2023 3rd International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON)
Karnataka, India. Dec 29-31, 2023

2023 3rd International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON) | 979-8-3503-1912-5/23/$31.00 ©2023 IEEE | DOI: 10.1109/SMARTGENCON60755.2023.10441931

Exploring Deep Convolutional Neural Networks: A Grad-CAM Enhanced Comparative Study for Automated COVID-19 Diagnosis from Chest X-ray Images

Md. Sazid Reza*, Mir Md. Jahangir Kabir†, Md. Abdur Rakib Mollah‡, Nazmul Islam Nahin§
Department of Computer Science & Engineering
Data Science and Innovation, TD School†
Department of Electronics & Telecommunication Engineering§
Rajshahi University of Engineering & Technology, Rajshahi, Bangladesh
University of Technology, Sydney, Sydney, Australia†
Emails: *dihansazid@gmail.com, †mmjahangir.kabir@gmail.com, ‡rakib1703115@gmail.com, §nahinete.ruet@gmail.com

***Abstract*—The emergence of COVID-19 in late 2019 sparked a global health crisis. Identifying infections quickly is key to providing proper care and limiting transmission. The Reverse Transcription-Polymerase Chain Reaction (RT-PCR) test, a widely recognized primary diagnostic method for detecting COVID-19, faces challenges associated with cost and compatibility in resource-constrained environments. This issue is resolved by chest X-ray imaging since it is non-invasive and inexpensive. In this domain, research often utilizes pre-trained models on relatively small datasets of chest X-ray images, limiting the potential for optimal generalization and accuracy achievable with larger training sets. Also, these research works limit the exploration of crucial information on salient regions within CXR images, constraining the depth of analysis and interpretation. This study addresses current research challenges through a detailed comparative analysis of convolutional neural network (CNN) architectures, including pre-trained models like Inception V3, Xception, DenseNet-121, VGG-19, and a Custom CNN developed from scratch. The evaluation focuses on their effectiveness in detecting COVID-19 using a relatively vast and comprehensive dataset. The evaluation focuses on assessing the efficacy of these models in detecting COVID-19 across binary-class and multi-class classifications within a large dataset, offering valuable insights into their overall effectiveness. Grad-CAM is employed in this study, which visualizes picture salient areas affecting deep neural network predictions and improves model interpretability. The heatmaps reveal the model's decision-making processes, making their behavior more understandable to medical practitioners. With resulting accuracies of 99.13% and 97.66% in binary-class and multi-class classification, Inception V3 demonstrated the most optimal performance. This proposed approach surpasses existing benchmark methods, underscoring its superiority.**

***Index Terms*—COVID-19, Deep Learning, CNN, RT-PCR, chest X-ray, Inception V3, Xception, DenseNet-121, Grad-CAM, ResNet-50, VGG-19.**

I. Introduction

A significant proportion of Wuhan residents exhibited symptoms suggestive of an unidentified disease resembling pneumonia in December 2019. After one month, many laboratories successfully demonstrated that the illness in question was not pneumonia, but rather a novel infectious disease referred to as COVID-19. The outbreak began with a newly identified coronavirus initially named "2019 novel coronavirus" (2019nCoV), Later recognized as "SARS-CoV-2" (Severe Acute Respiratory Syndrome Coronavirus-2). Common symptoms of this ailment include fever, fatigue, respiratory signs like coughing, sore throat, and difficulty breathing, along with distinct gastrointestinal problems such as diarrhea. [1]. The manifestation of the COVID-19 infection often occurs after a period of 14 days. The projected global mortality rates over a specific timeframe, as determined by the estimated delay in the incubation period, vary between 5% and 7% [2]. The timely identification of COVID-19 is imperative in order to effectively mitigate transmission and provide appropriate medical care to individuals affected by the disease. Real-time reverse transcription poly-reverse chain reaction(real-time RT-PCR) is the predominant method used for the initial detection of COVID-19 (rRT-PCR) [3]. Respiratory specimens are obtained from afflicted individuals, and this methodology yields outcomes within a mere two-day timeframe. However, in situations when resources are constrained, it may not always be readily accessible or cost-effective.

An alternate, less expensive, and non-invasive method of COVID-19 detection has been suggested: employing chest X-ray imaging to identify certain radiographic features associated with COVID-19 pneumonia. Although there are other imaging modalities available, it is widely acknowledged that chest radiography is a specific but not highly sensitive method for detecting clinically significant abnormalities [4]. In order to identify COVID-19 in chest X-rays, skilled radiologists and relevant patient background information are crucial for accurate detection. Deep learning, a subset of machine learning, shows significant promise in image analysis, aiding in the identification of COVID-19 pneumonia in chest X-rays [5]. Deep learning algorithms possess the capability to autonomously acquire knowledge and recognize patterns and characteristics

979-8-3503-1912-5/23/$31.00 ©2023 IEEE

in images without prior familiarity with specific radiographic features linked to COVID-19 pneumonia. Researchers have revealed that various optical marks, such as ground glass opacities—hazy darkish spots—are apparent in the affected patient's lung[6] which aids in disease detection. The objective of this study is to utilize binary-class and multi-class classifications of COVID-19. The multi-class classification consists of the following classes: Normal, Lung Opacity, Pneumonia, and COVID-19. The binary-class classification focuses on distinguishing between COVID-19 and Normal classes. The main contributions of this study are as follows:

- Several pre-trained Convolutional Neural Network (CNN) models are utilized for the purpose of detecting COVID-19 by using their transfer learning capabilities. Following the preliminary stage evaluation, a selection was made to utilize only four pre-trained models (namely Inception V3, Xception, DenseNet-121, and VGG-19) due to their notable performance capabilities.
- A Custom CNN model from scratch was designed specifically for COVID-19 detection. This allows the model to adapt more effectively to the nuances of the target task.
- Grad-Cam is employed in this study which can be regarded as a research gap in most of the related works. The Grad-CAM technique improves model interpretability and transparency in COVID detection by visualizing predictions in X-ray or CT scan pictures. This visualization tool helps radiologists validate model decisions, highlight clinically significant regions, and improve diagnosis confidence. Grad-CAM insights also help uncover and correct biases, inform model iterations, and contribute to open science by offering replicable techniques for researchers and healthcare professional education.

The current article is organized into distinct parts. Section II presents a thorough examination of the current literature on the detection of COVID-19. Section III presents a brief portrayal of the methods and dataset associated with the comparative study. The experimental design is detailed in Section IV, along with a comprehensive comparative analysis of related research papers. The findings derived from the assessment of the proposed methodology are summarised in Section V.

II. Related Works

Early in 2020, Sethi et al. [7] conducted the first significant study in the field of COVID-19 detection. In the early stages of the COVID-19 outbreak, the virus prevalence was rising gradually, and the available datasets contained a limited number of photos. The dataset had 381 CXR images, and a Support Vector Machine (SVM) + ResNet-50 model was employed, yielding an accuracy rate of 95.33%. ResNet-50 was used as the feature selector in this study. Tang et al. [8] proposed a weighted average ensemble method that was implemented on the COVID-X dataset comprising 15477 CXR images of three classes (6053 Pneumonia, 8851 Normal, and 573 COVID-19). They trained six deep learning models on the dataset and the final ensemble strategy outperformed all the initial models by 0.3% margin. The final resultant accuracy in detecting COVID-19 was 94.1%. Senan et al. [9] conducted a study on the COVID-19 dataset of 21165 CXR images, which served as the primary dataset in this study. The data underwent preprocessing and augmentation in a phase. In the context of multiclass classification, ResNet-50 achieved a 95% accuracy, while in the scenario of binary classification, the overall accuracy reached 99%. Dubey et al. [10] conducted a study on a COVID-CTSI dataset comprising of 746 data samples of two classes(COVID and NON-COVID). Out of the evaluated models, which encompassed CTnet-10 and an ensemble approach, the model based on VGG-19 emerged as the leading performer, achieving a commendable accuracy rate of 95%. Ismael [11] initiated research on a dataset of 380 images and conducted a binary classification in 2021. Pre-trained CNN models (VGG16, VGG19, ResNet18, ResNet50, and ResNet101) were used for feature extraction and a support vector machine was used as a classifier. The ResNet50 model, combined with an SVM classifier using the Linear kernel function, had the greatest accuracy of 94.7%. The ResNet50 model achieved a somewhat reduced accuracy of 92.6% after fine-tuning, whilst the CNN model attained a 91.6% accuracy during end-to-end training. In 2022 Kokilavani et al. [12] conducted a comparative analysis of many pre-trained CNN architectures known for their high efficiency, including VGG16, DeseNet121, MobileNet, NASNet, Xception, and EfficientNet. The results of the study revealed that VGG-16 exhibited the highest accuracy of 97.68% in detecting cases of COVID-19. The binary classification was performed on a dataset consisting of 3873 chest X-RAY images. The limitations identified in earlier studies are highlighted in this research. Through the employed methodology, the highest accuracy rates of 99.13% and 97.66% were achieved for binary-class and multi-class classification, respectively, utilizing Inception V3.

III. Methodologies

This section provides a detailed explanation of the methodology adopted in this comparative study. It covers various aspects, including information on the dataset, CNN architectures, transfer learning, and other techniques, as well as details on the experimental design, performance metrics, and visualization methods.

A. Dataset Description and Data Preprocessing

A publicly accessible CXR dataset called the "COVID-19 Radiography Database" was utilized in this study. The dataset comprised a total of 21,165 pictures. Among the photos, about 3616 were classified as COVID-19, 6012 as Lung Opacity, 10192 as Normal patients, and 1345 as Viral Pneumonia. Researchers from Qatar University in Doha and Dhaka University in Bangladesh, together with partners from Malaysia and Pakistan, assembled the database of chest X-ray pictures [13]. Fig. 1 illustrates the classwise image distribution.

Initially, the images were of 224*224 size, but they were resized to 150*150 as resizing CXR images to 150x150 in COVID detection studies offers computational efficiency,

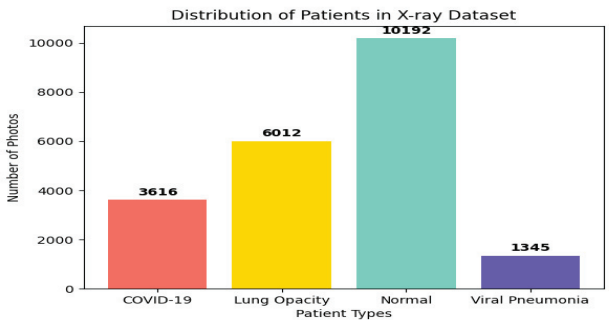


Fig. 1: Dataset image distribution

reduces noise, and strikes a balance between accuracy and resource efficiency.

B. Convolutional Neural Network

CNNs excel in image and video processing by hierarchically extracting local features, efficiently learning through shared parameters, and ensuring translation invariance. The hierarchical architecture, pre-trained models, and specialized 3D CNNs further enhance their effectiveness in capturing spatial and temporal dependencies in visual data. Convolution layers perform linear operations in feature extraction. Applying a kernel to a tensor input, followed by an activation function for non-linear adjustments, CNNs capture intricate patterns and relationships in the data [14]. The pooling layer in CNNs simplifies network calculations by progressively reducing the spatial size of input images, freeing up space, and providing only the most crucial information to subsequent CNN layers [15]. The fully connected layer is essentially learning a (potentially non-linear) function in a feature space that the convolutional layers have essentially provided: one that is meaningful, low-dimensional, and reasonably invariant. The dropout layer, which is the last layer in neural networks prevents overfitting by randomly deactivating neurons during training, promoting robust feature learning and enhancing generalization to new data. In this research, four ImageNet pre-trained CNN models were utilized, namely Xception, Inception V3, DenseNet-121, and VGG-19, along with one Custom CNN model.

C. Transfer Learning

Transfer learning refers to utilizing the knowledge from an existing model. Four pre-trained models (Inception V3, Xception, DenseNet-121, and VGG-19) were employed in this study. All of these models were pre-trained on the ImageNet dataset, denoting that these models have been trained with 1.2 million images from 1000 classes. These are beneficial for classification tasks because, when applied to a big dataset, the models have learned certain feature mappings beforehand. The optimal performance of these models is achieved when they undergo the process of finetuning, specifically tailored for a particular classification job. These models were fine-tuned specifically for COVID-19 detection. Initially, the models were loaded with weights trained on the ImageNet. The final classification layer was frozen to allow training of the fully connected layer for two-class and four-class classification of chest X-ray (CXR) images. In contrast, the classification layer of the pre-trained model is configured to discern 1000 classes from the ImageNet dataset. Then a convolutional base is created which is used as a feature extractor for all these models and a final class-specific (two and four) dense layer with softmax activation is added. The convolutional base was frozen before training to prevent the layer's weight update during the training phase. Fig. 2 illustrates the sequence of steps involved in the modified transfer learning procedure.

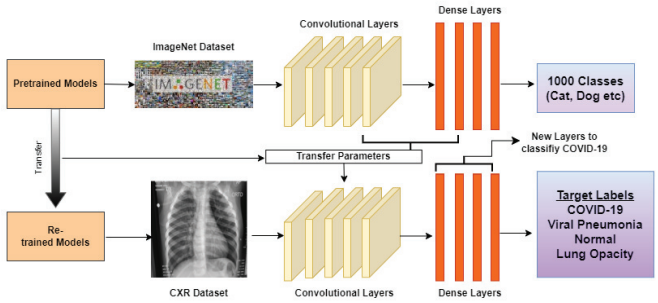


Fig. 2: Modified transfer learning models architecture.

D. Inception V3

Inception V3 is a 48-layer architecture with around 21 million parameters. The primary component of Inception V1 consists of four concurrent layers: 1×1 convolution, 3×3 convolution, 5×5 convolution, and 3×3 max pooling. Inception V3 improves on Inception V1 by increasing efficiency, and network depth without sacrificing speed, and computational cost [16]. The Inception V3 architecture uses auxiliary classifiers for regularization.

E. Custom CNN

The model architecture comprises three convolutional layers with 64, 32, and 32 filters. MaxPooling layers are employed to downsample data while preserving crucial features after each convolution. The third convolutional layer is dedicated to capturing advanced disease patterns and biomarkers. After the convolutional layers, the input undergoes flattening, leading to two dense layers with 50 and 100 neurons, utilizing ReLU activation to identify patterns and intricate feature combinations. The final dense layer employs softmax activation for categorical classification.

F. DenseNet-121

DenseNet-121, a neural network with 121 layers, has approximately 2 million parameters. As the depth of CNNs increases, they encounter the 'vanishing gradient' issue. This problem arises when information dissipates from input to output layers of the network, making training less successful. DenseNet-121 addresses the issue of vanishing gradients in convolutional neural networks (CNNs) by restructuring the architecture, resulting in a more direct and linked relationship between layers.

G. Xception

Xception is a 71-layer deep CNN architecture with depth-wise separable convolutions, enhancing feature learning and parameter efficiency, and promoting better generalization. Xception's 22.8 million parameters play a vital role in COVID-19 detection by enabling the model to learn intricate features from chest X-rays, enhancing its ability to identify patterns associated with the virus.

H. VGG-19

Developed in 2014, VGG-19 is characterized by a complex architecture. Featuring 16 convolutional layers with a 3x3 kernel size and 3 fully connected layers, VGG-19 has been a dominant model in image classification, known for its simple structure but susceptible to the vanishing gradient issue.

I. Grad-CAM Visualization

Grad-CAM stands for Gradient-Class Activation Map. It aids in identifying the regions that a particular model bases its classification on [17]. This procedure produces a heatmap for individual classes using CXR images, substantiating the notion that the pivotal role lies in the final convolution layer, as it retains the crucial information essential for capturing visual patterns within the CXR images.

IV. Experimental Analysis

With the dataset, models, and methodological pipeline established, this section presents the empirical evaluation, comparison, and analysis of the five CNN architectures Inception V3, Xception, DenseNet-121, VGG-19, and the custom CNN.

A. Experimental Setup

In this study, version 3.7 of the Python programming language was utilized, and the Google Colab Pro was utilized to run the GPU Tesla A-100 and V100.

B. Experimental Design

A batch size of 32 and 50 epochs was employed to enhance accuracy and generalization in pattern recognition on challenging datasets. Overfitting was mitigated by utilizing a dropout rate of 0.5, which facilitated diverse feature learning by randomly deactivating half of the neurons during training. Categorical cross entropy was employed for loss calculation. The Adam optimizer, with a default learning rate of 1×10^{-3}, was chosen for its adaptability in dynamically adjusting learning rates based on squared gradients and previous gradients, effectively managing varying feature sizes and expediting convergence during neural network training.

C. Performance metrics

Five diverse performance metrics were implemented for a thorough evaluation of model effectiveness. To assess overall correctness, the Accuracy measure (Acc) was utilized to calculate the ratio of true positives and true negatives to the total instances. Accuracy refers to the proximity of a measured value to a standard or real value :

$$Acc = \frac{TP + TN}{TP + TN + FP + FN}$$

In the formula, TP denotes true positives, TN denotes true negatives, FP represents false positives, and FN stands for false negatives.

Precision ($Prec$) and Recall (Rec) were employed to scrutinize the balance between true positives and the occurrences of false positives and false negatives. Precision specifically assessed the precision of positive predictions:

$$Prec = \frac{TP}{TP + FP}$$

Recall (Rec), which is often referred to as sensitivity or true positive rate, quantifies the ratio of accurately detected real positive cases by a classification model:

$$Rec = \frac{TP}{TP + FN}$$

The F1 Score ($F1$) offered a holistic evaluation of model performance by combining Precision and Recall into a single metric and also reflects the balance between the ability to correctly identify positive instances and the control over false positives:

$$F1 = \frac{2 \times Prec \times Rec}{Prec + Rec}$$

The Loss metric ($Loss$) measures the disparity between predicted and actual values in a machine learning model. Used as an optimization objective during training, it guides the model to minimize errors and enhance overall performance. The loss graph depicts changes in prediction errors during training, providing insights into the learning process.

D. Result Analysis

On Dataset-I, an 80:20 split was used, with 80% of the images going to training and 20% to testing. A 90:10 split of the training dataset was made, with 10% going towards validation data. The validation set, which is utilized throughout the training phase, serves the purpose of minimizing both training loss and validation loss. This distinguishes it from the test set contrast, the testing set operates independently and does not impact the training process. The practice of shuffling was employed to provide randomness in the arrangement of training data prior to its presentation to a model with the aim of mitigating any biases. To prevent biases, training data was shuffled in 42 random states before being fed to a model at random. The final predicted values for this dataset for binary-class and multi-class classification are depicted in TABLE I and II.

Inception V3 emerged as the top-performing pre-trained model, with 99.13% and 97.66% accuracy respectively in the binary-class and multi-class classification tasks. Xception achieved 99.09% and 95.35% accuracy in binary-class and multi-class classification, respectively, making it the second best-performing model. The VGG-19 model had the lowest performance in the context of multi-class classification,

TABLE I: Model performance metrics (binary-class classification)

Model	F1 Score	Recall	Precision	Accuracy
DenseNet-121	0.9815	0.9810	0.9800	0.9819
Custom CNN	0.9518	0.9545	0.9530	0.9553
Inception V3	0.9831	0.9804	0.9859	0.9913
Xception	0.9903	0.9906	0.9900	0.9909
VGG-19	0.9755	0.9765	0.9760	0.9769

TABLE II: Model performance metrics (multi-class classification)

Model	F1 Score	Recall	Precision	Accuracy
DenseNet-121	0.9481	0.9477	0.9404	0.9477
Custom CNN	0.9314	0.9321	0.9387	0.9321
Inception V3	0.9764	0.9766	0.9774	0.9766
Xception	0.9534	0.9535	0.9536	0.9535
VGG-19	0.9104	0.9110	0.9116	0.9110

whereas the Custom CNN model exhibited the lowest performance in the case of binary-class classification. In all cases, the validation accuracy shows small oscillations at first, but after 50 epochs, these oscillations become less pronounced and point toward the ideal solution. Fig. 3(a), 3(b), 3(c), 3(d), 3(e), and 3(f) depict the accuracy, loss curves, and Confusion Matrix of Inception V3 in both binary and multiclass classification, respectively.

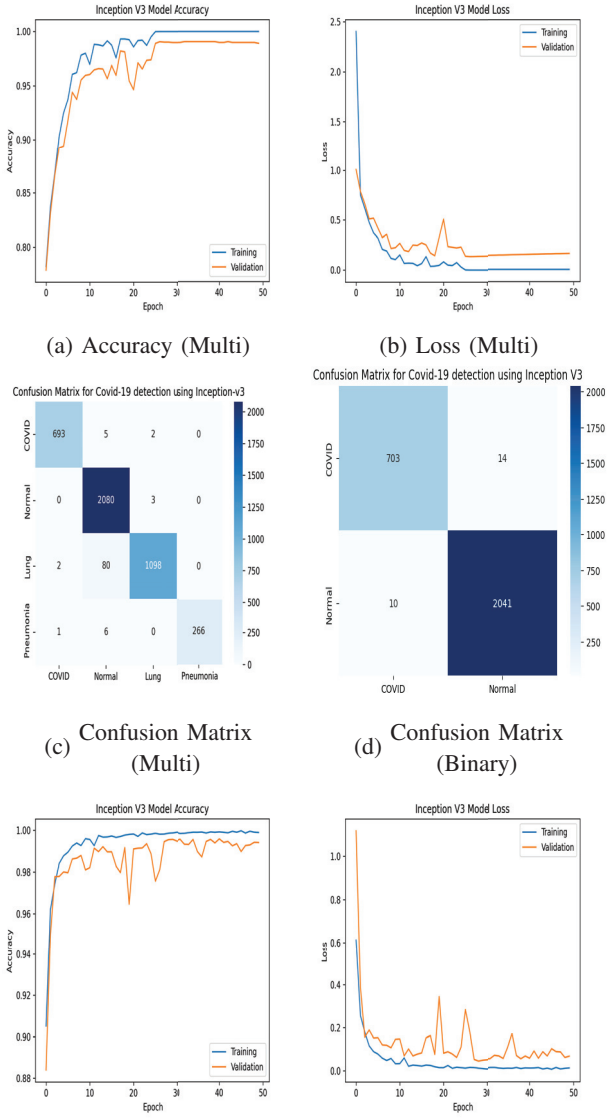
Table III compares the accuracies of the Inception V3 with relevant studies, showcasing better performance than its counterparts. In contrast to prior studies concentrating on a limited set of three classes in small datasets without detailing the detection technique, this study comprises of a vast dataset, employs four-class classification, and integrates Grad-CAM visualization.

TABLE III: Comparative analysis among modified models and existing approaches

Method	**Classes**	**Accuracy**
Sethi et al. [7]	(Three-class)	95.33%
Tang et al. [8]	(Three-class)	94.1%
Senan et al. [9]	(Four-class)	95%
Dubey et al. [10]	(Binary-class)	95%
Ismael [11]	(Binary-class)	94.7%
Kokilavani et al. [12]	(Binary-class)	97.68%
Inception V3	(Binary-class)	99.13%
Inception V3	(Four-class)	97.66%

E. Grad-CAM Visualization Result Analysis

Fig. 4 illustrates the initial image and the resulting output of the Inception V3 classification task for each of the four classes. The overlay heatmap on the input image is shown in the output image. Notably, a Jet color scheme (Red, Blue, Green) is employed, with each color representing values corresponding to the extracted features. Within this color scheme, blue tones indicate lower values, signifying the absence of feature extraction for a particular class. On the other hand, yellow and green tones depict moderate values, suggesting relatively limited feature extraction. Finally, red and dark red tones represent higher values, indicating that features in the corresponding region characterize the specific class.



(a) Accuracy (Multi) (b) Loss (Multi)

(c) Confusion Matrix (Multi) (d) Confusion Matrix (Binary)

(e) Accuracy (Binary) (f) Loss (Binary)

Fig. 3: Accuracy, Loss, and Confusion Matrix of Inception V3

V. Conclusion

In this work, an extensive comparative analysis of deep learning techniques for automated COVID-19 detection from chest X-ray images is presented. A large COVID dataset of over 21,000 images was utilized in this comparative study. The models, comprising DenseNet-121, Custom CNN, Inception V3, Xception, and VGG-19, exhibited binary-class classification accuracies of 98.19%, 95.53%, 99.13%, 99.09%, and 97.69%, respectively. Correspondingly, their multi-class classification accuracies aligned with these results, also yielding accuracies of 94.77%, 93.21%, 97.66%, 95.35%, and 91.10%. Inception V3 demonstrated superior performance

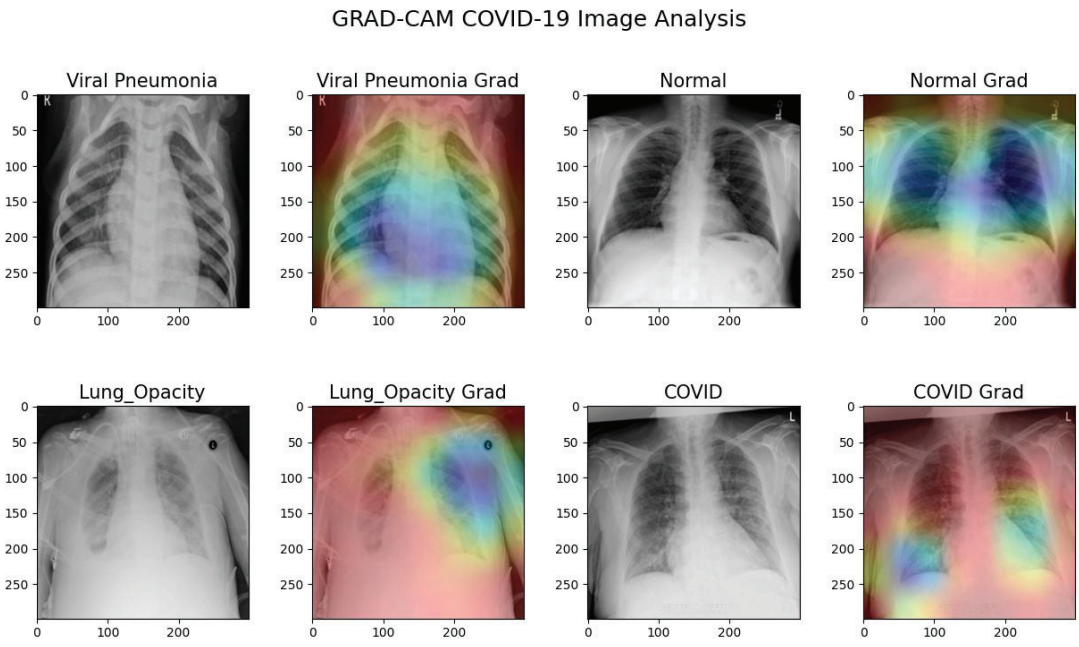


Fig. 4: Grad-CAM Visualization

in both multi-class and binary-class classification across all performance metrics. The proposed fine-tuned method of the Inception-V3 model exhibited notable efficacy on a large and diverse dataset, contrasting with recent studies on COVID-19 identification that primarily focused on smaller datasets. In order to assess the transparency of COVID-19 detection, the Grad-CAM approach was used to identify the most significant regions in the CXR images. In future research, the transition from 2D-CNN to 3D-CNN by incorporating an attention module to accurately identify COVID-19 may be investigated.

.

REFERENCES

[1] A. Cossarizza, S. De Biasi, G. Guaraldi, M. Girardis, C. Mussini, M. C.-. W. Group *et al.*, "Sars-cov-2, the virus that causes covid-19: cytometry and the new challenge for global health," *Cytometry*, vol. 97, no. 4, p. 340, 2020.

[2] D. Baud, X. Qi, K. Nielsen-Saines, D. Musso, L. Pomar, and G. Favre, "Real estimates of mortality following covid-19 infection," *The Lancet infectious diseases*, vol. 20, no. 7, p. 773, 2020.

[3] W. Wang, Y. Xu, R. Gao, R. Lu, K. Han, G. Wu, and W. Tan, "Detection of sars-cov-2 in different types of clinical specimens," *Jama*, vol. 323, no. 18, pp. 1843–1844, 2020.

[4] M. Annarumma, S. J. Withey, R. J. Bakewell, E. Pesce, V. Goh, and G. Montana, "Automated triaging of adult chest radiographs with deep artificial neural networks," *Radiology*, vol. 291, no. 1, pp. 196–202, 2019.

[5] L. Huang, R. Han, T. Ai, P. Yu, H. Kang, Q. Tao, and L. Xia, "Serial quantitative chest ct assessment of covid-19: a deep learning approach," *Radiology: Cardiothoracic Imaging*, vol. 2, no. 2, p. e200075, 2020.

[6] Y. Fang, H. Zhang, J. Xie, M. Lin, L. Ying, P. Pang, and W. Ji, "Sensitivity of chest ct for covid-19: comparison to rt-pcr," *Radiology*, vol. 296, no. 2, pp. E115–E117, 2020.

[7] P. K. Sethy and S. K. Behera, "Detection of coronavirus disease (covid-19) based on deep features," 2020.

[8] S. Tang, C. Wang, J. Nie, N. Kumar, Y. Zhang, Z. Xiong, and A. Barnawi, "Edl-covid: Ensemble deep learning for covid-19 case detection from chest x-ray images," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 9, pp. 6539–6549, 2021.

[9] E. M. Senan, A. Alzahrani, M. Y. Alzahrani, N. Alsharif, and T. H. Aldhyani, "Automated diagnosis of chest x-ray for early detection of covid-19 disease," *Computational and Mathematical Methods in Medicine*, vol. 2021, pp. 1–10, 2021.

[10] A. K. Dubey and K. K. Mohbey, "Enabling ct-scans for covid detection using transfer learning-based neural networks," *Journal of Biomolecular Structure and Dynamics*, vol. 41, no. 6, pp. 2528–2539, 2023.

[11] A. M. Ismael and A. Şengür, "Deep learning approaches for covid-19 detection based on chest x-ray images," *Expert Systems with Applications*, vol. 164, p. 114054, 2021.

[12] S. V. Kogilavani, J. Prabhu, R. Sandhiya, M. S. Kumar, U. Subramaniam, A. Karthick, M. Muhibbullah, S. B. S. Imam *et al.*, "Covid-19 detection based on lung ct scan using deep learning techniques," *Computational and Mathematical Methods in Medicine*, vol. 2022, 2022.

[13] M. E. Chowdhury, T. Rahman, A. Khandakar, R. Mazhar, M. A. Kadir, Z. B. Mahbub, K. R. Islam, M. S. Khan, A. Iqbal, N. Al Emadi *et al.*, "Can ai help in screening viral and covid-19 pneumonia?" *Ieee Access*, vol. 8, pp. 132 665–132 676, 2020.

[14] R. Tharsanee, R. Soundariya, A. S. Kumar, M. Karthiga, and S. Sountharrajan, "Deep convolutional neural network–based image classification for covid-19 diagnosis," in *Data Science for COVID-19*. Elsevier, 2021, pp. 117–145.

[15] V. Lakkavaram, L. Raghuveer, C. Satish Kumar, G. Sai Sri, and S. Habeeb, "A review on practical diagnostic of tomato plant diseases," *Suraj Punj J Multidiscip Res*, vol. 9, pp. 432–435, 2019.

[16] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2818–2826.

[17] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-cam: Visual explanations from deep networks via gradient-based localization," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 618–626.