 DOI: 10.35382/TVUJS.14.8.2024.105

CONVOLUTIONAL NEURAL NETWORK MODEL APPLIED TO PNEUMONIA DETECTION

Tran Quang Quy[1*], Nguyen Duc Binh[2], Quach Xuan Truong[3], Nguyen Vu Hai[4]

Abstract – *Deep learning has become increasingly applicable across various domains, with the recent surge in available data, especially within the medical field. Its primary role is to enhance decision-making by making it more efficient, accurate, and reliable. Both deep learning and machine learning are growing applications in medicine. This is especially evident in medical areas involving different biomedical imaging types and relying heavily on collecting and analyzing vast quantities of digital images. This paper explores the utilization of advanced deep learning models, specifically convolutional neural networks, to analyze chest X-ray images and aid in accurate diagnosis. The convolutional neural network model is designed to tackle the classification challenge of identifying whether chest X-ray images indicate the presence of pneumonia, utilizing a dataset that includes both normal chest X-rays and those exhibiting signs of viral pneumonia. The study experimentally applies and assesses the accuracy of the VGG16 and ResNet50 models.*

Keywords: convolutional neural network, deep learning pneumonia, ResNet50, VGG16, X-rays.

I. INTRODUCTION

Machine learning (ML) is a subfield of artificial intelligence (AI) that enables machines to acquire knowledge and improve their performance by analyzing and learning from previous events. The process involves analyzing data, seeing trends, and addressing intricate problems. Machine learning employs two primary techniques, supervised and unsupervised learning, to attain the desired outcome. AI, machine learning, and deep learning provide additional platforms for interpreting medical images. Medical imaging requires careful consideration in various aspects such as image acquisition, pre-processing, enhancement, dynamic range adjustment, contour detection, restoration, smoothing, conversion of 2D to 3D images, artifact removal, feature extraction, segmentation, classification, object identification, detection, and visualization. AI enhances radiology, making it a more promising field for multidisciplinary clinical diagnosis. It enables physicians to do minimally invasive procedures by minimizing the extent of bodily incisions. It aids in the interpretation of sensory input, analysis of web searches, and the development of self-driving autos. Among other applications, imaging processing in medical science provides precise and high-quality visualization of many medical imaging modalities such as MRI, CT, X-ray, and others. Figure 1 expresses the fundamental category of machine learning methods.

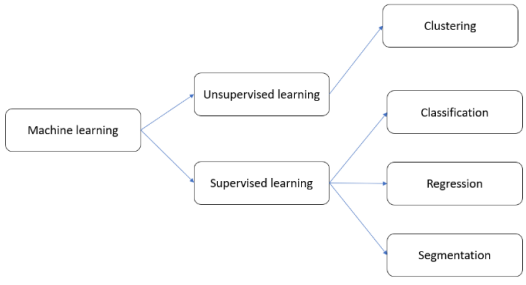

Fig. 1: Machine learning methods

The use of ML in medical imaging has fundamentally transformed the domain of diagnostic

[1,2,3,4]Thai Nguyen University of Information and Communication Technology (ICTU), Vietnam
*Corresponding author: tqquy@ictu.edu.vn
Received date: 02nd July 2024; Revised date: 24th December 2024; Accepted date: 26th December 2024

radiology, specifically in examining chest X-ray pictures. Chest X-rays are the most prevalent and extensively utilized radiography procedure, serving a critical function in the diagnosis of several illnesses, including pneumonia, lung cancer, and problems connected to COVID-19.

Figure 1 shows that supervised learning utilizes a dataset that is labeled, and from this dataset, it generates an output based on the input.

Unsupervised learning is employed to identify latent patterns in the data. It acquires knowledge and detects patterns from data that are not labeled. It adopts the internal organization of data that is not understood. Clustering and dimensionality reduction fall into this area.

The convolutional neural network (CNN) is a fundamental architecture in machine learning methodologies, particularly in image analysis tasks. CNN aims to preserve significant features within images. However, in the medical domain, datasets often lack the necessary volume to train machine-learning models effectively.

Data augmentation techniques are employed through various pre-processing methods to address this challenge. CNN utilizes both the same padding, which increases dimensionality, and valid padding, which decreases dimensionality. Max-pooling is employed to down-sample input representations, thereby reducing learning time and computational costs by selecting essential features using filters and reducing the output's pixel count, effectively diminishing dimensionality.

Rasheed et al. [1] used a CNN and Logistic regression (LR) that showed a remarkable result of 97.6–100% with principal component analysis (PCA) in COVID-19 patient identification. With COVID-19, another approach from Wang et al. [2] for applying Xception and SVM resulted in 99.33% accuracy. Several machine learning methods were analyzed and trained by Ahammed et al. [3], with average results recorded at 80.01% accuracy, except for CNN at 94.03%.

A pneumonia diagnosis used InceptionV3 CNN and other machine learning methods such as K-Nearest Neighbor, Neural Network, and Support Vector Machines, the results showed that CNN achieved the highest sensitivity of 84.1%n [4]. Later, Kim et al. [5] proposed a method that is inputted into many novel deep learning models (EfficientNet v2-M) to classify lung diseases with the U.S. National Institutes of Health (NIH) data set, the results returned as 69.33% loss and 82.15% accuracy. In 2023, Farhan et al. [6] published a new hybrid deep learning algorithm (HDLA) framework for automatic lung disease classification from X-ray images. The proposed model has improved classification performances by 95% to 99%. Thirteen thoracic lung diseases from chest X-rays were classified using the method of Kabiraj et al. [7], he proposed CX-Ultranet by using cross-entropy loss function and Efficient Net as an architecture baseline; the results achieved 88% accuracy on the average prediction.

Basic CNN has poor performance, while hybrid systems have their strength in improved accuracy without increasing the training time, Bharati et al. [8] used a hybrid deep learning framework by combining VGG, data augmentation, and spatial transformer network (STN) with CNN (VDSNet), which outperformed all existing methods as showing validation accuracy of 73%. Yimer et al. [9] developed a method using Xception deep learning and tested it on Jimma University Medical Center Radiology Department and National Institute of Health (NIH), with the accuracy recorded at 97.3%. YOLOv3 was used by Chen et al. [10] to crop the appropriate location of the lung field automatically, the work got 92.42% accuracy on abnormal X-ray images and some diseased with 71.94∼85.71% accuracy.

LungNet22 is a fine-tuned model based on VGG16 architecture, it was introduced by Shamrat et al. [11], the accuracy was 98.89%.

A. Collection of dataset

The deep learning model designed to support the diagnosis of pneumonia was experimentally trained on the Chest X-ray dataset. This dataset comprises 5,876 chest X-ray images curated by

Mooney [12], updated in 2018, and hosted on Kaggle. The chest X-ray images were collected from pediatric patients aged one to five years at the Women and Children's Health Center in Guangzhou, China.

The photos had quality screening to remove scans that were of low quality or illegible. Two proficient physicians assigned the photos diagnostic labels, and a third specialist assessed an evaluation set to guarantee the accuracy of the diagnoses.

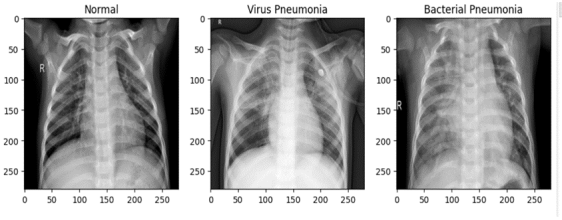

Fig. 2: X-ray image of pneumonia

In Figure 2, the chest X-ray portrayal of a healthy lung (left illustration) delineates distinct pulmonary structures devoid of any aberrant manifestations. Contrarily, the chest X-ray depiction of a patient afflicted with viral pneumonia (central illustration) exhibits disseminated 'streak-like' infiltrations across both lung fields.

The dataset has been handled by two following preparation procedures:

- Image normalization has been performed to standardize pixel values and improve model convergence during training.

- Image resizing: The dimensions of all images have been adjusted to 256 x 256 pixels, maintaining the original proportions.

While bacterial pneumonia (right illustration) typically presents as localized consolidations within pulmonary lobes. These X-ray representations are systematically categorized into three directories: train, test, and valid, respectively designating training, testing, and validation datasets. The chest X-ray dataset comprises 5,876 samples, categorized into bacterial pneumonia (2,800), viral pneumonia (1,493), and normal (1,583). It is divided into three subsets: the training set contains 5,236 samples (2,550 bacterial, 1,345 viral, 1,341 normal), the testing set includes 624 samples (242 bacterial, 148 viral, 234 normal), and the validation set has 16 samples (8 bacterial, 8 normal). This dataset provides a solid foundation for training and evaluating deep-learning models in pneumonia detection.

Dataset	Pneumonia label		Normal label	Total number of samples
	Bacterial	Virus		
Chest X-Ray	2,800	1,493	1,583	5,876
Training	2,550	1,345	1,341	5,236
Test	242	148	234	624
Validation	8	0	8	16

Fig. 3: Description of dataset

B. Data preprocessing

Image reading: Before image processing, it is imperative to read and ascertain the format consistency of the images. Verification of uniformity in format and adherence to requisite specifications is essential. While grayscale images suffice for specific tasks due to their simplicity in object recognition, color images offer richer information. When color images are necessitated for specific operations, conversion from grayscale to color may be required.

Image resizing: Variations in image sizes within datasets are commonplace. ML algorithms necessitate scaling options to standardize image dimensions. Typically, smaller-sized images yield superior performance compared to larger dimensions. Convolutional neural networks (CNNs) yield optimal and expedited outcomes when trained on datasets with uniform dimensions.

Noise reduction: Machine learning algorithms generally require minimal noise removal procedures. However, specific applications mandate the isolation of either background or foreground elements, necessitating noise reduction for image refinement and adjustment.

Data augmentation: The efficacy of machine learning techniques is contingent upon the volume of training data. Data augmentation tech-

niques, such as rotation, scaling, and transformation, are employed to augment dataset size. By applying various degrees of rotation, scaling, and transformations, the dataset can be expanded. Data augmentation, the process of generating additional data from existing datasets, is pivotal in enhancing the performance of machine learning models for segmentation and classification tasks. Figure 3 illustrates data augmentation applied to a lung chest X-ray (CXR) image.

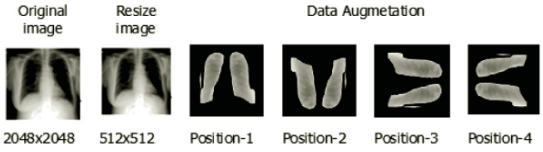

Fig. 4: Pre-processing in a lung CXR image

C. Feature extraction

The techniques are employed to detect defects in each lung X-ray image. Latent features associated with the images can also be analyzed using the Bayesian method described by Barber [13]. This approach allows for representing the conditional probability for an image sample and class, where the class is part of a set encompassing positive (1) or negative (0) cases.

$$p(c|a) = \frac{p(a|c)p(c)}{p(a)} \quad (1)$$

Given the two-dimensional grayscale chest X-ray image a, its maximum a posteriori (MAP) estimate for the most likely class is determined using a Bayesian approach within the set of classes.

$$c_{MAP} = argmax_{c \in C}\, p(c|a) \quad (2)$$

The most likely class c is associated with the image if Bayes's rule (1) allows to determine p(cla).

$$c_{MAP} = argmax_{c \in C} \frac{p(a|c)p(c)}{p(a)} \quad (3)$$

However, it is often recommended to omit the denominator p(a) as in Equation (4).

$$c_{MAP} = argmax_{c \in C}\, p(a|c)p(c) \quad (4)$$

The schema in Figure 4 illustrates a chest X-ray image with patterns in the lung region that may be affected by diseases, with the class indicating either positive or negative.

Now, a set of latent features f of a chest X-ray image is considered and discovered through the learning process and allow the association of features f to an image in the form of a vector as in Equation (5).

$$f = latent\_feature(a), f \in R^1 \quad (5)$$

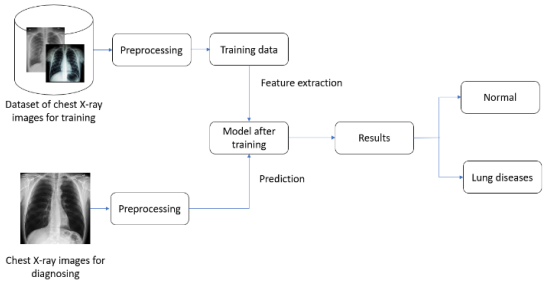

Fig. 5: Steps to prediction using CNN

Therefore, the joint probability distribution p(c,a) can be seen in detail, where the probability of class c given features f and measure of feature f given image sample a are introduced in estimating Equation (6).

$$p(c,a) \approx \sum_{f} p(c|f)p(f|a) \quad (6)$$

In the following Equation (7), detailed information about the image region surrounding each pixel x of the image a is required for the convolution operation with a kernel h that has a size of (2k + 1) * (2k + 1).

$$f(x) = h * a(x) = \sum_{dx=-k}^{k} h(dx)a(x + dx) \quad (7)$$

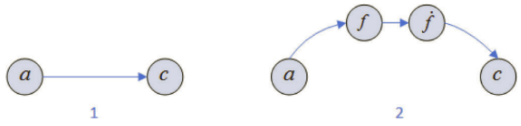

Fig. 6: (1) Relation between chest X-ray image and class; (2) Image feature is determined by CNN and then is reduced to have a lower-dimensional representation, facilitating the classification process

Since the convolution operation forms the basis of CNNs with multiple hidden layers, various CNN models can be explored for generating features f as described in Equation (7). This approach inherently facilitates the discovery of latent features from original chest X-ray images. Although these models are computationally intensive, their deep learning capabilities across multiple convolutional layers enable the extraction of specific features necessary for classification. In this study, as will be demonstrated, the task of feature creation, as per Equation (7), is executed using CNN, including ResNet50 and VGG16.

II. DEEP LEARNING IN CLASSIFICATION X-RAY IMAGES

This work employed deep learning models to categorize X-ray images for the classification of lung diseases. The study employed two widely recognized CNN structures: ResNet50 and VGG16.

A. Methodology

The methodology includes the following phases in the classification of lung pictures. Every medical analysis necessitates adherence to specific obligatory steps. Pathological tests require high-resolution images to identify areas of abnormality in the human body accurately. Efficient medical image analysis is crucial for accurate and timely treatment of patients. Each discipline emerges to aid individuals in leading a tranquil and harmonious existence. Physicians and radiologists rely on the machine learning techniques they utilize to provide timely assistance in the field of medicine. Machine learning techniques achieve their success by utilizing CNN with various activation functions, several filters, and the Softmax function.

Computer algorithms are utilized to process digital photographs using the technique of image processing. It offers a broader selection of algorithms for image processing to improve image quality and highlight key elements.

An image is a two-dimensional array with values ranging from 0 to 255. Mathematically, it can be expressed as the function *f(x,y)*. The variables x and y represent the horizontal and vertical coordinates, respectively. These coordinates indicate the pixel value of a picture.

As can be seen in Figure 5, the process of problem-solving entails a series of fundamental steps:

Data input: Input data are divided into two parts, the first one is for training data, and the second is for predicting images.

Preprocessing: This step is critical in building deep-learning models for image classification.

Model after training: After being trained with deep learning, the model received from training saves the best parameters for X-ray image prediction.

Result: This step is to evaluate the model, and the doctor can see the prediction of chest X-ray image diagnosis.

B. Preprocessing images

Data preprocessing is an essential step in building deep-learning models for image classification. In this study, we utilized several preprocessing techniques to enhance the quality and variability of the X-ray images used in training our models. These techniques include resizing, rescaling, and augmentation.

The dataset contained images with different resolutions, demanding the standardization of resolutions to provide uniformity and comparison among different models. All photos were adjusted to a consistent resolution of 180 x 180 pixels to

achieve this objective. The size of this uniform met the input specifications of the CNNs employed in this investigation, namely the VGG16 and ResNet50 architectures. Resizing the photos guarantees that the models are provided with input images of a uniform size, which is essential for efficient feature extraction and classification.

Rescaling is a crucial preprocessing step. During this procedure, the pixel values of the photos were reduced by dividing each value by 255. This process converts the pixel values, which are initially within the range of 0 to 255, to a range of 0 to 1. Rescaling data is beneficial as it normalizes the dataset, reducing computational complexity and improving model efficiency. Rescaling the pixel values standardizes them, which improves convergence during training and promotes the stability and performance of CNN models.

C. CNN architectures and hyperparameters

The two CNN models applied in the study (VGG16, ResNet50) were trained on the same hardware configuration, support libraries, and model parameter settings.

The research used a computer with the settings as follows:

- CPU: Intel Xeon 2690 V0, 8 cores and 16 threads
- GPU: RTX 3070 8GB VRAM
- RAM: 32 GB

This study employed the TensorFlow 2.15 library and the Keras API to implement the CNN algorithms rapidly. This library is specifically designed and developed for deep learning purposes, providing essential tools and functions for building and training neural network models. This combination of hardware and support libraries allowed the study to efficiently and quickly train the CNN models to classify lung X-ray images. Regarding parameter settings, the CNN models were configured with a final fully connected layer comprising two output nodes to perform the classification. The Softmax activation function was used. The training process for the models was conducted over 20 epochs, with a learning rate of 0.001 for the entire training process.

The selection of ResNet50 and VGG16 architectures for pneumonia detection is based on their proven effectiveness in image classification tasks. ResNet50 is known for its deep residual learning capabilities, allowing it to effectively address vanishing gradient issues in deep networks, which is critical for accurately detecting subtle features in chest X-ray images. VGG16, on the other hand, offers a simpler, yet highly effective architecture with smaller convolutional filters and a deep network structure, making it efficient for learning spatial hierarchies in image data, which is beneficial for distinguishing between healthy and pneumonia-affected regions in medical images.

VGG Architectures

The VGG16 deep convolutional neural network (DCNN) model was created by Simonyan et al [14]. Expanding the depth of the VGG model can enhance the ability of the kernels to acquire more intricate characteristics. This network has been trained using a dataset of over a million photos from the ImageNet collection. VGG16 is a deep convolutional network consisting of 16 layers, including 13 convolutional layers and three fully connected layers, with ReLU (Rectified Linear Unit) activation functions used throughout the network. The conv1 layer receives input in the form of 224×224 RGB images of a standardized size. The image is processed using a series of filters specifically designed to produce the smallest receptive field, which is 3×3. In addition, one of the pairs employs a 1 x 1 convolutional filter, which alters the input channels in a linear manner. The convolution stride is set to one pixel, and the spatial padding of the input to the convolution layers is modified to maintain the spatial resolution after convolution. The padding for 3×3 convolutional layers is set at 1 pixel. After the convolutional layers, five max-pooling layers are used to accomplish spatial pooling. The max-pooling operations are being performed with a stride of 2 on 2×2 pixel frames. Three consecutive convolutional layers are stacked on top of each other. The first two containers have a total of 4,096 channels each, while the third container

performs ILSVRC classification on a 1000-way basis. Ultimately, a SoftMax layer is present. The structure of the complete interconnected tiers is uniform across all networks.

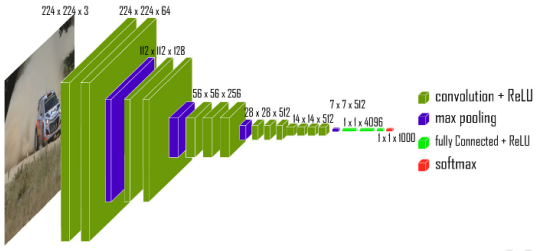

Fig. 7: VGG16 Architecture [15]

ResNet50 architecture

ResNet50 used in this work consists of 48 fully connected layers, along with a max pool layer and an average pool layer, resulting in a total of 50 layers [16]. It utilizes residual connections (skip connections) to allow the gradient to flow more effectively through the network during training, and the ReLU (Rectified Linear Unit) activation function is applied to introduce non-linearity, enabling the model to learn complex patterns in the data. It can perform floating-point calculations exceeding 3.8 $\times$ 109. The ResNet50 architecture employs a combination of convolutional filters of different sizes to address the problem of decay inherent in CNN models and reduce the training time associated with the deep structure. ResNets have a smaller number of filters, which allows them to execute operations more rapidly. The 34-layer ResNet achieves a performance of 3.6 billion FLOPS, whereas the shorter 18-layer ResNets get a performance of 1.8 billion FLOPS. This design may be trained using around 23 million variables. The network can process an input image with divisible dimensions by 32 in terms of width, height, and channel width. The ResNet design uses a 7 $\times$ 7 kernel size for the first convolution and a 3 $\times$ 3 kernel size for max pooling. Each pair of layers in the 34-layer network is replaced with a three-layer bottleneck block, resulting in a final 50-layer ResNet.

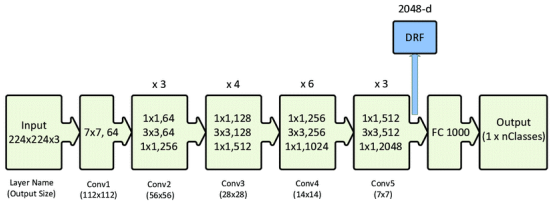

Fig. 8: ResNet50 Architecture [17]

D. Evaluation metrics

The diagnosis of lung diseases is a task of classification. To evaluate the performance of the models in this picture recognition task, the study employ the metrics of Accuracy, Precision, and Recall. Assume that photos without lung disorders are labeled as Positive (+1), whereas those with lung diseases are labeled as Negative (-1). When the categorization model analyzes the photos, four scenarios will arise, namely: TP (True Positive): samples correctly classified as positive (actually positive).

- FP (False Positive): samples incorrectly classified as positive (actually negative).

- TN (True Negative): samples correctly classified as negative (actually negative).

- FN (False Negative): samples incorrectly classified as negative (actually positive).

The metrics including accuracy, precision and recall are provided in Equations (8), (9), and (10), respectively.

$$Accuracy = \frac{\sum TP + TN}{\sum TP + FP + FN + TN} \quad (8)$$

With precision:

$$precision = \frac{\sum TP}{\sum TP + FP} \quad (9)$$

With recall:

$$recall = \frac{\sum TP}{\sum TP + FN} \quad (10)$$

E. Results

The team achieved positive results after training the CNN models using a training dataset consisting of 3,242 chest X-ray pictures that were classified into two distinct class labels. The results were evaluated using four fundamental metrics: loss, accuracy, precision, and recall throughout the training and validation procedure.

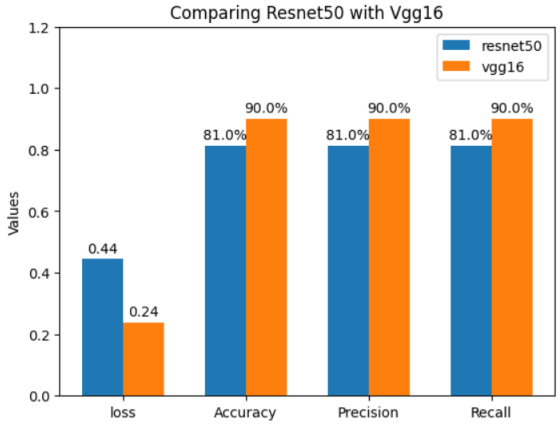

Fig. 9: Comparing CNNs between Resnet150 and VGG16

Regarding the evaluation results of the models using ResNet50 and VGG16 architectures on the training and validation sets in the above images, the accuracy and loss functions of VGG16 are more stable. The performance of VGG16 is higher than that of ResNet50.

The impressive results of the deep learning model using the VGG16 architecture highlight its potential for developing an application to assist in the diagnosis of pneumonia through X-ray images. This can effectively shorten the examination time, ensure patients receive timely treatment, and reduce the risk of disease progression.

Based on the figures provided, this section evaluates the performance of ResNet50 and VGG16 using Precision, Recall, F1-score, and Accuracy. Both models improve during training, but VGG16 maintains more stable performance across epochs, with higher accuracy, precision,

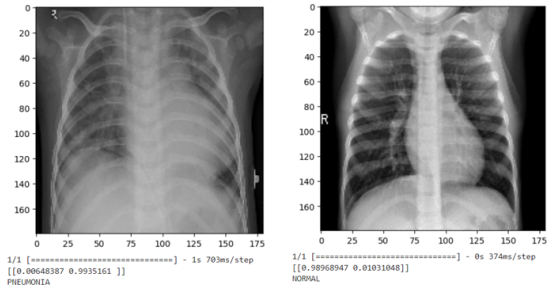

Fig. 10: Output of the model with the input of (left) PNEUMONIA (right) NORMAL

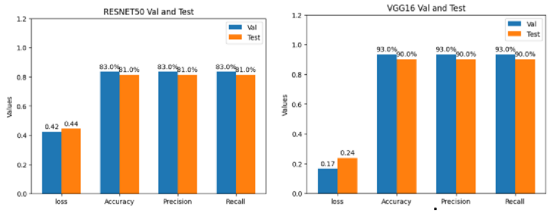

Fig. 11: Compare overall performances of ResNet50 and VGG16

and recall values. ResNet50 shows initial improvement but stabilizes at a lower level.

- Accuracy: VGG16 has higher accuracy on both training and test sets.

- Precision: VGG16 outperforms ResNet50 in correctly classifying positive samples.

- Recall: ResNet50 improves recall but remains lower than VGG16.

- F1-score: VGG16 achieves a higher F1-score, indicating better overall performance.

Following an extensive research and development phase, a program was created to compare and assess the performance of two prominent CNN models, ResNet50 and VGG16, in classifying X-ray pictures into pneumonia and normal. The findings indicated that the ResNet50 model exhibited a mere 81% accuracy, whereas the VGG16 model had a superior accuracy of up to 90%, hence significantly surpassing the ResNet50 model in its ability to classify lung diseases.

Regarding performance, VGG16 may outper-

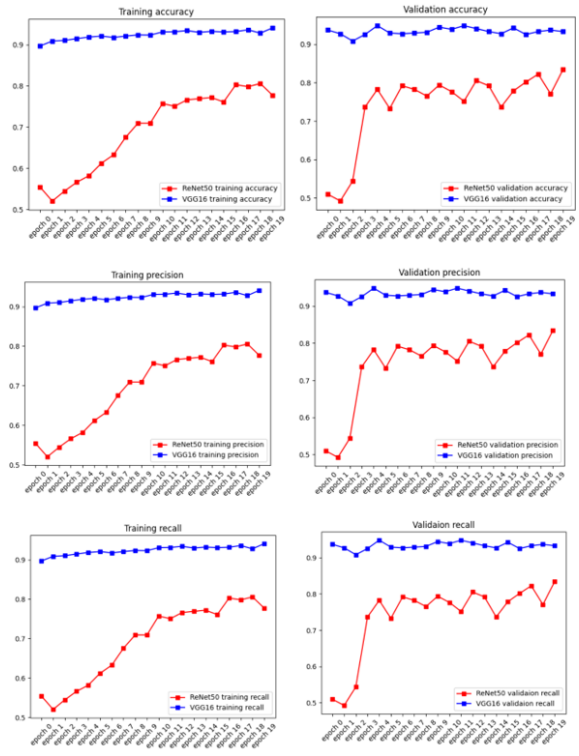

Fig. 12: Performances of Resnet150 and VGG16

form ResNet50 in pneumonia detection due to its more straightforward yet highly effective architecture with smaller convolutional filters and deeper layers, which allows it to capture fine-grained spatial features essential for distinguishing subtle differences in chest X-ray images. ResNet50, while incorporating residual connections to mitigate vanishing gradient issues and enable more profound network training, may face challenges in this specific task due to its more complex structure and the increased difficulty of learning from noisy or subtle medical image data. Additionally, VGG16's consistent performance with relatively fewer parameters than ResNet50 might make it more robust in scenarios with limited training data, such as medical image datasets.

III. CONCLUSION

The study demonstrated the utilization of deep learning models, specifically CNNs, for categorizing chest X-ray images to diagnose pneumonia. The performance of two architectures, ResNet50 and VGG16, was evaluated and compared using a dataset consisting of 3,242 X-ray pictures. The findings revealed that the VGG16 model exhibited superior performance compared to ResNet50 regarding stability and overall effectiveness. This was evident from the consistent accuracy and loss functions during training and validation.

The exceptional performance of the VGG16 architecture indicates its potential usefulness in creating a diagnostic tool for pneumonia utilizing X-ray pictures. An instrument of this nature has the potential to diminish significantly the duration needed for diagnosis, guaranteeing prompt medical intervention for patients and lowering the likelihood of advancement of disease. This work highlights the potential of deep learning models in advancing medical diagnostics and improving patient care by utilizing more efficient and accurate illness detection approaches.

ACKNOWLEDGEMENTS

This research was sponsored by the Ministry of Education and Training, project 'Research on application of machine learning model in analysis electronic medical records gastrointestinal disease', B2022-TNA-24.

REFERENCES

[1] Rasheed J, Hameed AA, Djeddi C, Jamil A, Al-Turjman F. A machine learning-based framework for diagnosis of COVID-19 from chest X-ray images. *Interdisciplinary Sciences: Computational Life Sciences*. 2021;13: 103–117. https://doi.org/10.1007/s12539-020-00403-6.

[2] Wang D, Mo J, Zhou G, Xu L, Liu Y. An efficient mixture of deep and machine learning models for COVID-19 diagnosis in chest X-ray images. *PLoS One*. 2020;15(11): e0242535. https://doi.org/10.1371/journal.pone.0242535.

[3] Ahammed K, Satu M, Abedin M, Rahaman M, Shariful Islam SM. Early detection of coronavirus cases using chest X-ray images employing machine learning and deep learning. *MedRxiv* [Preprint] 2020. https://doi.org/10.13140/RG.2.2.13579.11045.

[4] Yee SLK, Raymond WJK. Pneumonia diagnosis using chest X-ray images and machine learning. In: *Proceedings of the 2020 10th International Conference on Biomedical Engineering and Technology*. New York,

USA: Association for Computing Machinery; 2020. p.101–105.

[5] Kim S, Rim B, Choi S, Lee A, Min S, Hong M. Deep learning in multi-class lung diseases' classification on chest X-ray images. *Diagnostics*. 2022;12(4): 915. https://doi.org/10.3390/diagnostics12040915.

[6] Farhan AMQ, Yang S. Automatic lung disease classification from the chest X-ray images using hybrid deep learning algorithm. *Multimedia Tools and Applications*. 2023;82(25): 38561–38587. https://doi.org/10.1007/s11042-023-15047-z.

[7] Kabiraj A, Meena T, Reddy PB, Roy S. Detection and classification of lung disease using deep learning architecture from X-ray images. In: *Advances in Visual Computing*. Cham: Springer International Publishing. 2022; p. 444–455. https://doi.org/10.1007/978-3-031-20713-6_34.

[8] Bharati S, Podder P, Mondal MRH. Hybrid deep learning for detecting lung diseases from X-ray images. *Informatics in Medicine Unlocked*. 2020;20: 100391. https://doi.org/10.1016/j.imu.2020.100391.

[9] Yimer F, Tessema AW, Simegn GL. Multiple lung diseases classification from chest X- ray images using deep learning approach. *International Journal of Advanced Trends in Computer Science and Engineering*. 2021;10(5): 2936–2946. https://doi.org/10.30534/ijatcse/2021/021052021.

[10] Chen KC, Yu HR, Chen WS, Lin WC, Lee YC, Chen HH, et al. Diagnosis of common pulmonary diseases in children by X-ray images and deep learning. *Scientific Reports*. 2020;10(1): 17374. https://doi.org/10.1038/s41598-020-73831-5.

[11] Shamrat FJM, Azam S, Karim A, Islam R, Tasnim Z, Ghosh P, et al. LungNet22: a fine-tuned model for multiclass classification and prediction of lung disease using X-ray images. *Journal of Personalized Medicine*. 2022;12(5): 680. https://doi.org/10.3390/jpm12050680.

[12] Mooney P. *Chest X-ray images (Pneumonia)*. https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia. [Accessed 25th March 2018].

[13] Barber D. *Bayesian reasoning and machine learning*. Cambridge, United Kingdom: Cambridge University Press; 2012.

[14] Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv* [Preprint] 2014. https://doi.org/10.48550/arXiv.1409.1556.

[15] GeeksforGeeks. *VGG-16 and CNN model*. https://www.geeksforgeeks.org/vgg-16-cnn-model/ [Accessed 1st June 2024].

[16] Muhathir, Dwi Ryandra, MF, Syah RB, Khairina N, Muliono R. Convolutional neural network (CNN) of Resnet-50 with Inceptionv3 architecture in classification on X-ray image. In: *Computer Science On-line Conference*. Cham: Springer International Publishing; 2023. p.208–221. https://doi.org/10.1007/978-3-031-35314-7_20.

[17] Mascarenhas S, Agarwal M. A comparison between VGG16, VGG19 and ResNet50 architecture frameworks for image classification. In: *2021 International conference on disruptive technologies for multi-disciplinary research and applications (CENTCON)*. 19th–21st November 2021; Bengaluru, India. IEEE; 2021. p.96–99. https://doi.org/10.1109/CENTCON52345.2021.9687944.