Deep Learning-Powered Mobile App for Fast and Accurate COVID-19 Detection from Chest X-rays

Rahhal Errattahi, Fatima Zahra Salmam, Mohamed Lachgar, Asmaa El Hannani, Abdelhak Aqqal
Chouaib Doukkali University of El Jadida, National School of Applied Sciences, Laboratory of Information Technologies, El Jadida, Morocco

Abstract—The COVID-19 pandemic has imposed significant challenges on healthcare systems globally, necessitating swift and precise screening methods to curb transmission. Traditional screening approaches are time-consuming and prone to errors, prompting the development of an innovative solution - a mobile application employing machine learning for automated COVID-19 screening. This application harnesses computer vision and deep learning algorithms to analyze X-ray images, rapidly detecting virus-related symptoms. This solution aims to enhance the accuracy and speed of COVID-19 screening, particularly in resource-constrained or densely populated settings. The paper details the use of convolutional neural networks (CNNs) and transfer learning in diagnosing COVID-19 from chest X-rays, highlighting their efficacy in image classification. The trained model is deployed in a mobile application for real-world testing, aiming to aid healthcare professionals in the battle against the pandemic. The paper provides a comprehensive overview of the background, methodology, results, and the application's architecture and functionalities, concluding with avenues for future research.

Keywords—COVID-19 diagnosis; computer vision; deep learning; X-ray images; mobile application.

I. INTRODUCTION

The COVID-19 pandemic has brought unprecedented challenges to healthcare systems worldwide. Early and efficient screening is crucial to prevent the spread of the virus. However, traditional screening methods such as manual assessments and laboratory tests are time-consuming and often have a high error rate [1]. Therefore, there is an urgent need for a more effective and rapid screening solution. To address this challenge, we propose an innovative solution - a mobile application that leverages machine learning for automatic screening for COVID-19. The application uses computer vision and deep learning algorithms to analyze X-ray images of patients and detect symptoms of the virus in real-time. The proposed application has the potential to improve the accuracy of screening results and reduce the time required for patient diagnosis by automating the screening process.

The proposed solution aims to overcome the limitations of traditional COVID-19 screening methods, especially in settings with limited resources or high population density. By leveraging the capabilities of machine learning, the screening process is broken down into smaller components that can be easily updated and maintained. The ultimate objective of this solution is to enhance the speed and accuracy of COVID-19 screening and support healthcare professionals in their efforts to combat the pandemic.

Over the last few decades, Computer Vision (CV) and Deep Learning (DL) have been widely used in various Computer Aided Diagnosis (CAD) applications especially in radiology, where CAD system is used to help radiologists analyze and interpret medical images like mammography, X-ray, CT-scan, etc. Recently, radiologists have noted that chest X-ray images show distinctive marks of pneumonia caused by COVID-19 [2]. Several studies investigated automatic COVID-19 detection from chest X-ray images [2], [3]. However, despite the promising results achieved in the reported works, the deployment of the developed DL models has not been comprehensively studied [4]. Deploying DL models in mobile applications will allow the scientific community to experiment those models in real conditions by reaching a large population. On the other hand, the proposed application will be highly portable, simple to use, and inexpensive. Thus, such software will allow primary screening for COVID-19 in areas where medical expertise and test kits are not available or insufficient (rural and remote areas, hot spots, and clusters of at-risk populations).

Our study had the objective of assessing the use of deep learning, in particular convolutional neural networks (CNNs), in the context of chest X-rays for patients suspected of having COVID-19 pneumonia. This included its capability for directly diagnosing COVID-19 and its ability to distinguish COVID-19 from other community-acquired types of pneumonia. The overarching goal was to introduce an effective tool that could aid in COVID-19 pneumonia diagnosis, either by offering a second opinion to radiologists or by providing a preliminary assessment when a radiologist is not readily accessible.

To achieve these objectives and considering the existing body of research that consistently demonstrates the superior performance of CNNs in image classification tasks, we undertook the training and cross-validation of an Inception-V3 based architecture. We assessed the model’s performance using two distinct sets of independent datasets. This model was then deployed in a mobile application to assist in the differentiation of COVID-19 cases.

The paper is structured as follows: Section II presents background and related works. Materials and methods are presented in Section III followed by results and discussion in Section IV. Section V describes the software, including its architecture and functionalities. Finally, Section VI concludes the paper and discusses future directions for research.

II. BACKGROUND AND RELATED WORKS

Multiple studies have been conducted to explore the automatic detection of COVID-19 from CT-scan and chest X-ray images, showing encouraging outcomes. The initial set
of research focuses on CT-scan images [3], [5] which can effectively identify to detect deep-rooted changes in lung tissue. Despite their promising detection rates, these methods involve higher costs and expose patients to potentially harmful radiation doses [6].

The second group of studies revolves around the investigation of chest X-ray radiographs [7], [8], [9]. These images are the most commonly used radiographs in the medical field. Besides detecting changes in the lungs, they are valuable for swiftly assessing breathing or heart-related issues without causing discomfort to the patient.

Due to these advantages, our paper specifically focuses on the second group of approaches that utilize chest X-ray images. This type of medical images is prioritized because it offers the lowest cost and are widely accessible in medical institutions for detecting thoracic pathologies.

To date, numerous methods have been developed and introduced for the automatic detection of COVID-19 using X-ray Images. In our prior publication [10] we presented a comprehensive summary of the most significant works in this area. It is evident that the majority of the proposed approaches rely on Transfer Learning (TL), where models are pre-trained on external large datasets and adapted to the target task. The use of this training technique is mainly due to the lack of COVID-19 data available for the community.

Recently, Cohen et al. [11] conducted a study investigating the cross-domain performance using DenseNet [12]. They found that the model’s performance was unsatisfactory when evaluated on datasets from different sources. On the other hand, several researchers have utilized various datasets from multiple origins. For example, Narin et al. [9] proposed a pre-trained convolutional neural network based on Res-Net50 with a binary classification approach (COVID-19 and Normal). The dataset they used was obtained from the open-source GitHub repository proposed by Cohen et al. [13] and the “ChestX-ray8” database proposed by Wang et al. [14]. For evaluation, they employed a 5-fold cross-validation method and achieved an accuracy of 98%.

Apostolopoulos et al. [8] conducted a study where they gathered X-ray images from different sources [13], [15]. They assessed the performance of state-of-the-art convolutional neural network architectures using a Transfer Learning strategy. They demonstrated that VGG19 and MobileNetV2 could extract significant biomarkers related to COVID-19 with accuracies of 98.75% and 96.78%, respectively, using binary classification.

In another work, Wang et al. [16] introduced a tailored convolutional network architecture called COVID-Net for detecting COVID-19 cases from chest X-ray images. They collected a dataset of 13,975 chest X-ray images from 13,870 patient cases to train and evaluate their model, achieving an accuracy of 93.3%. Meanwhile, Zhang et al. [17] conducted their experiments on two combined datasets [13], [14]. They proposed a deep learning model for COVID-19 detection, comprising a backbone network for high-level feature extraction, a classification head, and an anomaly detection head. Their model achieved a sensitivity of 96% and a specificity of 70.65%.

Consequently, existing studies have focused on using transfer learning for the COVID-19 detection from medical images. Moreover, they present potential limitations, including:

- **Limited Data Availability:** Many existing works suffer from a scarcity of labeled data, making it challenging to train accurate and robust models. COVID-19 X-ray images are relatively rare, and obtaining large datasets for training can be difficult;
- **Data Imbalance:** The datasets used for training often suffer from class imbalance, with a disproportionate number of negative cases (non-COVID-19) compared to positive cases (COVID-19). This can lead to biased models that are better at detecting the majority class but perform poorly on COVID-19 cases;
- **Generalization Issues:** Models developed in one region or using specific equipment may not generalize well to other regions or types of X-ray machines. There can be significant variability in how X-ray images are captured and processed;
- **Evaluation Bias:** Some studies may suffer from evaluation bias, where models are tested on the same datasets from which they were trained or on datasets that are too similar. This can result in overly optimistic performance metrics;
- **Limited Clinical Validation:** While models may achieve high accuracy in identifying COVID-19 from X-ray images in research settings, their clinical validation and real-world performance may be less certain. More rigorous clinical testing and validation are needed;
- **Ethnic and Age Bias:** Some models may not perform equally well across different ethnic groups or age ranges. Bias in the training data can lead to disparities in model performance;
- **Overfitting and Noise:** Overfitting to noise or irrelevant features in the data can be a problem, particularly when dealing with small datasets. Models may learn to exploit artifacts in the images rather than true COVID-19 patterns;
- **Resource Intensive:** Deep learning models for image classification, particularly those used in healthcare can be computationally intensive and require substantial hardware resources, limiting their practical deployment in resource-constrained settings;
- **Continuous Updates:** The evolving nature of the COVID-19 virus and the emergence of new variants may require continuous updates and retraining of models to maintain their effectiveness.

Addressing these limitations is crucial for the development of reliable and clinically viable automatic COVID-19 detection solutions from X-ray images. Along this direction, and to overcome these limitations, the main contributions of this work can be summarized as follows:

- The experiments in this paper utilize a fairly large and balanced dataset in terms of classes, collected from
different data sources, including various ethnic groups and age ranges;

- The model’s performance was evaluated using two distinct sets of independent datasets to assess its generalization capabilities.
- The CLAHE histogram equalization method was applied to enhance data quality;
- The trained model was deployed in a mobile/web application, offering the potential for rigorous clinical testing and validation. This application will also facilitate the collection of other clinical data related to COVID-19.

III. MATERIALS AND METHODS

A. Dataset

In order to assess the model's predictive ability and generalization capability, we established two distinct datasets. The first dataset gathers 3429 X-ray images from the COVID-19 Radiography Database [18], [19], and Chest Imaging data collection [20] sources. This dataset is divided into three subsets: training, development, and test1. While for the second dataset, we used images from the RSNA Pneumonia Detection Challenge [21] and Chest X-rays Radiopaedia [22] including 758 X-ray images. This set, named test2, will be used to make a double evaluation of the model in order to check the model’s ability to generalize.

For both datasets, we consider three classes: Normal, COVID-19, and other pneumonia, given that the last class gathers bacterial and viral pneumonia. Data distribution over training, evaluation, and both test sets is presented in Table I.

<table>
<thead>
<tr>
<th>TABLE I. LABEL DISTRIBUTIONS IN TRAINING, DEVELOPMENT AND TEST SETS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class</td>
</tr>
<tr>
<td>------------</td>
</tr>
<tr>
<td>NORMAL</td>
</tr>
<tr>
<td>COVID-19</td>
</tr>
<tr>
<td>OTHER PNEUMONIA</td>
</tr>
</tbody>
</table>

B. Model Architecture

The experimental dataset used in this study consists of 4,187 X-ray images, which is considered insufficient for training a neural network with a deep architecture. Consequently, the COVID-19 detection system proposed in this research relies on the Inception-V3 model architecture that is pre-trained on the ImageNet dataset [23].

The original Inception-V3 architecture is composed of two main sections: feature extraction and classification. The feature extraction part includes five convolutional layers, each followed by batch normalization, two pooling layers, and 11 inception modules. Each inception module consists of a combination of parallel convolutional and pooling layers. The second part of the architecture comprises fully-connected and softmax layers, responsible for the final classification task.

To customize the Inception-V3 model for our specific task, we made modifications by removing its original classification part. Instead, we constructed a custom classification section to suit the number of classes in our dataset. The customized part consists of an average-pooling layer, followed by two fully connected layers. To prevent overfitting during training, each of these three added layers is accompanied by a dropout layer, randomly deactivating activations with a probability of 0.4.

Finally, as the output for classification, a softmax layer was incorporated, providing the probabilities for each class. The model architecture is visually depicted in Fig. 1.

C. Performance Evaluation

In addition to measuring classification accuracy, we also employed sensitivity in Eq. (1) and specificity in Eq. (2) as metrics for evaluating the COVID-19 detection performance. Sensitivity, also known as True Positive Rate or Recall, represents the model’s ability to correctly identify COVID-19 cases, and can be calculated using Eq. (1):

\[
Sensitivity = \frac{TP}{TP + FN} \quad (1)
\]

While Specificity measures the model’s ability to correctly identify normal (non-COVID-19) cases and can be calculated using Eq. (2):

\[
Specificity = \frac{TN}{TN + FP} \quad (2)
\]

Where:

- TP (True Positives) refers to the number of correctly identified COVID-19 cases;
- FN (False Negatives) refers to the number of actual COVID-19 cases that were incorrectly classified as non-COVID-19 cases by the model;
- TN (True Negatives) refers to the number of correctly identified normal cases;
- FP (False Positives) refers to the number of normal cases that were incorrectly classified as COVID-19 cases by the model.
In the first set of experiments, we investigated the fine-tuning of the Inception-V3 model. Table II shows the customized Inception-V3 model’s performance of the test set with and without adaptation of the Inception-V3 convolution part. In the first line, the Inception-V3 model was pre-trained on ImageNet dataset and only the customized layers were trained on the chest X-ray images as explained in Section III. While in the second line, the pre-trained Inception-V3 model is fine-tuned by gradually unfreezing the top layers of the convolutional part of the model on the chest X-ray images and the customized layers are trained from scratch.

The results, in Table II, show that using Inception-V3 convolution parts without any adaptation on the new task give acceptable results with an accuracy of 95.62%. However, when the model is fine-tuned (adapted) on the new task the results are even better with an accuracy of 98.98% and a sensitivity and specificity of 98.25% and 99.34%, respectively.

| TABLE II. COVID-19 DETECTION RESULTS USING INCEPTION-V3 WITH AND WITHOUT ADAPTATION ON THE TEST1 SUBSET |
|-----------------------------------|-----------|-----------|
| Inception-V3 without adaptation   | 95.62     | 92.11     | 97.23     |
| Inception-V3 with adaptation      | 98.98     | 98.25     | 99.34     |

In the next experiments, we look at the effect of data preprocessing on the model’s generalization while preserving fine-tuning when training our model. We apply the well-known Contrast Limited Adaptive Histogram Equalization (CLAHE) on our X-ray images while analyzing the model’s performance on the second test set (Test2). The results reported in Table III.

| TABLE III. COVID-19 DETECTION RESULTS USING INCEPTION-V3 WITH AND WITHOUT CLAHE PREPROCESSING ON TEST2 SUBSET |
|---------------------------------|-----------|-----------|
| Inception-V3 with adaptation    | 81.62     | 40.91     | 11.50     |
| Inception-V3 with adaptation + CLAHE | 81.27 | 93.18 | 78.75 |

It can be seen by comparing the classification accuracies achieved when evaluating the model on the second test set that without data preprocessing, the model shows a poor generalization, where the classification accuracy drops from 98.98% on the first test set to only 16.62% on the second one. On the other hand, adopting data preprocessing leads to a significant improvement in performance generalization on the second test set, with an accuracy of 81.27%.

This improvement in the model’s generalization could be explained by the consistent quality of X-ray images from each class coming from the same sources. Therefore, the model is very sensitive to image quality, and when re-evaluated in a real-world test set (test2) where it encounters images of varying quality, it may not perform as well as it did during the initial evaluation.

In summary, experiments show that the adaptation of the CNN layers of the pre-trained Inception-V3 model on COVID-19 X-ray images leads to better results as compared to using the model without any adaptation. Furthermore, the results confirmed that image preprocessing is essential in medical image analysis as it enhances the quality of medical images, improves the performance of classification models, and ensures the accuracy and reliability of diagnostic and analytical results.

V. SOFTWARE DESCRIPTION

We embedded the trained Inception-V3 model into a web and Android mobile application for automated COVID-19 screening. The proposed application is called CovidChecker and it is publicly available on GitHub [24].

The app is designed to identify anomalies and estimate the extent of lung infection in real-time, and automatically detect high-risk patients with pneumonia or other COVID-19 related pulmonary symptoms.

The app enables users to complete a questionnaire and submit chest X-ray images for COVID-19 screening. The images and accompanying metadata will be used to augment the training dataset of DL models, thereby improving the accuracy and reliability of COVID-19 tests over time. To protect privacy, the data collected from users will be stored on secure servers for research purposes only.

A. Software Architecture

The software architecture, as shown in Fig. 2, illustrates the implementation of the web and mobile application utilizing Flask, MongoDB, and the DL module in the backend, jQuery Ajax in the web frontend, and Retrofit for the Android mobile application. The architecture is designed as follows:

Web Frontend: The web frontend is developed using jQuery Ajax to communicate with the Flask backend. Through the web user interface, the users can capture images using their computer’s webcam and complete a questionnaire to assist the decision-making.

Backend: The backend is built using Flask, a RESTful API for the Python web server. It receives requests from the frontend and process them. MongoDB is used as a database to store images, classification results, and user data. All data transferred to the application is fully anonymized, including metadata. It has been hypothesized that an individual is not identifiable from their chest radiographs.

DL Module: The best DL module trained in section IV is loaded into the Flask backend and used to process images submitted by users.

Android Mobile Application: The Android mobile application uses Retrofit to communicate with the Flask backend. The mobile user interface allows the user to take images via their smartphone camera and fill out a questionnaire, similar to the web application.

The architecture is designed to be fast, scalable, and user-friendly for both web and mobile users. It can also handle scenarios where the mobile application loses connectivity to the backend.

B. Data Privacy Policy

The information that is collected will be transmitted from the user’s phone, when they are connected to the Internet, to a secure database server. Once the data has been successfully transmitted, it will be deleted from the user’s device. We do not
collect any personally identifiable information such as email addresses or other explicit personal identifiers. Since the data collected is sensitive and contains demographic information, all data transferred to the database server is fully anonymized, including metadata. Furthermore, we established a sharing agreement for the data under the terms and conditions agreement of the mobile app.

C. Illustrative Example

The main functionality of this app is self-screening, which provides an automatic and rapid computer-aided diagnosis of COVID-19 from chest X-ray images. During this stage, the app will collect some basic demographic and medical history data, as well as the X-ray image through a quick questionnaire (see Fig. 3).

Before proceeding with the self-screening, users are required to accept the terms and conditions governing the app’s use and agree to the data privacy policy that outlines how data is collected, stored, and managed. Second, after providing basic demographic information, such as their gender, age, region, and smoking history, the users will be asked whether they have previously tested positive for COVID-19. Then, they will be prompted to enter their medical history and any current symptoms they may be experiencing.

Finally, the user uploads his/her X-ray image and would be able to find the results within a matter of seconds. The uploaded image is sent to the backend server for analysis, where it undergoes preprocessing before being passed to a pre-trained image classification model. The model determines whether the input image shows signs of COVID-19 or not, and sends the results back to the mobile app in real-time. The app user receives the screening results and guidelines based on the X-ray analysis results. The results are presented as the probability of having COVID-19, which is determined by the output of the DL module. If the probability is less than 50%, the results page will appear in green to indicate safety (see Fig. 4), while a probability exceeding 50% will cause the results page to turn red (see Fig. 5).

D. Software Impact

The study on using transfer learning for COVID-19 detection from X-ray images can have a significant impact on various parties involved in healthcare.

For patients, this system could provide a faster and more accessible diagnosis, which could lead to earlier treatment and
better outcomes. Additionally, the use of X-ray images can be less invasive than other diagnostic methods, such as RT-PCR testing, and can be done onsite, reducing the need for patients to travel to a medical facility.

For healthcare providers, the proposed system could serve as an aid in the diagnosis process, reducing the time and effort required for manual analysis of chest radiographs and allowing healthcare professionals to focus on tasks that are more critical. This could also help to improve the accuracy and consistency of COVID-19 diagnoses, as the system can provide a second opinion when a radiologist is not immediately available.

Moreover, the successful implementation of this system in the larger healthcare industry could lead to the wider adoption of computer vision and deep learning techniques in the diagnosis of various diseases. This could significantly improve the efficiency and accuracy of medical diagnoses, ultimately leading to better patient outcomes and reduced costs for the healthcare system.

Overall, the proposed system has the potential to improve the efficiency and accuracy of COVID-19 diagnoses, providing benefits to patients, healthcare providers, and the broader healthcare industry.
VI. Conclusion

In conclusion, this research paper presents an innovative solution in the form of a mobile application that employs machine learning, computer vision, and deep learning to analyze X-ray images for the rapid detection of COVID-19 symptoms. The study focuses on the use of convolutional neural networks (CNNs) and transfer learning to diagnose COVID-19 from chest X-ray images, showcasing their effectiveness in image classification. The results demonstrate that adapting the Inception-V3 model with fine-tuning leads to superior accuracy, sensitivity, and specificity in COVID-19 detection.

The paper addresses various challenges and limitations encountered in the development of automatic COVID-19 detection solutions, such as data availability, data imbalance, generalization issues, and the need for clinical validation. It also highlights the importance of data preprocessing, such as CLAHE histogram equalization, in enhancing the quality of medical images and improving classification model performance.

The research introduces “CovidChecker”, a mobile application that enables users to self-screen for COVID-19, submit chest X-ray images, and provide valuable data for improving the accuracy of COVID-19 tests over time. The application is designed with a strong focus on data privacy, ensuring that user data is anonymized and securely stored.

ACKNOWLEDGMENT

This work was supported by the CNRST Morocco under contract no UCD-CNRST-5.

REFERENCES


[25] www.ijacsa.thesai.org 1260 | P a g e