Deep Learning-Based COVID-19 Detection from Chest X-ray Images: A Comparative Study

Duc M Cao¹, Md Shahedul Amin², Md Tanvir Islam³, Sabbir Ahmad⁴, Md Sabbirul Haque⁵, Minister Rahman⁷ and Tahera Koli⁸

¹Department of Economics, University of Tennessee, Knoxville, TN, USA
²Department of Finance & Economics, University of Tennessee, Chattanooga, TN, USA
³Department of Computer Science, Monroe College, New Rochelle, New York, USA
⁴Department of Mathematical Science, University of South Dakota, Vermillion, USA
⁵IEEE Professional Community, IEEE
⁶Department of Professional Security Studies, New Jersey City University, Jersey City, New Jersey, USA
⁷Department of Computer & Info Science, Gannon University, Erie, Pennsylvania, USA

Corresponding Author: Md Sabbirul Haque, E-mail: sabbir465@ieee.org

ABSTRACT

The COVID-19 pandemic, caused by the SARS-CoV-2 virus, has rapidly spread across the globe, leading to a significant number of illnesses and fatalities. Effective containment of the virus relies on the timely and accurate identification of infected individuals. While methods like RT-PCR assays are considered the gold standard for COVID-19 diagnosis due to their accuracy, they can be limited in their use due to cost and availability issues, particularly in resource-constrained regions. To address this challenge, our study presents a set of deep learning techniques for predicting COVID-19 using chest X-ray images. Chest X-ray imaging has emerged as a valuable and cost-effective diagnostic tool for managing COVID-19 because it is non-invasive and widely accessible. However, interpreting chest X-rays for COVID-19 detection can be complex, as the radiographic features of COVID-19 pneumonia can be subtle and may overlap with those of other respiratory illnesses. In this research, we evaluated the performance of various deep learning models, including VGG16, VGG19, DenseNet121, and ResNet50, to determine their ability to differentiate between cases of coronavirus pneumonia and non-COVID-19 pneumonia. Our dataset comprised 4,649 chest X-ray images, with 1,123 of them depicting COVID-19 cases and 3,526 representing pneumonia cases. We used performance metrics and confusion matrices to assess the models’ performance. Our study’s results showed that DenseNet121 outperformed the other models, achieving an impressive accuracy rate of 99.44%.

KEYWORDS

COVID-19, X-RAY, Pneumonia.

ARTICLE INFORMATION

ACCEPTED: 05 November 2023 PUBLISHED: 28 November 2023 DOI: 10.32996/jcsts.2023.5.4.13

1. Introduction

The COVID-19 pandemic, caused by the rapid spread of the SARS-CoV-2 virus, has had a profound impact on global public health. The urgent need for timely and accurate identification of infected individuals has led to various diagnostic approaches, with RT-PCR assays being the gold standard for COVID-19 detection. However, the availability and cost of these tests have presented challenges, particularly in resource-constrained regions. To address this issue, the study presented here explores the potential of deep learning techniques for predicting COVID-19 using chest X-ray images. Chest X-ray imaging offers a non-invasive and cost-effective diagnostic tool for managing COVID-19, making it an attractive alternative.

Copyright: © 2023 the Author(s). This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) 4.0 license (https://creativecommons.org/licenses/by/4.0/). Published by Al-Kindi Centre for Research and Development, London, United Kingdom.
Nevertheless, interpreting chest X-rays for COVID-19 detection is a complex task, as the radiographic features of COVID-19 pneumonia can be subtle and overlapping with other respiratory conditions. In response to this challenge, the study evaluates the performance of various deep learning models, including VGG16, VGG19, DenseNet121, and ResNet50, to distinguish between COVID-19 pneumonia and non-COVID-19 pneumonia. The research uses a substantial dataset of 4,649 chest X-ray images, of which 1,123 depict COVID-19 cases, and 3,526 represent pneumonia cases. Key performance metrics and confusion matrices are employed to assess the models’ effectiveness.

The COVID-19 pandemic has posed significant challenges for the global healthcare system, necessitating the urgent need to identify affected people accurately and quickly. This study focuses on using chest X-rays to find COVID-19 in patients with symptoms or illnesses that have been diagnosed. To control the transmission of the virus, it is essential to identify COVID-19-affected individuals as soon as possible. The most effective approach for COVID-19 detection is RT-PCR, although, in areas with limited resources, it may not always be available or economical. The identification of distinctive radiographic characteristics connected to COVID-19 pneumonia using chest X-ray imaging has been proposed as an alternative, affordable, and non-invasive approach to COVID-19 detection. However, the sensitivity and specificity of chest X-rays in COVID-19 detection are highly variable, and the interpretation of radiographic findings requires expertise and may be subject to interobserver variability. Deep learning, a subset of machine learning, has shown significant promise. In image analysis, it has the potential to aid in the detection of COVID-19 pneumonia on chest X-ray images. Deep learning algorithms can automatically learn and identify image patterns and features without prior knowledge of specific radiographic features associated with COVID-19 pneumonia. Detecting COVID-19 through radiological means can be challenging because its appearance on radiographs is like that of viral pneumonia. Distinguishing between the two requires specific expertise, which may not always be readily available, especially with the high number of suspected cases daily. As a solution, (khan, 2022,2023,2023,2023) have turned to automation and machine learning to bridge this gap. In this study, we aim to review the current state of knowledge on the use of deep learning algorithms in the detection of COVID-19 pneumonia on chest X-ray images. We will discuss the performance of deep learning algorithms and their potential clinical implications for COVID-19 detection in resource-limited settings. The results of this study could have important consequences for the creation of precise and affordable methods for detecting COVID-19. The use of deep learning algorithms in radiological imaging could significantly enhance the accuracy of diagnosis for this virus.

The findings of this study reveal that DenseNet121 outperforms the other models, achieving an impressive accuracy rate of 99.44%. These results have important implications for improving the accuracy and efficiency of COVID-19 diagnosis, particularly in regions with limited resources and the need for rapid screening. Furthermore, the use of deep learning algorithms in radiological imaging may significantly enhance the precision of COVID-19 detection, marking a significant step forward in the ongoing battle against the pandemic. This article explores the study’s methodology, results, and implications, shedding light on the potential of deep learning in the field of medical image analysis and COVID-19 diagnosis.

2. Literature review
The study by Taher et al. (2022) focuses on the use of chest X-ray (CXR) pictures to categorize instances as viral pneumonia, normal pneumonia, or Covid-19 pneumonia. The technique requires computing the Discrete Cosine Transform (DCT) for each smaller subblock created from the CXR pictures. The DCT makes it possible to create compressed copies of each CXR image, which is helpful for energy compaction. The final feature vectors’ dimension is decreased by using average pooling windows after the images have been compressed. A small number of spectral DCT components that are chosen as features for each image make up these vectors. The photos are split into three categories using a multilayer artificial neural network. The accuracy of the suggested method is 95% on average.

The study by Jonayet et al. (2023) builds upon the existing literature by employing state-of-the-art deep learning models, a substantial dataset, and comprehensive evaluation methods. The notable performance of DenseNet121 in accurately identifying COVID-19 cases from chest X-ray images highlights the potential of deep learning in aiding medical diagnosis, particularly in the context of the COVID-19 pandemic. Further research and validation on larger and more diverse datasets will be essential to confirm the generalizability of these findings to clinical practice.

Tora man et al. (2020) introduced a novel artificial neural network, the Convolutional CapsNet, for the detection of COVID-19 using chest X-ray images. This new approach is developed with the aim of providing both rapid and accurate diagnostics for COVID-19, offering binary classification (COVID-19 vs. No-Findings) and multi-class classification (COVID-19 vs. No-Findings vs. Pneumonia). The reported results indicate an accuracy of 97.24% for binary classification and 84.22% for multi-class classification. The study suggests that this approach may assist physicians in diagnosing COVID-19 and potentially enhance diagnostic performance while also serving as a fast-screening tool.
Abraham et al. (2020) present an innovative approach to COVID-19 detection, combining multiple pre-trained CNNs with feature selection and a classifier. The reported high AUC and accuracy values in the two datasets support the effectiveness of this method, particularly in comparison to using single CNNs. The study’s findings contribute to the ongoing efforts to develop accurate and reliable tools for COVID-19 diagnosis from medical images. However, this work takes a lot of time and is expensive for the medical sector.

Aslan et al. (2022) research paper contributes to the growing body of work that utilizes machine learning and deep learning techniques to diagnose COVID-19 from medical images. The use of various CNN models, hyperparameter optimization, ANN-based image segmentation, and high classification accuracy underscore the potential of these methods in aiding COVID-19 diagnosis and clinical decision-making.

The article by Rasheed et al. (2021) explored the potential of machine learning techniques to automatically diagnose coronavirus infections with a high degree of accuracy using X-ray images. They specifically chose two widely used classifiers, logistic regression (LR) and convolutional neural networks (CNN), aiming to create a fast and efficient system. Furthermore, the researchers investigated a dimensionality reduction strategy based on principal component analysis (PCA) to accelerate the learning process and enhance classification accuracy by selecting the most distinctive features. Deep learning methods typically require a substantial amount of training data, but for COVID-19 X-ray images, an adequate number of labeled training samples was not readily available. To address this challenge, the study employed data augmentation techniques involving generative adversarial networks (GAN) to expand the training dataset and mitigate overfitting issues. The authors used an openly accessible dataset and incorporated GAN to amass a total of 500 X-ray images for their study. Both LR and CNN models yielded promising results in identifying COVID-19 patients. In particular, the LR and CNN models demonstrated an overall accuracy of 95.2% to 97.6% without PCA and 97.6% to 100% with PCA for the identification of positive COVID-19 cases.

The existing research has concentrated solely on datasets without providing a sustainable model for evaluating the performance across a broader range of datasets. Additionally, none of the previous studies have undertaken comparative analyses with other current models to gauge the competitiveness of the model in question. In our study, we emphasized these two aspects to draw attention to this important health concern.

To address this challenge, the study employed data augmentation techniques involving generative adversarial networks (GAN) to expand the training dataset and mitigate overfitting issues. The authors used an openly accessible dataset and incorporated GAN to amass a total of 500 X-ray images for their study. Both LR and CNN models yielded promising results in identifying COVID-19 patients. In particular, the LR and CNN models demonstrated an overall accuracy of 95.2% to 97.6% without PCA and 97.6% to 100% with PCA for the identification of positive COVID-19 cases.

3. Methodology

3.1 VGG 16

This section offers an overview of the VGG16 deep-learning model employed in this study. VGG16 comprises a total of 16 layers, with 13 of them being convolutional layers, and the remaining three are fully connected layers. It is designed to process input images with a predetermined size of 224 × 224 pixels in RGB format. The image dimensions are progressively reduced through max-pooling operations, and at the conclusion of the network, a SoftMax classifier is typically applied after the last fully connected layer. However, for the purposes of this investigation, a custom-designed classifier was utilized in lieu of the final fully connected layer with SoftMax activation, as depicted in Figure 1.
3.2 VGG 19

The VGG19 model that was utilized in the research. The VGG19 model employed in this study consists of a total of 19 layers and is designed to process RGB images with dimensions of 224 x 224 pixels. Within these 19 layers, there are 16 convolutional layers and 3 fully connected layers. The dimension of the output is reduced through max-pooling operations, and at the end of the network, a SoftMax classifier is typically applied after the final fully connected layer. However, in the context of this study, a custom-designed classifier was used instead of the standard final fully connected layer with SoftMax activation, as illustrated in Figure 2.

A dataset for pneumonia was formed by randomly selecting an equal number of samples for both training and validation purposes, ensuring a balanced distribution. In this setup, 20% of these samples were allocated for validation, while the remaining 80% were designated for training. To enhance the quality of testing results and prevent the model from becoming overly biased towards pneumonia cases, the remaining samples were held in reserve [Shazia et al. 2021, Elaziz et al. 2020]. The rationale behind this approach was to achieve optimal accuracy, as prior research has consistently emphasized the importance of having a balanced training set to obtain precise and reliable results [Haque et al., 2023].

3.3 Dense Net 121

The DenseNet121 model, a powerful deep learning architecture, requires input in the form of RGB images with precise dimensions of 224x224 pixels. Within this model, an extensive network of parameters is distributed across 121 layers, totaling more than 8 million parameters. The model’s distinctive feature lies in its Dense Blocks, fundamental components that play a pivotal role in retaining the feature map dimensions while dynamically adjusting the quantity of filters. These Dense Blocks facilitate the efficient learning of intricate features from the input data. To ensure the effective processing of information, Transition layers are strategically positioned between these Dense Blocks [Zheng et al. 2020, Das et al. 2021]. These Transition layers incorporate batch normalization techniques to facilitate the crucial process of down sampling, optimizing the flow of data through the network.
What sets this study apart is the innovative approach taken in the utilization of the DenseNet121 model. Instead of adhering to the conventional approach of employing a final fully connected layer equipped with SoftMax activation for classification, a specially tailored classifier was implemented. This customized classifier, as depicted in Figure 3, is an integral component of this research, aimed at enhancing the model’s performance in the context of COVID-19 [Qaid et al. 2021, Miah et al. 2023] and pneumonia categorization. By making these thoughtful adaptations, this study contributes to the ongoing effort to improve the accuracy and effectiveness of COVID-19 diagnosis, particularly in the realm of medical image analysis.

### 3.4 Resnet 50

The ResNet50 architecture, a prominent member of residual networks, is characterized by the inclusion of a Max-Pool layer, an Average Pool layer, and an impressive total of 48 Convolutional Layers. This design offers a robust foundation for deep learning applications. Within ResNet50 [Ezzat et al. 2021, Khan et al. 2021], each convolution block consists of three convolutional layers and an identification block is also integrated. The model comprises a substantial parameter space, with over 23 million distinct parameters that can be fine-tuned during training. In the context of this study, as depicted in Figure 4, specific modifications were made to the ResNet50 model to tailor it for the precise classification of COVID-19 and pneumonia cases. These adaptations represent a crucial step in the customization of the model to address the specific diagnostic challenges posed by these medical conditions, ultimately contributing to more accurate and reliable results in COVID-19 detection.

### 3.5 Model Training and Validation

In this comprehensive study, a total of four deep learning models, namely VGG16, VGG19, DenseNet121, and ResNet50, were harnessed to address the critical task of COVID-19 detection using medical images. To ensure uniformity and consistency in the
dataset, all images were meticulously scaled down to a standardized resolution of 224 × 224 pixels. The study harnessed the power of cutting-edge technology, employing TensorFlow 2.4 and the Keras API to design and implement the convolutional neural network (CNN) [Haque et al. 2023, Ukwuoma et al. 2022] algorithm. The computational muscle behind the model training was provided by a robust 12 GB NVIDIA Tesla K80 GPU [Attallah 2022, Rajpal et al. 2022, Yang et al. 2022], enabling the efficient processing of a substantial volume of data. The performance assessment of these deep learning models was conducted throughout the training process, with a specific focus on their capacity to accurately predict ground truth probabilities. This evaluation was carried out using the categorical cross-entropy loss function [Islam et al. 2021], a widely accepted metric for gauging model performance and fine-tuning its capabilities.

For the study dataset, a meticulous collection of images was performed, drawing from multiple publicly available databases. A total of 3,525 images depicting cases of pneumonia, along with 1,123 images representing COVID-19 cases, were strategically curated for the purpose of model training, validation, and testing. To ensure consistent processing, all images underwent standardization through resizing, unifying their dimensions to 224 by 224 pixels from their original, diverse range of sizes. To preserve the integrity of the data and facilitate comprehensive evaluation, 10% of the COVID-19 samples were randomly selected for testing, while the remaining samples were thoughtfully divided into an 80% training set and a 20% validation set. This rigorous and systematic approach allowed the study to harness the full potential of deep learning models in the crucial task of COVID-19 detection, contributing to the development of accurate and effective diagnostic tools for addressing this pressing global health challenge.

4. Result and Discussion
Table 3 and Figs. 5, 6, 7, and 8 demonstrate the accuracy and loss values of each fine-tuned model during the training and validation processes. ResNet50 achieved the minimum validation loss in just 4 epochs, with a validation accuracy of 99% or higher. According to the findings, the models can immediately recognize the distinctive features of pneumonia and COVID-19. However, it was shown that DenseNet121 and ResNet50 had the highest training accuracy when accuracy and loss were considered.

<table>
<thead>
<tr>
<th>Model</th>
<th>Accuracy (%)</th>
<th>Precision (%)</th>
<th>Recall (%)</th>
<th>F1 Score (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>VGG 15</td>
<td>99.22</td>
<td>99.14</td>
<td>99.07</td>
<td>99.01</td>
</tr>
<tr>
<td>VGG 19</td>
<td>98.29</td>
<td>99.08</td>
<td>99.1</td>
<td>99.10</td>
</tr>
<tr>
<td>Dense Net 121</td>
<td>99.44</td>
<td>98.70</td>
<td>99.30</td>
<td>99.48</td>
</tr>
<tr>
<td>Resnet 50</td>
<td>99.16</td>
<td>99.40</td>
<td>99.21</td>
<td>99.24</td>
</tr>
</tbody>
</table>

If we consider the combination of both training and validation accuracy and loss, then the DenseNet121 model achieved the lowest validation loss and the highest validation accuracy, indicating that it may have performed the best among the four models. It should be underlined that more testing and verification are required to ensure the comparative efficacy of the models shown in Table 1 before they can be used.
In terms of accuracy, DenseNet121 outperformed the other models, with the highest accuracy rate of 99.44%, showcasing its exceptional ability to correctly classify COVID-19 cases from medical images. VGG16 also demonstrated strong accuracy at 99.22%.

When considering precision, recall, and F1 Score, DenseNet121 continued to excel, demonstrating precision, recall, and F1 Score values of 98.70%, 99.30%, and 99.48%, respectively. ResNet50 is closely followed with precision, recall, and F1 Score values above 99%, highlighting its remarkable performance in identifying COVID-19 cases. VGG16 and VGG19 exhibited impressive precision and recall values above 99%, indicating their strong capability in correctly identifying COVID-19 cases. However, their F1 Score values were slightly lower than DenseNet121 and ResNet50.

In summary, DenseNet121 and ResNet50 stand out as the top-performing models in terms of both accuracy and comprehensive performance metrics, making them highly promising tools for the accurate detection of COVID-19 from medical images. VGG16 and VGG19 also demonstrated strong performance, with a particular strength in precision and recall. These results emphasize the potential of deep learning models in enhancing the accuracy and efficiency of COVID-19 diagnosis, contributing to the ongoing efforts to combat the pandemic.
Fig 5: VGG 16 model accuracy graph

Fig 6: VGG 19 model accuracy graph

Fig 7: Dense Net model accuracy graph

Fig 8: Resnet model accuracy graph
5. Conclusion and Future Work
Overall, DenseNet121 ability to accurately detect COVID-19 through chest X-rays is a testament to the power of machine learning in healthcare and highlights the importance of continued research and innovation in this field. While further research is needed to validate the effectiveness of DenseNet121 in detecting COVID-19 on a larger scale, the results so far are very promising. The use of DenseNet121 in detecting COVID-19 through chest X-rays has the potential to revolutionize how we diagnose and manage this disease, particularly in areas with limited access to testing and medical resources.

In this study, we harnessed the power of deep learning models, including VGG16, VGG19, DenseNet121, and ResNet50, to tackle the critical challenge of COVID-19 detection from medical images. The results have been highly promising, showcasing the potential of these models in accurately and efficiently identifying COVID-19 cases. DenseNet121 emerged as the top-performing model, achieving an impressive accuracy rate of 99.44% and excelling in precision, recall, and F1 Score. ResNet50 closely followed, demonstrating strong performance in all metrics. VGG16 and VGG19 also displayed robust capabilities, particularly in precision and recall.

These findings hold significant implications for the field of medical image analysis and COVID-19 diagnosis. The accurate and rapid identification of COVID-19 cases is crucial in the ongoing global battle against the pandemic. Deep learning models, as demonstrated in this study, offer a powerful tool for enhancing diagnostic accuracy and efficiency, particularly in resource-constrained regions where alternative methods such as RT-PCR may face limitations.

The results of this study open exciting possibilities for the application of deep learning in COVID-19 diagnosis, and future work should focus on refining these models for real-world clinical use and expanding their impact on global healthcare.

Funding Statement: The authors received no specific funding for this study.

Conflicts of Interest: The authors declare no conflicts of interest to report regarding the present study.

References


