Published by :
http://www.ijert.org

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 12 Issue 11, November-2023

Automatic COVID-19 Detection Systems using Deep Learning Techniques from Chest X-ray Images

Mohammed Khaled Bin Jaah
Faculty of Electrical and Computer Engineering department,
King Abdulaziz University, Saudi Arabia, Jeddah

Abstract— **Deep learning has made significant advances in the previous several years, leading to remarkable advancements in a variety of applications, including images categorization. Deep learning has emerged as a crucial methodology in the domain of medical imaging that has significantly improved the categorization of medical pictures. Convolutional neural networks (CNNs) have shown very high effectiveness in the detection of several illnesses, including different dental conditions, Parkinson's disease, Alzheimer's disease, malaria, and coronary artery disease. Similar to earlier instances, CNN shows a lot of potential for diagnosing COVID-19 patients via the use of medical imaging methods like computed tomography and chest X-rays. The World Health Organization (WHO) has officially designated the coronavirus illness COVID-19 as a worldwide pandemic. Identification of persons who have tested positive for COVID-19 is crucial for stopping the spread of this contagious illness. In this study, models like ResNet50, Alex Net, Google Net, Mobile Net, and Modified ResNet50s are developed for using chest X-ray pictures to detect COVID-19 patients. A dataset including 3,000 chest X-ray (CXR) images is used, encompassing examples of both COVID-19 positivity and negativity. All models are trained and confirmed using COVID-19 chest X-ray images and normal chest X-ray images. The Modified ResNet50 model with Channel Shuffle has an F1-score of 100%, a Receiver Operating Characteristic (ROC) curve area of 100%, and precision and accuracy of 100% and 100%, respectively. This research also compares the effects of increasing dataset size and convolutional layer alterations on classification performance.**

Keywords- Deep learning · Convolutional Neural Networks (CNNs) · COVID-19 · Normal · X-ray images · Receiver Operating Characteristic (ROC).

I. INTRODUCTION

The infectious disease known as COVID-19, which emerged in 2019, is caused by the novel virus SARS-CoV-2 and is characterized by its high spread. This virus was initially observed in Wuhan, Hubei Province, China, in December 2019 and has subsequently disseminated rapidly to nearly all nations and regions worldwide. As a consequence, an unparalleled global pandemic has ensued, marking a significant event in human history [1] [2]. The declaration of the COVID-19 epidemic in China as a Public Health Emergency of International Concern (PHEIC) was made by the Director-General of the World Health Organization (WHO) on January 30, 2020. This decision was based on the recognition of the outbreak's capacity to have a substantial effect on countries with insufficient healthcare infrastructure. In March 2020, the World Health Organization (WHO) officially classified the outbreak as a pandemic. [3].

As of March 2021, global reports indicate that the number of COVID-19 cases has exceeded 119 million, with a verified death toll over 2.6 million. Since the onset of the global pandemic, around 94 million individuals have seen recovery from the infection, although a substantial number of almost 20 million cases remain active throughout various regions worldwide. The ongoing epidemic continues to exacerbate and exhibits no indications of abating or decelerating. The rapid spread of the pandemic is being seen in many countries, including the United States, India, and Brazil [4].

Since the emergence of technology in the field of medical science, there has been a notable advancement in medical gadgets and diagnostic techniques. The current generation of Internet of Things (IoT) devices has significantly simplified intricate processes and facilitated real-time monitoring. The field of technology is constantly advancing [5]. In contemporary times, the use of deep learning techniques has enabled the generation of comprehensive and varied feature sets beyond the capabilities of even domain experts [6].

Over the last several decades, machine learning (ML) and deep learning (DL) have gained significant attention and have proven to be very valuable in addressing complex medical scenarios. In order to effectively discern between pneumonia-infected lungs and healthy lungs, machine learning algorithms rely on well-structured and ordered data. To predict the result, a collection of features is necessary. In contrast, within the context of deep learning (DL), the classifier is capable of discerning between normal and pneumonia images by using network properties that are created automatically [7]. There is no need to provide sustenance to a domestically produced filtration system. Deep learning (DL) has emerged as a powerful technology that obviates the need for manually creating features, owing to its capability to create features

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 12 Issue 11, November-2023

autonomously. Numerous breakthroughs have been achieved in deep learning algorithms over the years, resulting in their extensive use for the detection of abnormalities in medical images. There are deep learning algorithms that are capable of using previously obtained knowledge in order to address a novel challenge [8].

Computer-Aided Detection (CAD) provides a considerable degree of precision in the identification and diagnosis of COVID-19, while simultaneously reducing the workload and time required for these tasks. The use of imaging analysis and processing, in combination with the suitable picture modality and diagnostic techniques. This thesis makes a dual contribution of significant importance. Firstly, it presents the development of an automated detection system that exhibits high precision and efficiency in identifying COVID-19 from CXR images. The CAD system reported in this study, however, has been trained and optimized using data from digital X-ray datasets via the implementation of five-fold validations [9].

I. LITERATURE REVIEWE

In [10], the authors proposed and evaluate various architectures for the task of COVID-19 detection using X-ray images. Their proposed architectures include the pre-trained MobileNetv2 and ResNet50 models. They tested these computer designs using a big collection of X-ray images. Then, they compared how well their designs worked with other methods that people already use to find COVID-19 in X-rays. Their proposed models are compared to existing COVID-19 detection algorithms in terms of accuracy. The proposed models achieve an accuracy of 96.71% and an F1-score of 91.89%. These results demonstrate that the proposed models outperform the most advanced algorithms currently available.

In [11], they conducted a thorough assessment of ten different CNN architectures, considering both their performance with and without enhancement techniques. These models encompass InceptionV3, InceptionResNetV2, MobileNet, MobileNetV2, Vgg19, NASNetMobile, ResNet101, DenseNet121, DenseNet169, and DenseNet201. Their proposed methodology, when augmented with TVF + Gamma, demonstrated notably superior results in terms of classification accuracy and sensitivity. In the 4-way classification scenario, MobileNet with TVF + Gamma stood out, achieving an impressive accuracy of 93.25%. This represented a substantial improvement of 1.91% in accuracy score. Moreover, it demonstrated a remarkable COVID-19 sensitivity of 98.72% and an F1-score of 92.14%. This indicates that their proposed approach substantially enhanced the model's ability to accurately identify COVID-19 cases.

In [12], they used an examination of image features extracted from MobileNet to ascertain their validity and medical significance. The process demonstrated the ability to identify abnormal X-rays with an accuracy of 95.45 ± 1.54% and effectively differentiate COVID-19 cases with an accuracy of 89.88 ± 3.66%. The visual representations generated by the Grad-CAM algorithm offered compelling proof that the methodology pinpointed meaningful regions within the images.

In [13], they employed various deep learning techniques, including deep feature extraction, fine-tuning of pre-trained convolutional neural networks (CNNs), and end-to-end training of a customized CNN model, to differentiate between COVID-19 and healthy chest X-ray images. Pre-trained deep CNN models such as ResNet18, ResNet50, ResNet101, VGG16, and VGG19 were utilized for deep feature extraction. The experiment utilized a dataset containing 180 COVID-19 and 200 healthy chest X-ray images. The fine-tuned ResNet50 model demonstrated an impressive accuracy of 92.6%.

In [14], they introduced a new AI-powered deep learning model for automatically detecting COVID-19 in chest X-ray images. They employed advanced deep learning architectures like ResNet50, VGG19, Xception, and DarkNet19. The model was evaluated accessible datasets and demonstrated an accuracy of 96.10%. Their findings suggested that this approach held significant promise in healthcare, offering a quicker, cost-effective, and highly accurate method for COVID-19 detection.

In [4], they classified COVID-19, non-COVID-19 viral pneumonia, bacterial pneumonia, and normal CXR scans gathered from several public sources using a deep learning strategy based on a pre-trained AlexNet model. Their model was prepared to implement two-way classification (i.e., Coronavirus versus ordinary, bacterial pneumonia versus typical, non-Coronavirus viral pneumonia versus ordinary, and Coronavirus versus bacterial pneumonia), three-way classification (i.e., Coronavirus versus bacterial pneumonia versus typical), and four-way classification (i.e., Coronavirus versus bacterial pneumonia versus non-Coronavirus viral pneumonia versus typical). Their model obtained 94.43% accuracy, 98.19% sensitivity, and 95.78% specificity for non-COVID-19 viral pneumonia and normal (healthy) CXR images. Their model obtained 91.43% accuracy, 91.94% sensitivity, and 100% specificity for bacterial pneumonia and ordinary CXR images. Their model obtained 99.62% accuracy, 90.63% sensitivity, and 99.89% specificity for classification CXR images of COVID-19 pneumonia and non-COVID-19 viral pneumonia. Their model obtained 94.00% accuracy, 91.30% sensitiv- ity, and 84.78% for the three-way classification. Lastly, for the four-way classification, their model obtained 93.42%, sensitivity of 89.18%, and specificity of 98.92%.

In [15], they used eight different deep learning techniques to identify people with COVID-19 in a set of 400 chest X-ray images. The techniques were VGG16,

InceptionResNetV2, ResNet50, DenseNet201, VGG19, MobileNetV2, NasNetMobile, and ResNet15V2. Among these, Net-Mobile outperformed all other models on the chest X-ray datasets, achieving an impressive accuracy of 93.94%.

In [16], in this paper, they aimed to diagnose COVID-19 disease from X-ray images using deep learning

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 12 Issue 11, November-2023

architectures. Additionally, a 96.30% accuracy rate was achieved with the hybrid architecture they had improved. While developing the hybrid model, the last 5 layers of the Resnet 50 architecture were ejected. 10 layers were added in place of the 5 layers that were removed. The count of layers, which was 177 in the Resnet50 architecture, was increased to 182 in the hybrid model. Thanks to these layer changes made in Resnet50, the accuracy rate was increased more. Classification was performed with AlexNet, Resnet50, GoogLeNet, VGG16 and developed hybrid architectures using COVID-19 Chest X-Ray dataset and Chest X-Ray images (Pneumonia) datasets. In [17], they devised a deep learning system tailored to extract features and identify COVID-19 in chest X-ray images. They fine-tuned three potent networks, namely ResNet50, InceptionV3, and VGG16, using an enriched dataset constructed by pooling COVID-19 and normal chest X-ray images from various public databases. To augment the dataset, they employed techniques like random rotations within -10 to 10 degrees, the addition of random noise, and horizontal flips to artificially generate a substantial number of chest X-ray images. The experimental outcomes were exceptionally encouraging: the proposed models achieved remarkable accuracy, specifically 97.20% for Resnet50, 98.10% for InceptionV3, and 98.30% for VGG16, in the classification of chest X-ray images as Normal or COVID-19. These results underscore the effectiveness of transfer learning, offering robust performance and easily implementable methods for COVID-19 detection.

TABLE 1 AN OVERVIEW OF RESULTS FROM EARLIER STUDIES, RESNET50 (RN50), AND MOBILENET (MON), AND NOT CALCULATED (NC).

References No.	***[10]***	***[11]***	***[12]***	***[13]***	***[4]***
Methods Used	RN50 and MON	MON	MON	RN50	AlexNet
Accuracy	91.54 % 95.73 %	93.25%	90%	94.7%	94.43%
Sensitivity	NC	98.72%	NC	91%	98.19%
Specificity	NC	NC	NC	98%	95.78%
Extra Evaluation Metrics	F1-score = 91% F1-score = 96%	F1-score= 92.3%	NC	F1-score= 94.79%	NC

References No.	***[14]***	***[15]***	***[16]***	***[17]***
Methods Used	RN50	RN50 and MON	RN50 and GoogLeNet AlexNet	RN50
Accuracy	90.68%	RN50= 64% MON= 97%	RN50= 92.59 GoogLeNet =90.74% AlexNet= 88.89%	97.20%
Sensitivity	NC	NC	97%	98.25%
Specificity	NC	NC	98%	97%
Extra Evaluation Metrics	NC	RN50: Precision= 79% Recall= 64% F1-score= 58% MON: Precision= 97% Recall= 97% F1-score= 97%	NC	Precision= 97% Recall= 96% F1-score= 97%

II. METHODOLOGY

In this study, MATLAB was used to evaluate data from 3000 x-ray images. Three stages were used throughout the study to distinguish between normal and abnormal COVID-19. These stages are shown in figure 1.

3.1 DATABASE SOURCE:

The databases of the University of Dhaka and Qatar University will be utilized in this study [18]. A total of 13,808 chest X-ray (CXR) images are included in the database (including both COVID-19 positive and negative cases). 10,192 normal and 3,616 COVID-19 cases make up the datasets.

3.1 PREPROCESSING:

Preprocessing is essentially the stage for locating and minimizing images artifacts. This step is required in CXR images because many CXR images consist of noise and undesired artifacts such as patient clothing and wire which must be eliminated in order to appropriately diagnose COVID-19. Initially, the RGB images were transformed into grayscale using the MATLAB tool, and then they were resized to prepare them for input into the system.

The region of interest (ROI) for training and testing was extracted to eliminate superfluous text and machine annotations around the images. Meaningful information was obtained by defining the ROI on the chest X-ray images, mainly covering the lung region [19].

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 12 Issue 11, November-2023

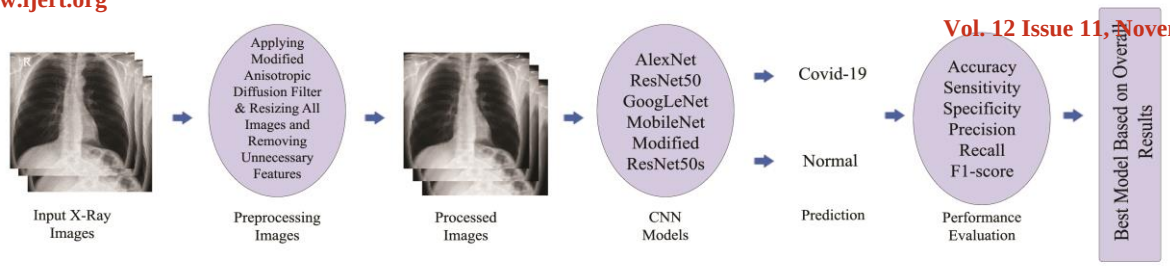

Fig. 1 Visual representation of the experiment's workflow

Initially, a rectangle is employed to delineate the region of interest (ROI), and a mask is generated based on this rectangle. Subsequently, through logical indexing, the area outside the ROI is zeroed out, and the isolated portion is presented. Filtering methods maintain valuable information while effectively addressing noise present in an image. When it comes to extracting significant features from noisy images, information-preserving filtering techniques are particularly well-suited [20].

In this study, the Modified Anisotropic Diffusion Filtering (MADF) technique was employed to preserve detailed information while reducing noise and distortion in the images. This filtering technique is noted for its superior performance compared to other filtering methods, attributed to its capacity for eliminating multiplicative speckle noise in plane regions.

The method being introduced relies on correlation and kurtosis measurements of noise to retain valuable edge information. In Equation (1), Io represents a noisy image that combines speckle noise (n) with the original image I [21, 22]. The noise component is described in Equation (1), where the noise intensity, denoted as G, is computed based on image characteristics using MATLAB. The average noise intensity, denoted as μ and determined by Equation (2), along with the kurtosis value, calculated using Equation (3), play essential roles. The objective is to minimize the correlation between the image class and noise class, serving as the stopping criterion for the iteration. This speckle reduction process persists until the noise component of the image approaches a Gaussian value, at which point the kurtosis value should ideally reach zero. The iterative process concludes when the kurtosis value drops below 0.001 (as specified in Equation (4)),

Signifying effective speckle reduction while preserving edges. The iteration halts once the correlation between the image class and noise class reaches its minimum. Equation (5) computes the correlation of image intensities (ρI), and Equation (6) calculates the correlation of noise intensities (ρG). The filtering method's optimal outcome is achieved when ρI and ρG exhibit the least deviation from each other.

$$I_0 = I \cdot n \quad (1)$$

$$n = \frac{I - G}{\sqrt{G}} \quad (2)$$

$$\mu = \frac{\sum_{i=1}^{N} G_i}{N} \quad (3)$$

$$k = \frac{\frac{1}{N}\sum_{i=0}^{N}(G-\mu)^4}{[\frac{1}{N}\sum_{i=0}^{N}(G-\mu)^2]^2} - 3 \quad (4)$$

$$\rho_I = \frac{\sum_{i=0}^{M-1}\sum_{j=0}^{N-1} i \cdot j \cdot p_I(i.j) - \mu I_x \mu I_y}{\frac{\sum_{i=1}^{N}(I_{ix} - \mu_{Ix})(I_{iy} - \mu_{Iy})}{N}} \quad (5)$$

$$\rho_G = \frac{\sum_{i=0}^{M-1}\sum_{j=0}^{N-1} i \cdot j \cdot p_G(i.j) - \mu G_x \mu G_y}{\frac{\sum_{i=1}^{N}(G_{ix} - \mu_{Ix})(G_{iy} - \mu_{Iy})}{N}} \quad (6)$$

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 12 Issue 11, November-2023

Figure 2 illustrates original images, noisy images, and the results after applying modified anisotropic diffusion filtering. It is evident that the proposed MADF technique effectively preserves edges.

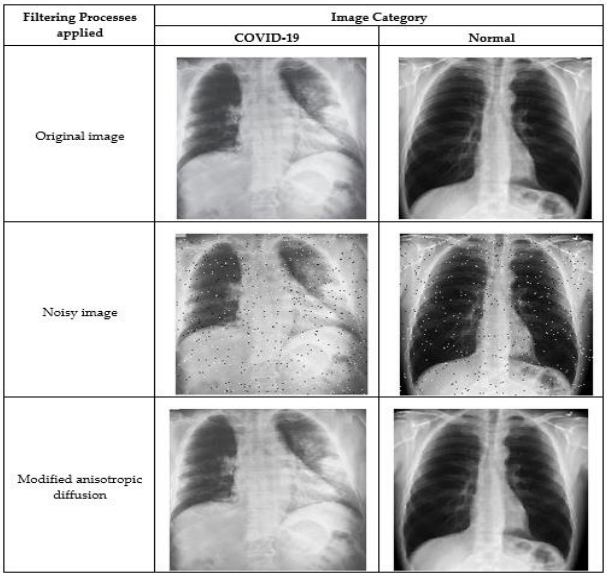

Fig. 2 A visual representation of the results of applying modified anisotropic diffusion on the images.

3.1 Predicting COVID-19 Using Pre-Trained CNN Models:

Convolutional neural networks (CNNs) have demonstrated exceptional performance in various medical image processing applications. However, training CNN models from scratch for COVID-19 case prediction poses a challenge due to the scarcity of X-ray samples. This is where transfer learning (TL) comes into play. TL leverages the knowledge acquired from a deep learning (DL) model trained on a massive dataset to tackle a related task with a relatively smaller dataset. This approach eliminates the need for a large dataset and lengthy training times, which are inherent in training DL models from scratch [23]. This study employs four pre-trained models, namely ResNet50, AlexNet, GoogLeNet, and MobileNet, to classify COVID-19 cases from normal ones. These models have proven highly effective in various computer vision and medical image analysis tasks, making them suitable for differentiating COVID-19 infections from normal cases. Notably, these models were initially trained on a large-scale labeled dataset called ImageNet [11] and subsequently fine-tuned using chest X-ray images. The final layer of these models was removed and replaced with a new Fully Connected (FC) layer having an output size of two, corresponding to the two classes (normal and COVID-19). In these refined models, only the final Fully Connected (FC) layer undergoes training, while the remaining layers retain their pre-trained weights. The hyperparameters, which are crucial for optimizing these deep learning models, were held constant to ensure a fair comparison.

The architectural summary of the pre-trained CNN models is presented in Table 2.

Table 2 Descriptions of the architecture of the pre-trained CNN models utilized in this study.

Model	*Layers*	*Parameters (in million)*	*Input layer size*	*Output layer size*
AlexNet	8	60	(224,224,3)	(2,1)
GoogleNet	22	5	(224,224,3)	(2,1)
MobileNet	53	3.4	(224,224,3)	(2,1)
ResNet50	50	25.6	(224,224,3)	(2,1)

3.1 Performance Evaluation:

Following model training experiments, this study has yielded six models (ResNet50, AlexNet, GoogLeNet ,and MobileNet) that exhibit strong COVID-19 detection capabilities when applied to x-ray images. Subsequently, this paper introduces a set of evaluation metrics tailored to medical image classification. These metrics encompass accuracy, sensitivity, specificity, precision, recall, F1-score, a confusion matrix, and a graphical representation illustrating how accuracy evolves with the number of training epochs. Each classification metric was computed, enabling us to provide a detailed evaluation of the detection and classification prowess of each model. These calculations are outlined as follows [24]:

- **True positives (TP):** The image actually has COVID-19, and the model correctly predicts that it does.
- **False positives (FP):** The image actually does not have COVID-19, but the model incorrectly predicts that it does.
- **True negatives (TN):** The image actually does not have COVID-19, and the model correctly predicts that it does not.
- **False negatives (FN):** The image actually has COVID-19, but the model incorrectly predicts that it does not.
- **Receiver Operating Characteristic Curve:** The graph that plots the true positive rate (TPR) against the false positive rate (FPR). The area under the curve (AUC) is a measure of the overall performance of the classifier [25].
- **Confusion Matrix:** The table that summarizes the performance of a classifier. It shows the number of instances that were correctly classified and incorrectly classified for each class.
- **Accuracy (ACC):** The percentage of instances that are correctly classified. In the context of mass lesions, accuracy measures how many patients with mass lesions are correctly predicted to have mass lesions. In the context of diagnosis, accuracy measures how many patients with malignant pathology are correctly classified as having malignant pathology, or how many patients with benign pathology are correctly classified as having benign pathology.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 12 Issue 11, November-2023

$$ACC = \frac{(TP + TN)}{(TP + TN + FP + FN)} \quad (7)$$

- **Sensitivity (SEN) or Recall (REC):** The proportion of positive instances that are correctly classified. In other words, it shows how many patients with malignant pathology are correctly classified as having malignant pathology, out of all patients with malignant pathology.

$$SEN\ or\ REC = \frac{(TP)}{(TP + FN)} \quad (8)$$

- **Specificity**: the proportion of negative instances that are correctly classified. In other words, it shows how many patients without malignant pathology are correctly classified as not having malignant pathology, out of all patients without malignant pathology.

$$SPE = \frac{TN}{(TN + FN)} \quad (9)$$

- **Precision:** The proportion of positive predictions that are actually positive. In other words, it shows how many patients who are predicted to have malignant pathology actually have malignant pathology.

$$PRE = \frac{TP}{(TP + FP)} \quad (10)$$

- **F1-score:** The harmonic mean of precision and recall. It is a measure of how well a model performs at both detecting positive instances and avoiding false positives.

$$F1 - SCORE = \frac{2 \times PERCISION \times RECALL}{(PERCISION + RECALL)} \quad (11)$$

Table 3 All performances of 4 different models on each fold for Normal / COVID-19 cases, Accuracy (ACC), Sensitivity (SEN), Recall (REC), Specificity (SPE), Precision (PRE), F1-Score (F1), and Average (AVE).

Models/Fold		*ACC %*	*SEN %*	*PRE %*	*SPE %*	*REC %*	*F1 %*
AlexNet	1	99.3	98.7	98.6	98.3	98.7	99.3
	2	96.5	98.3	93.2	98.3	98.3	99.1
	3	99.3	98.7	98.6	98.7	98.7	98.7
	4	98.5	99.7	100	99.7	99.7	99.8
	5	100	100	100	100	100	98.1
	AVE	98.7	99.1	98.0	99.0	99.1	99.0
ResNet50	1	99.9	100	100	100	100	100
	2	99.8	100	100	100	100	99.9
	3	99.9	100	100	100	100	100
	4	99.9	99.5	100	100	100	100
	5	100	100	100	100	100	100
	AVE	99.9	99.9	100	100	99.9	99.9
MobileNet	1	97.5	97.5	99.3	97.5	98.3	99.5
	2	99.6	99.6	99.1	98.6	98.4	98.6
	3	98.7	98.7	98.7	97.7	98.7	98.7
	4	98.3	98.3	99.8	98.3	97.5	98.3
	5	97.9	98.4	98.1	98.9	99.6	98.4
	AVE	98.4	98.5	99.0	98.2	98.5	98.7
GogLeNet	1	99.9	98.3	99.9	99.0	98.3	99.8
	2	99.8	99.6	99.9	99.9	99.6	99.7
	3	99.9	98.4	99.5	99.8	98.4	99.9
	4	99.9	97.5	99.8	99.9	97.5	99.9
	5	99.5	98.7	99.9	99.9	98.7	99.7
	AVE	99.8	98.5	99.8	99.7	98.5	99.8

II. RESULTS

4.1 Configuration of Experimental Parameters and Results:

MATLAB software was used to train the suggested deep transfer learning models. Pre-trained CNN models included ResNet50, AlexNet, GoogLeNet, and MobileNet. The dataset was divided into two separate datasets at random, with 20% and 80% of each dataset utilized for testing and training, respectively. K-fold was selected as the cross validation technique, and results were obtained for 5 distinct k values (k=1-5).

In this study, two classes were used for classifications (2000 Normal and 1000 COVID-19 cases). In order to get a robust result in this research, the 5-fold cross validation approach was used to the four pre-trained models ResNet50, Alex Net, Google Net and Mobile Net. Table 2 provides a detailed breakdown of performance metrics for every fold value of each model.

As demonstrated in the comprehensive evaluation presented in the table 2, it became evident that ResNet50 consistently delivered outstanding performance. The results displayed in the table above clearly indicate that, among the various models examined, ResNet50 consistently achieved the highest level of accuracy. In Figure 3, the presented confusion matrix serves as a visual aid to assess the overall performance of the models. This includes the computation of metrics such as accuracy, sensitivity, specificity, precision, recall, and F1-score, which are determined using the values from Equations (7)–(11). The dataset comprises 3000 data samples, with 1000 being COVID-19 samples and 2000 being normal samples. COVID-19 and normal cases are represented by the labels ‘C’ and ‘N’, respectively. According to the matrix, the modified ResNet50 successfully distinguishes 1998 normal x-ray images and 1000 COVID-19 x-ray images. In total, the model accurately identifies 2998 images (comprising both normal and COVID-19 images), resulting in an impressive accuracy rate of 99.9%. The suggested model can handle much deeper networks compared to other models that used in this study. This is because it uses special connections that help with training very deep networks. It also uses shortcuts that speed up training by making it easier for the network to learn from its mistakes. Additionally, the classification accuracy of the proposed framework is competitive with recent research (see

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 12 Issue 11, November-2023

Table 4). It's important to note that our work deals with a large collection of X-ray images, which introduces additional challenges for the classification model.

Table 4 Evaluating our best outcomes in light of other studies, Accuracy (ACC), Sensitivity (SEN), Recall (REC), Specificity (SPE), Precision (PRE), and F1-Score (F1).

Reference s No.	*Model used*	*ACC %*	*SEN %*	*PRE %*	*SPE %*	*REC %*	*F1 %*
[10]	ResNet50	91.5	NC	NC	NC	NC	91
[13]	ResNet50	94.7	91	NC	98	NC	94.7
[14]	ResNet50	90.6	NC	NC	NC	NC	NC
[15]	ResNet50	64.0	NC	79.0	NC	64.0	58.0
[16]	ResNet50	92.5	97.0	NC	98.0	NC	NC
[17]	ResNet50	97.2	98.2	97.0	97.0	96.0	97.0
Our study	ResNet50	99.9	99.9	100	100	99.9	99.9

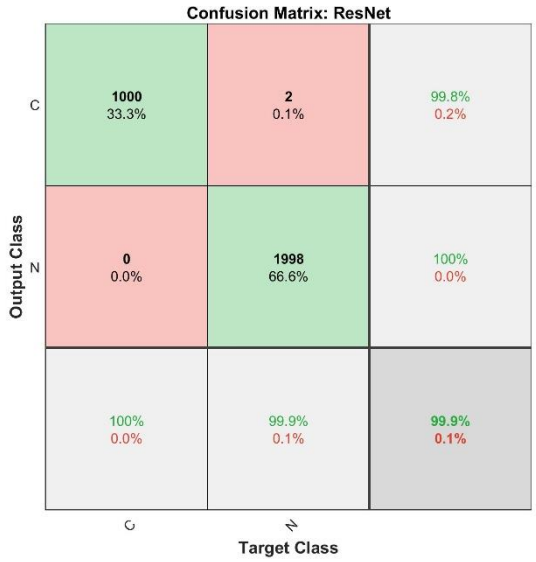

Fig. 3 Confusion matrix of ResNet50.

4.2 Ablation Experiments:

In the exploration of deep learning model enhancements, ablation experiments assume a pivotal role as they involve the systematic removal or modification of specific elements to evaluate their impact on overall performance [26]. Based on the experimental results, ResNet50 model demonstrated the most favorable overall performance. The architecture under consideration, ResNet50, is well-known for its depth and performance. ResNet50 includes input and output layers, convolutional layers for feature extraction, and pooling layers to reduce spatial dimensions. Residual blocks, each with multiple convolutional layers, contribute to deep representation learning. Fully connected layers transform features for classification. The network's unique design has proven effective in diverse computer vision tasks, highlighting its significance in deep learning applications [27]. However, there is a desire to investigate potential modifications that could enhance its capabilities. Three key modifications were introduced into our proposed model: the addition of a Channel Shuffle layer, the introduction of a Softmax layer, and the removal of the Batch Normalization layer.

- Channel Shuffle Layer: The Channel Shuffle layer is introduced to promote cross-channel information flow, fostering the improvement of the model's ability to capture diverse features. By rearranging feature maps across channels, the layer encourages richer representations, allowing the model to better extract intricate patterns within the data.

- Softmax Layer: Incorporation of a Softmax layer into the model is aimed at the enhancement of its classification capabilities. The Softmax function is utilized to normalize raw output scores into probabilities, facilitating a more robust and interpretable prediction mechanism. This addition seeks to improve the model's ability to provide confident and well-calibrated class probabilities.

- Batch Normalization Removal: The decision to omit the Batch Normalization layer, a common technique for stabilizing and accelerating training, is made to explore the impact of its absence. Removal of the Batch Normalization layer aims to evaluate whether the model can maintain or even enhance its performance in terms of convergence speed and generalization.

Through this ablation experiment, the objective is to assess the individual contributions of the Channel Shuffle layer, Softmax layer, and the absence of Batch Normalization to the ResNet50 model. Results obtained from these modifications will offer insights into the importance of these components, potentially guiding future architectural choices and optimizations in deep learning models. The following table illustrates the comparison between the results of the original ResNet50 model and those of the modified model.

Model used	*ACC %*	*SEN %*	*PRE %*	*SPE %*	*REC %*	*F1 %*
ResNet50+ Channel Shuffle	100	100	100	100	100	100
ResNet50+ Softmax	99.99	99.99	100	100	99.9	99.99
ResNet50- Batch Normalization	99.83	99.75	100	100	99.7	99.87
Original ResNet50	99.99	99.99	100	100	99.9	99.99

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 12 Issue 11, November-2023

Table 5 Comparison between the results of the original ResNet50 model and modified ResNet50 model.

As shown in the table above, the ResNet50 model, when enhanced with the Channel Shuffle layer, emerged as a significant improvement over the original architecture, showing a perfect score of 100% across all evaluated metrics. This emphasized the effectiveness of promoting cross-channel information flow through the strategic inclusion of the Channel Shuffle layer, thus improving the model's ability to capture diverse features and patterns within the dataset.

In contrast, the addition of the Softmax layer into ResNet50 seemed to have negligible effects on the overall performance, with results closely mirroring those of the original ResNet50 model. This observation suggests that while Softmax normalization contributed to refining class probabilities, its impact on the specific metrics under consideration might have been limited in this context.

Remarkably, removing the Batch Normalization layer added an intriguing aspect to the experiment. While Batch Normalization is typically used to stabilize training and enhance convergence, its absence in this modified model slightly decreased performance. This subtle result implies that Batch Normalization, despite its computational overhead, plays a positive role in the model's capacity to generalize and maintain robust training dynamics.

The ablation experiment offered valuable insights into the specific impacts of the Channel Shuffle layer, Softmax layer, and the absence of Batch Normalization on the ResNet50 model. These discoveries not only improve our understanding of these architectural components but also act as a guide for future optimizations and architectural decisions in the field of deep learning.

III. DISSCUSSION

In this work, a completely automated diagnostic technique for separating COVID-19 instances from healthy individuals and pneumonia cases was suggested. The model was evaluated using a sizable, unbalanced test dataset, and deeper CNN models were added as a benchmark study. As shown in table 5, higher accuracy was obtained than the one achieved in [10]. Our model showed an accuracy of 100% in ResNet50+Channel Shuffle model (compared to their study which achieved 91.5% accuracy). Additionally, an 9% improvement in F1-score were seen.

In [15], they used a smaller dataset of 400 x-ray images for model evaluation, compared to our dataset of 3000 x-ray images. However, this smaller dataset may not be representative of the actual clinical setting, where a small number of cases in the testing pool can lead to a large imbalance in the classes. In [17] reported an overall accuracy of 97.2% using the ResNet50 model, which is 2.8% less than our proposed model. Their precision rate is 97%, which is 3% lower than ours. Precision is an important metric to consider in classification, as it measures the proportion of positive instances that are correctly classified.

Our method showed a significant improvement in overall accuracy over previous studies. Additionally, our dataset was larger than those used in other studies. We ran different models on this dataset to differentiate between normal and pneumonia cases, all of which came from the same source. Our method is highly robust to variations in image preprocessing methods, such as removing noise, determining regions of interest, and adjusting the contrast of images. This robustness was achieved through the use of a deep learning model that was trained on a large dataset of images collected using a variety of methods. As a result, our model was able to achieve a classification accuracy of 100%, which is significantly higher than the accuracy of other models that have been reported in the literature.

Despite the good outcomes achieved, deep learning models come with a set of limitations. One significant challenge arises from their sensitivity to image artifacts and noise. These limitations in the input data can significantly impact the model's performance, leading to wrong classifications or reduced accuracy. Additionally, the computational demands of deep learning models are noteworthy. They often require high-quality central processing units (CPUs) and can be computationally intensive, particularly when processing large datasets or complex architectures. This can lead to extended processing times, potentially impeding real-time applications or requiring significant computing resources. Another critical limitation related to the understanding of these models. Understanding why a particular classification decision was made can be a complex and difficult task. Deep learning models operate by learning complex relationships within the data, often in high-dimensional spaces, making it challenging for humans to comprehend the exact features or patterns that led to a specific outcome. This lack of understanding can be a big problem in important situations where it's really important to know why a decision was made. Therefore, while deep learning models have demonstrated exceptional capabilities, it's very essential to recognize and find solutions for these challenges so that we can use these models in a good and responsible way in different areas.

III. CONCLUSION

Deep learning techniques have demonstrated remarkable achievements by drawing insights from vast collections of labeled images. This unique capability enables them to distinguish complex relationships between the various elements within the images and their respective labels. This means they can recognize difficult patterns and associations that might be challenging for conventional machine learning approaches. In this study, the ResNet50 pre-trained model demonstrated remarkable accuracy, reaching an impressive 100%, surpassing the performance of the other three models. This highlights its potential as a valuable aid for radiologists in clinical settings, enhancing their ability to provide precise diagnoses.

In future research, it is considered important to extend the model's capabilities to process CT scan images for COVID-19 detection. This expansion would allow a more comprehensive evaluation of the model's effectiveness

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 12 Issue 11, November-2023

across various imaging modalities. Furthermore, the potential for adapting the model to identify a range of diseases beyond COVID-19, utilizing radiography images, is being recognized. This adaptation could significantly enhance the model's clinical utility, offering radiologists a versatile tool for diagnosing and evaluating various medical conditions. By training the model on a diverse dataset encompassing different pathologies, it can be refined to recognize specific patterns indicative of a wide array of diseases.

RFERENCES

[1] S. Bakheet and A. Al-Hamadi, “Automatic detection of COVID-19 using pruned GLCM-Based texture features and LDCRF classification,” Comput. Biol. Med., vol. 137, no. June, p. 104781, 2021.

[2] P. Kumar, P. Banjarey, R. Malik, A. N. Tikle, and R. P. Verma, “Population structure and diversity assessment of barley (Hordeum vulgare L.) introduction from ICARDA,” Journal of Genetics, vol. 99, no. 1, 2020.

[3] I. Iqbal, J. Latief, and M. Mudasir, "CoroNet: A deep neural network for detection and diagnosis of COVID-19 from chest x-ray images," Comput. Methods Programs Biomed., vol. 196, p. 105581, Jan. 2020.

[4] W. Zouch, D. Sagga, A. Echtioui, R. Khemakhem, M. Ghorbel, C. Mhiri, and A. B. Hamida, "Detection of COVID-19 from CT and Chest X-ray Images Using Deep Learning Models," Ann. Biomed. Eng., vol. 50, no. 7, pp. 825-835, Jul. 2022.

[5] G. A. Palmer, G. Tomkin, H. E. Martín-Alcalá, G. Mendizabal-Ruiz, J. Cohen, "The Internet of Things in assisted reproduction," Reprod. Biomed. Online, vol. 47, no. 5, pp. 103338, Nov. 2023.

[6] P. Uppamma and S. Bhattacharya, "Deep Learning and Medical Image Processing Techniques for Diabetic Retinopathy: A Survey of Applications, Challenges, and Future Trends," J. Healthcare Eng., vol. 2023, 2728719, Feb. 2, 2023.

[7] Chris Huntingford, Elizabeth S Jeffers, Michael B Bonsall, Hannah M Christensen, Thomas Lees and Hui Yang, “Machine learning and artificial intelligence to aid climate change research and preparedness”, Environmental Research Letters, 14(12), Nov 2019.

[8] X. Liu, W. Wu, J. Chun-Wei Lin, S. Liu, "A Deep Learning Model for Diagnosing COVID-19 and Pneumonia through X-ray," Curr. Med. Imaging, vol. 19, no. 4, pp. 333-346, 2023.

[9] F. Azour and A. Boukerche, “Design Guidelines for Mammogram-Based Computer-Aided Systems Using Deep Learning Techniques,” IEEE Access, vol. 10, pp. 21701–21726, 2022.

[10] M. Emin Sahin, “Deep learning-based approach for detecting COVID-19 in chest X-rays,” Biomed. Signal Process. Control, vol. 78, no. January, p. 103977, 2022.

[11] A. Sharma and P. K. Mishra, Image enhancement techniques on deep learning approaches for automated diagnosis of COVID-19 features using CXR images, vol. 81, no. 29. Multimedia Tools and Applications, 2022.

[12] I. D. Apostolopoulos, D. J. Apostolopoulos, and N. D. Papathanasiou, “Deep Learning Methods to Reveal Important X-ray Features in COVID-19 Detection: Investigation of Explainability and Feature Reproducibility,” Reports, vol. 5, no. 2, p. 20, 2022.

[13] A. M. Ismael and A. Şengür, "Deep learning approaches for COVID-19 detection based on chest X-ray images," Expert Syst. Appl., vol. 164, p. 114054, Feb. 2021.

[14] R. Soundrapandiyan, H. Naidu, M. Karuppiah, M. Maheswari, and R. C. Poonia, “AI-based wavelet and stacked deep learning architecture for detecting coronavirus (COVID-19) from chest X-ray images,” Comput. Electr. Eng., vol. 108, no. March, p. 108711, 2023.

[15] M. M. Ahsan, K. D. Gupta, M. M. Islam, S. Sen, M. L. Rahman, and M. Shakhawat Hossain, “COVID-19 Symptoms Detection Based on NasNetMobile with Explainable AI Using Various Imaging Modalities,” Mach. Learn. Knowl. Extr., vol. 2, no. 4, pp. 490–504, 2020.

[16] M. Yildirim and A. Cinar, “A deep learning based hybrid approach for covid-19 disease detections,” Trait. du Signal, vol. 37, no. 3, pp. 461–468, 2020.

[17] S. Guefrechi, M. Ben Jabra, A. Ammar, A. Koubaa, and H. Hamam, “Deep learning based detection of COVID-19 from chest X-ray images,” Multimed. Tools Appl., vol. 80, no. 21–23, pp. 31803–31820, 2021.

[18] T. Rahman, "Covid-19 radiography database," Kaggle, [Online]. Available: https://www.kaggle.com/datasets/tawsifurrahman/covid19-radiography-database. Accessed: Nov. 19, 2023.

[19] M.S. Hossain, G.M. Shahriar, M.M.M. Syeed, M.F. Uddin, M. Hasan, S. Shivam, S. Advani, "Region of interest (ROI) selection using vision transformer for automatic analysis using whole slide images," Sci. Rep., vol. 13, no. 1, p. 11314, Jul. 13, 2023.

[20] H. E. Davies, C. G. Wathen, and F. V. Gleeson, "The risks of radiation exposure related to diagnostic imaging and how to minimise them," BMJ, vol. 342, p. d947, Feb. 25, 2011.

[21] M. Y. Ng, E. Y. P. Lee, J. Yang, F. Yang, X. Li, H. Wang, M. M. Lui, C. S. Lo, B. Leung, P. L. Khong, C. K. Hui, K. Y. Yuen, and M. D. Kuo, "Imaging Profile of the COVID-19 Infection: Radiologic Findings and Literature Review," Radiol Cardiothorac Imaging, vol. 2, no. 1, p. e200034, Feb. 13, 2020.

[22] M. Gao, B. Kang, X. Feng, W. Zhang, and W. Zhang, "Anisotropic Diffusion Based Multiplicative Speckle Noise Removal," Sensors, vol. 19, no. 14, p. 3164, Jul. 18, 2019.

[23] S. R. Syed and M. A. SD, "A diagnosis model for brain atrophy using deep learning and MRI of type 2 diabetes mellitus," Front. Neurosci., vol. 17, p. 1291753, Oct. 30, 2023.

[24] S. Liu, F. Roemer, Y. Ge, E.J. Bedrick, Z.M. Li, A. Guermazi, L. Sharma, C. Eaton, M.C. Hochberg, D.J. Hunter, M.C. Nevitt, W. Wirth, C.K. Kent Kwoh, X. Sun, "Comparison of evaluation metrics of deep learning for imbalanced imaging data in osteoarthritis studies," Osteoarthritis Cartilage, vol. 31, no. 9, pp. 1242-1248, Sep. 2023.

[25] J. Zhang, Z. Li, H. Lin, M. Xue, H. Wang, Y. Fang, S. Liu, T. Huo, H. Zhou, J. Yang, Y. Xie, M. Xie, L. Lu, P. Liu, Z. Ye, "Deep learning assisted diagnosis system: improving the diagnostic accuracy of distal radius fractures," Front. Med (Lausanne), vol. 10, p. 1224489, Aug. 17, 2023.

[26] X. Liu, L. L. Xu, Y. P. Lu, T. Yang, X. Y. Gu, L. Wang, and Y. Liu, "Deep_KsuccSite: A novel deep learning method for the identification of lysine succinylation sites," Front. Genet., vol. 13, p. 1007618, Sep. 29, 2022.

[27] X. Du, L. Si, P. Li, and Z. Yun, "A method for detecting the quality of cotton seeds based on an improved ResNet50 model," PLoS One, vol. 18, no. 2, p. e0273057, Feb. 15, 2023.