ngrok. User should be able to pass an input image of an CXR and expect to receive the output of a labeled CXR classifying it as Normal, COVID-19, Bacterial or Viral Pneumonia along with GRAD-CAM Visualization.

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 8 Issue 32 (September 2023) PP. 113-128
DOI: 10.35631/JISTM.832008

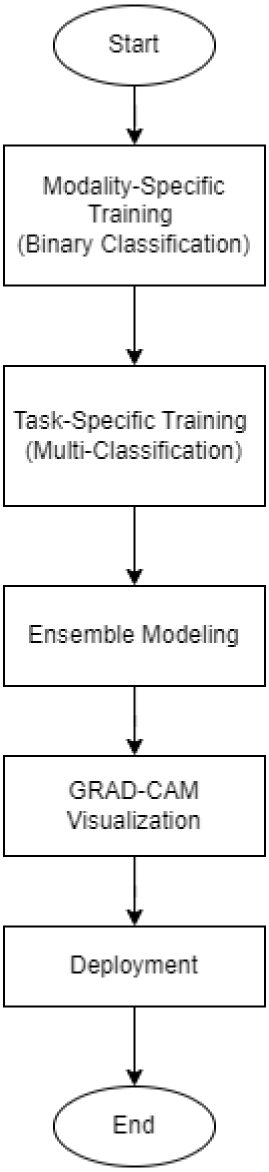

Figure 1: Methodology Of The Study

Data Description

There are two types of datasets used in this project; 'Normal_Pneumonia_xray' for the modality-specific training and 'Lung Disease Dataset' for task-specific training.

The first dataset named "Normal_Pneumonia_xray" used in this study was obtained from Prashant Patel, a Kaggle user. The data was gathered from three prominent public sources, including the COVID-19 Image Data Collection, COVID-NET, and Chest X-Ray Images (Pneumonia), which were created by Cohen et al, Wang et al, and Paul Mooney, respectively. The dataset comprises a total of 6432 chest X-ray images, with 1288 images dedicated for testing and 5144 images for training. Initially, the dataset had three categories: COVID-19, Pneumonia, and Normal. However, to simplify it, the COVID-19 data was merged with the Pneumonia data, resulting in only two classes: Pneumonia and Normal.

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 8 Issue 32 (September 2023) PP. 113-128
DOI: 10.35631/JISTM.832008

The second dataset used in this study, known as the Lung Disease Dataset, was sourced, compiled and modified from Rahman et al. (Rahman et al., 2020) and Kermany et al. (Kermany, 2018). The Lung Disease Dataset originally contained chest X-rays of four types of lung diseases (Coronavirus Disease, Bacterial Pneumonia, Viral Pneumonia, Tuberculosis, and normal patients). The dataset was then modified with Visipics software to remove duplicates and augmented the dataset by a factor of 6, resulting in 10095 images.

The scope of this study focuses on chest X-rays related to COVID-19, Bacterial Pneumonia, Viral Pneumonia, and Normal, and any folders with X-rays related to Tuberculosis were manually removed. The validation data was taken as a subset of the training data, so the validation folder was omitted. The images are in PNG format with varying resolutions, with 4834 images in the training set and 1617 images in the testing set. The folders for the four classes contain 2031, 2009, 2008, and 2013 images respectively.

Data Preparation

The images are in PNG format with varying resolutions, with 4834 images in the training set and 1617 images in the testing set. The folders for the four classes contain 2031, 2009, 2008, and 2013 images respectively.

Model

Modality-specific features are learned from the Normal_pneumonia_xray dataset where the system has to classify the CXR input between normal and pneumonia. To solve the problem of model bias and overfitting from high inter-class similarity and low intra-class variation found in medical images, Modality-Specific Knowledge Transfer is applied. Using a large CXR dataset to retrain CNN models to learn modality-specific feature representations can improve generalization and performance of the model.

The knowledge learned based on the weight layers during the modality specific training is transferred and fine-tuned to the multiclass classification models. The 3 task-specific models, Inception-V3, VGG16 and VGG19 perform multiclass classification individually on the Lung Disease Dataset by predicting whether the CXR belongs to a COVID-19, Bacterial Pneumonia, Viral Pneumonia, or a normal patient. The three models are instantiated with modality-specific weights, trimmed at fully connected layers, and appended with 4 class-specific heads.

The ensemble model is the result from the weighted averaging of 3 multi-classification models; Inception-V3, VGG16 and VGG19. These models are loaded and stacked into an ensemble with functions imported from the kerasSurgeon library. Different weights of 0.5, 0.4, and 0.1 are assigned to each multi-classification model according to its performance, indicating the significance of each model for prediction.

Visualization

Our work utilized Gradient-weighted Class Activation Mapping (Grad-CAM) as the visualization technique which is based on class-specific heatmaps (Selvaraju et al., 2017). Grad-CAM can be used for weakly-supervised localization, which is when a model is trained solely on whole-image labels rather than explicit location annotations to determine the location of specific objects. In this study, the aim is to localize the areas in the lungs affected by coronavirus, bacterial and viral pneumonia for the guidance of healthcare practitioners as mentioned in the problem statement. Bright red pixels on two-dimensional heat maps indicate

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 8 Issue 32 (September 2023) PP. 113-128
DOI: 10.35631/JISTM.832008

pixels with greater relevance and, therefore, weights for classifying the test sample as a specific class are reliable. In order to make the predictions, distinct colour transitions are seen for various pixel significance ranges. By overlaying the heat maps on the input sample CXR, salient Regions of Interest (ROIs) found in the CXR are discovered.

Web Application Deployment

The overview of the web application deployment is illustrated in **Figure 2**. The web app is hosted on one of the ngrok domains, therefore no public domain or IP address is required. A ngrok account must be registered at ngrok.com to obtain the authentication token (auth_token). After this is accomplished, the generated URL will redirect to the web app. The front end of the web application, Chest X-ray Classification, was developed using python and Streamlit to identify COVID-19 infection using a chest X-ray image, along with other types of diseases such as Bacterial and Viral Pneumonia. The "Upload File" option allows the user to submit an X-ray image for the results. As shown in **Figure 3**, the GRAD-CAM of the input image is shown beside the original image for comparison to the emphasized regions which the model has learned from. Upon clicking the ‘Show guide’ checkbox, users can see the Heatmap guide of the GRAD-CAM. The ‘Export Results’ button allows the user to download the GRAD-CAM image, Heatmap, and Prediction results as a ZIP file. After the image has been uploaded, the prediction results are displayed as shown in **Figure 4**. The tested image will also be assigned to the rest of the classes, along with its confidence score.

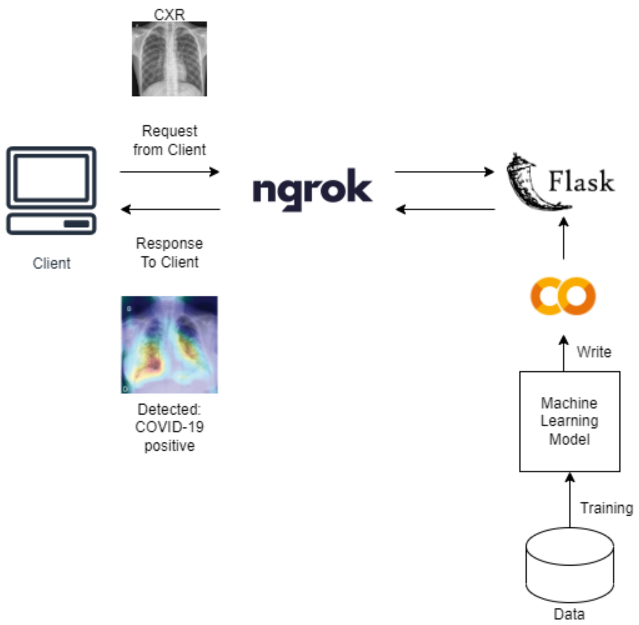

Figure 2: Web Application Deployment Overview

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 8 Issue 32 (September 2023) PP. 113-128
DOI: 10.35631/JISTM.832008

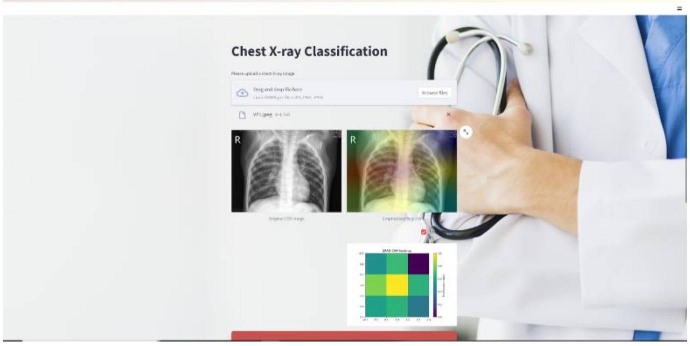

Figure 3: Web App Screenshot with GRAD-CAM

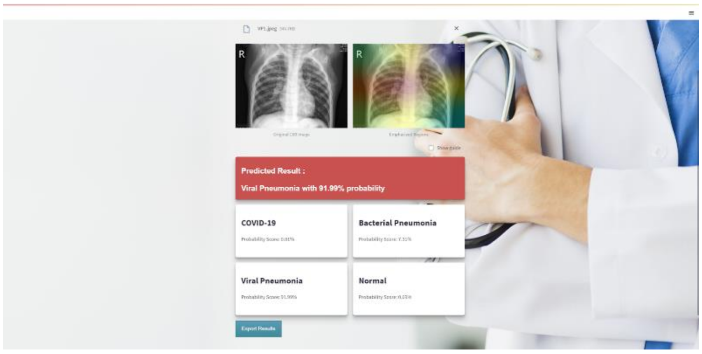

Figure 4: Web App Screenshot with Prediction Results

Results

The modality-specific models, task-specific models and ensemble model are evaluated using metrics such as recall, precision, F1 score, accuracy, and AUC score. **Table** shows the results of the modality-specific models during binary classification of Normal and Pneumonia class.

Table 2 : Modality-Specific Models Results

Models/Metric	Precision	Recall	F1 Score	AUC	Accuracy
Inception-V3	0.9404	0.9356	0.9368	0.9890	0.9356
VGG-16	0.9561	0.9557	0.9559	0.9897	0.9557
VGG-19	0.9512	0.9480	0.9488	0.9900	0.9480

Table shows the results obtained for the 3-task specific models, along with the ensemble model in classifying CXR as Normal, COVID-19, Viral Pneumonia or Bacterial Pneumonia. The ensemble model is the result of weighted averaging by combining Inception-V3, VGG-16, and VGG-19 models. Since a classification report containing the precision, recall, and f1 score can't handle a mix of multilabel-indicator and binary targets, only the accuracy of the Ensemble model is evaluated.

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 8 Issue 32 (September 2023) PP. 113-128
DOI: 10.35631/JISTM.832008

Table 3 : Task-Specific Models Results

Models/Metric	Precision	Recall	F1 Score	AUC	Accuracy
Inception-V3	0.8317	0.8349	0.8330	0.9598	0.8349
VGG-16	0.8298	0.8306	0.8297	0.9579	0.8306
VGG-19 (deployed model)	**0.8582**	**0.8559**	**0.8530**	**0.9653**	**0.8559**
Ensemble	-	-	-	-	0.8462

In the binary classification, the VGG-16 network outperformed the other models by achieving an accuracy of 95.57% while the VGG-19 was the best model in the multiclass classification, resulting in 85.59% accuracy. The ensemble model scored a higher accuracy of 84.62% than VGG-16 and Inception-V3 but lower than VGG-19. Hence, the VGG-19 model is chosen to be deployed as a web application. For better understanding of the results, the ROC-AUC curve is shown for all the models in **Figure 5**, **Figure 6**, **Figure 7Figure , Figure 8, Figure 9**, and **Figure 10** respectively.

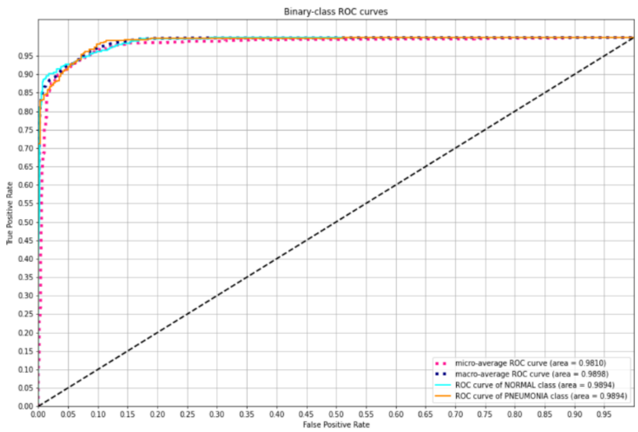

Figure 5: ROC Curve of Inception-V3(Binary Classification)

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 8 Issue 32 (September 2023) PP. 113-128
DOI: 10.35631/JISTM.832008

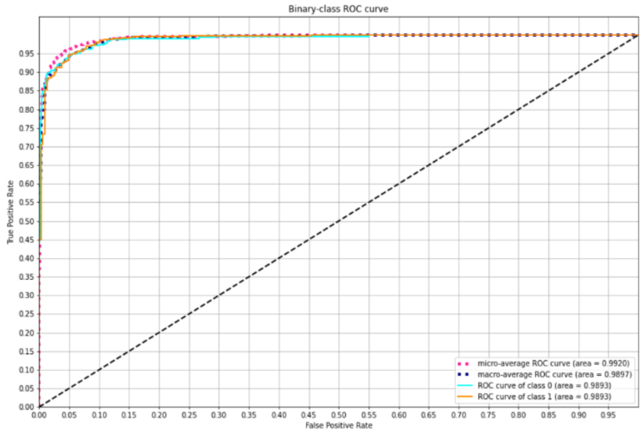

Figure 6: ROC Curve of VGG-16(Binary Classification)

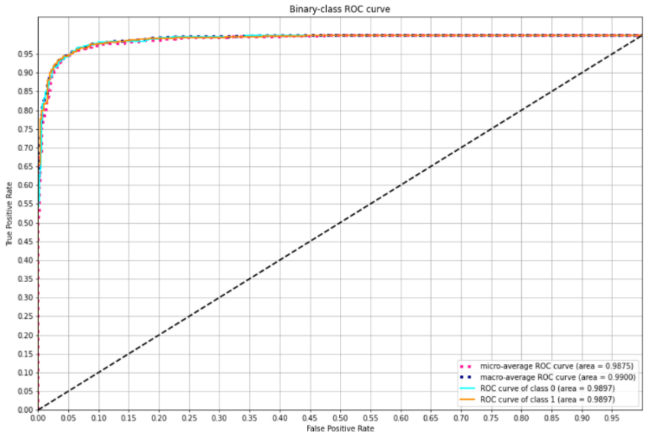

Figure 7: ROC Curve of VGG-19(Binary Classification)

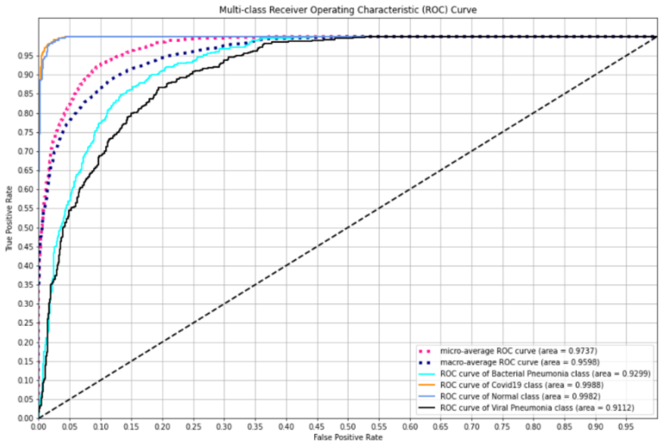

Figure 8: ROC Curve of Inception-V3(Binary Classification)

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 8 Issue 32 (September 2023) PP. 113-128
DOI: 10.35631/JISTM.832008

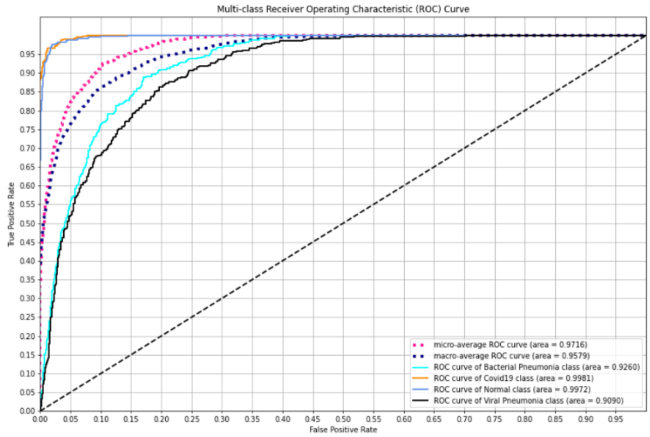

Figure 9: ROC Curve of VGG-16(Multiclass)

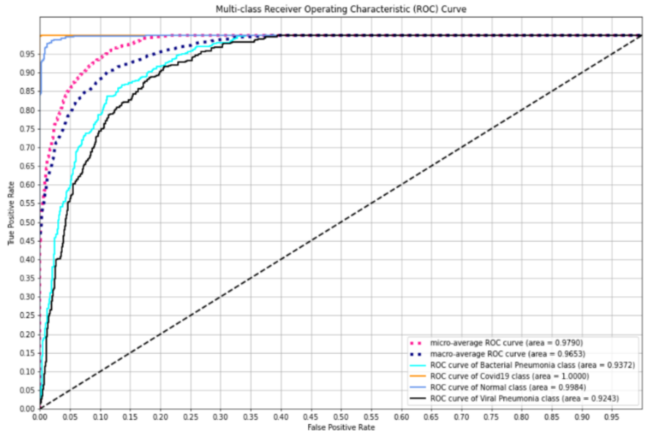

Figure 10: ROC Curve of VGG-19(Multiclass)

VGG-16 and VGG-19 are both deep convolutional neural networks (CNNs) with different architectures. The number of layers and their configurations in these models can make them more suitable for specific tasks. In this case, VGG-16 might have features that are more discriminative for binary classification, while VGG-19 excels in multiclass classification. The distribution of data points across classes can impact model performance. If one class is significantly larger or smaller than the others, it can affect accuracy. The choice of the "best" model can also depend on how well it handles the distribution of data in the multiclass problem. The distribution of data points across classes can impact model performance. If one class is significantly larger or smaller than the others, it can affect accuracy. The choice of the "best" model can also depend on how well it handles the distribution of data in the multiclass problem.

Discussion

The VGG-16 model is found to be more accurate than the other models based on Table 2. In comparison to VGG19 and inception-V3, the aforementioned model showed promising accuracy values compared to other models. The recall value obtained from this model is also higher than other models as it shows that the model is reliable in the healthcare industry due to the higher acceptance of false positives than false negatives. This is due to the architecture depths of the VGG-16 model being the best for learning the hierarchical representations of features from the CXR data and categorizing them into classes for normal and pneumonia. The

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 8 Issue 32 (September 2023) PP. 113-128
DOI: 10.35631/JISTM.832008

VGG-16 model also outperformed the competition as measured by the F-score, which provides a balanced measure of precision and recall.

In accordance with Table 3, it can be observed that the VGG-19 model was more accurate for multi-class classification with a higher AUC measure, indicating that it has the smallest margin for error and hence offers a more precise estimate. Interestingly, the ensemble model achieved higher accuracy than Inception-V3 and VGG-16 but not more than VGG-19. In comparison to the VGG-16, Ensemble and Inception-V3 models, the VGG-19 model performed better when F-score, Precision and Accuracy were taken into account.

The heatmap generated by Grad-CAM is typically color-coded, with warmer colors such as red,orange and yellow indicating regions that are highly relevant to the predicted class, while cooler colors like purple, blue, green indicate regions that are less relevant. The intensity of the color represents the degree of relevance, with brighter colors indicating higher relevance as shown in the heatmap guide Figure 3. By analyzing and interpreting the heatmap, insights can be gained into which regions of the input image are most relevant for the model's prediction, which in turn can help to identify areas for potential improvement. For example, if the model is making incorrect predictions, examining the heatmap can help researchers identify which regions of the input image the model is not attending to, and which regions it is overemphasizing. This information can be used to refine the model's architecture or to guide the selection of additional training data.

Conclusion

In conclusion, this project has successfully met its objective to develop an X-ray Classification system that identifies and localized COVID-19 in a multiclass environment including classes such as Normal, Bacterial Pneumonia, and Viral Pneumonia. VGG-19, the multiclass CNN model which was chosen to be deployed into the Web Application, has particularly achieved a 99.27% accuracy in detecting COVID-19 in X-rays. Although the overall accuracy of the model is 85.59%, the project's objective mainly focuses on COVID-19 disease. The proof-of-concept system built in this project eases the radiologists' workload and saves their time so that they can diagnose diseases early, accurately, and quickly which is significant to mitigate the COVID-19 outbreak due to its communicable nature. Ultimately, the disease localization feature through GRAD-CAM included in the developed system also helps doctors to take the next step in the diagnosis. Using CXR for all the individuals who are suspected to have Covid is unideal. Hence, the system can only cater to COVID-19 patients with moderate to severe respiratory issues where their cases require further investigation by healthcare experts. The project's localization capability, which is an additional but crucial feature to enforce reliability, would be considered its strength. On the other hand, the model's lower recall than the acquired precision can be seen as a drawback because, in the healthcare sector, it is typically recommended for a higher recall because erroneous positives are more acceptable than false negatives. One of the limitations of this project is GPU restriction on the total runtime, hindering from exploring other types of CNN models, fine-tuning its parameters and adding more data to the training. As a future improvement, it is recommended to add more data to the CNN models running on a reliable platform with faster and more dependable computational resources, aiming for better accuracy and recall.

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 8 Issue 32 (September 2023) PP. 113-128
DOI: 10.35631/JISTM.832008

Acknowledgment
The authors would like to thank the Institute of Visual Informatics, University Kebangsaan Malaysia (UKM) for providing the research facilities through Geran Galakan Penyelidik Muda (GGPM-2022-063).

References

Albert, E., Torres, I., Bueno, F., Huntley, D., Molla, E., Fernández-Fuentes, M. Á., … Navarro, D. (2021). Field evaluation of a rapid antigen test (Panbio™ COVID-19 Ag Rapid Test Device) for COVID-19 diagnosis in primary healthcare centres. *Clinical Microbiology and Infection : The Official Publication of the European Society of Clinical Microbiology and Infectious Diseases*, *27*(3), 472.e7-472.e10. https://doi.org/10.1016/j.cmi.2020.11.004

Augustine, R., Das, S., Hasan, A., S, A., Abdul Salam, S., Augustine, P., … Kevadiya, B. D. (2020). Rapid Antibody-Based COVID-19 Mass Surveillance: Relevance, Challenges, and Prospects in a Pandemic and Post-Pandemic World. *Journal of Clinical Medicine*, *9*(10). https://doi.org/10.3390/jcm9103372

Author: Tinsy John Perumanoor. (n.d.). What is VGG16? — Introduction to VGG16. *Medium*.

Cerutti, F., Burdino, E., Milia, M. G., Allice, T., Gregori, G., Bruzzone, B., & Ghisetti, V. (2020). Urgent need of rapid tests for SARS CoV-2 antigen detection: Evaluation of the SD-Biosensor antigen test for SARS-CoV-2. *Journal of Clinical Virology : The Official Publication of the Pan American Society for Clinical Virology*, *132*, 104654. https://doi.org/10.1016/j.jcv.2020.104654

Devi, S. K., A., D., A., R. S., & H., T. K. (2021). A Medical Decision Support System to Detect Covid-19 Pneumonia Using CNN. *2021 Innovations in Power and Advanced Computing Technologies (i-PACT)*, 1–6. https://doi.org/10.1109/i-PACT52855.2021.9696553

Fang, Y., Zhang, H., Xie, J., Lin, M., Ying, L., Pang, P., & Ji, W. (2020). Sensitivity of Chest CT for COVID-19: Comparison to RT-PCR. *Radiology*, *296*(2), E115–E117. https://doi.org/10.1148/radiol.2020200432

Harahwa, T. A., Lai Yau, T. H., Lim-Cooke, M.-S., Al-Haddi, S., Zeinah, M., & Harky, A. (2020). The optimal diagnostic methods for COVID-19. *Diagnosis*, *7*(4), 349–356. https://doi.org/doi:10.1515/dx-2020-0058

Jabber, B., Lingampalli, J., Basha, C. Z., & Krishna, A. (2020). Detection of Covid-19 Patients using Chest X-ray images with Convolution Neural Network and Mobile Net. *2020 3rd International Conference on Intelligent Sustainable Systems (ICISS)*, 1032–1035. https://doi.org/10.1109/ICISS49785.2020.9316100

Jin, C., Chen, W., Cao, Y., Xu, Z., Tan, Z., Zhang, X., … Feng, J. (2020). Development and evaluation of an artificial intelligence system for COVID-19 diagnosis. *Nature Communications*, *11*(1), 5088. https://doi.org/10.1038/s41467-020-18685-1

Kermali, M., Khalsa, R. K., Pillai, K., Ismail, Z., & Harky, A. (2020). The role of biomarkers in diagnosis of COVID-19 - A systematic review. *Life Sciences*, *254*, 117788. https://doi.org/10.1016/j.lfs.2020.117788

Kermany, D. (2018). *Labeled Optical Coherence Tomography (OCT) and Chest X-Ray Images for Classification*. Mendeley. https://doi.org/10.17632/rscbjbr9sj.2

Kucirka, L. M., Lauer, S. A., Laeyendecker, O., Boon, D., & Lessler, J. (2020). Variation in False-Negative Rate of Reverse Transcriptase Polymerase Chain Reaction-Based SARS-CoV-2 Tests by Time Since Exposure. *Annals of Internal Medicine*, *173*(4), 262–267. https://doi.org/10.7326/M20-1495

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 8 Issue 32 (September 2023) PP. 113-128
DOI: 10.35631/JISTM.832008
Mahendiratta, S., Batra, G., Sarma, P., Kumar, H., Bansal, S., Kumar, S., … Medhi, B. (2020). Molecular diagnosis of COVID-19 in different biologic matrix, their diagnostic validity and clinical relevance: A systematic review. *Life Sciences*, *258*, 118207. https://doi.org/10.1016/j.lfs.2020.118207

Musleh, A. A. w. A., & Maghari, A. Y. (2020). COVID-19 Detection in X-ray Images using CNN Algorithm. *2020 International Conference on Promising Electronic Technologies (ICPET)*, 5–9. https://doi.org/10.1109/ICPET51420.2020.00010

Narin, A. (2020). Detection of Covid-19 Patients with Convolutional Neural Network Based Features on Multi-class X-ray Chest Images. *2020 Medical Technologies Congress (TIPTEKNO)*, 1–4. https://doi.org/10.1109/TIPTEKNO50054.2020.9299289

Panwar, H., Gupta, P. K., Siddiqui, M. K., Morales-Menendez, R., Bhardwaj, P., & Singh, V. (2020). A deep learning and grad-CAM based color visualization approach for fast detection of COVID-19 cases using chest X-ray and CT-Scan images. *Chaos, Solitons & Fractals*, *140*, 110190. https://doi.org/https://doi.org/10.1016/j.chaos.2020.110190

Rahman, T., Khandakar, A., Kadir, M. A., Islam, K. R., Islam, K. F., Mazhar, R., … Chowdhury, M. E. H. (2020). Reliable Tuberculosis Detection Using Chest X-Ray With Deep Learning, Segmentation and Visualization. *IEEE Access*, *8*, 191586–191601. https://doi.org/10.1109/ACCESS.2020.3031384

Rajaraman, S., Siegelman, J., Alderson, P. O., Folio, L. S., Folio, L. R., & Antani, S. K. (2020). Iteratively Pruned Deep Learning Ensembles for COVID-19 Detection in Chest X-Rays. *IEEE Access*, *8*, 115041–115050. https://doi.org/10.1109/ACCESS.2020.3003810

Raptis, C. A., Hammer, M. M., Short, R. G., Shah, A., Bhalla, S., Bierhals, A. J., … Henry, T. S. (2020). Chest CT and Coronavirus Disease (COVID-19): A Critical Review of the Literature to Date. *AJR. American Journal of Roentgenology*, *215*(4), 839–842. https://doi.org/10.2214/AJR.20.23202

Renukadevi, N. T., Saraswathi, K., Vigneshwaran, N., Pradeep, V., & Arulsanthosh, T. B. (2021). Performance of CNN Architectures in the Detection of Covid-19 Disease. *2021 Innovations in Power and Advanced Computing Technologies (i-PACT)*, 1–5. https://doi.org/10.1109/i-PACT52855.2021.9696954

Santoso, F. Y., & Purnomo, H. D. (2020). A Modified Deep Convolutional Network for COVID-19 detection based on chest X-ray images. *2020 3rd International Seminar on Research of Information Technology and Intelligent Systems (ISRITI)*, 700–704. https://doi.org/10.1109/ISRITI51436.2020.9315479

Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. *2017 IEEE International Conference on Computer Vision (ICCV)*, 618–626. https://doi.org/10.1109/ICCV.2017.74

Sethuraman, N., Jeremiah, S., & Ryo, A. (2020). Interpretación de las pruebas de diagnóstico del SARS-CoV-2. *JAMA*, *323*(22), 2249–2251.

Singla, A. (2020). *Convolutional Neural Network Explained | CNN | Deep Learning*.

Umair, M., Khan, M. S., Ahmed, F., Baothman, F., Alqahtani, F., Alian, M., & Ahmad, J. (2021). Detection of COVID-19 Using Transfer Learning and Grad-CAM Visualization on Indigenously Collected X-ray Dataset. *Sensors*, Vol. 21. https://doi.org/10.3390/s21175813

Wang, W., Li, Y., Li, J., Zhang, P., & Wang, X. (2021). Detecting COVID-19 in Chest X-Ray Images via MCFF-Net. *Computational Intelligence and Neuroscience*, *2021*, 3604900. https://doi.org/10.1155/2021/3604900