LUNGINFORMER: A lung pneumonia disease detection using contrast enhancement and hybridization InceptionResNet and Transformer

Hanafi [a,1,*]

[a] Magister Informatic Engineering, Universitas Amikom Yogyakarta, Sleman, 55283, Indonesia
[1] hanafi@amikom.ac.id
* corresponding author

ARTICLE INFO

Article history
Received January 27, 2025
Revised May 1, 2025
Accepted May 7, 2025
Available online May 31, 2025

Keywords
Lung disease
CLAHE
InceptionResNet
X-ray images
Transformer

ABSTRACT

Lung pneumonia is categorically a serious disease on Earth. In December 2019, COVID-19 was first identified in Wuhan, China. COVID-19 caused severe lung pneumonia. The majority of lung pneumonia diseases are diagnosed using traditional medical tools and specialized medical personnel. This process is both time-consuming and expensive. To address the problem, many researchers have employed deep learning algorithms to develop an automated detection system for pneumonia. Deep learning faces the issue of low-quality X-ray images and biased X-ray image information. The X-ray image is the primary material for creating a transfer learning model. The problem in the dataset led to inaccurate classification results. Many previous works with a deep learning approach have faced inaccurate results. To address the situation mentioned, we propose a novel framework that utilizes two essential mechanisms: advanced image contrast enhancement based on Contrast Limited Adaptive Histogram Equalization (CLAHE) and a hybrid deep learning model combining InceptionResNet and Transformer. Our novel framework is named LUNGINFORMER. The experiment report demonstrated LUNGINFORMER achieved an accuracy of 0.98, a recall of 0.97, an F1-score of 0.98, and a precision of 0.96. According to the AUC test, LUNGINFORMER achieved a tremendous performance with a score of 1.00 for each class. We believed that our performance model was influenced by contrast enhancement and a hybrid deep learning model.

1. Introduction

Numerous studies have examined the efficacy of computerized diagnoses in the field of medicine, which need collaboration between doctors and computer scientists. Certain automation diagnostic tools used in medicine are capable of being classified as highly sophisticated as they strive to replicate the judgments provided by doctors and other healthcare workers. Moreover, machine-learning detection methods employed in medicine has the capacity to systematically examine complex and vast health information [1], [2]. Healthcare practitioners may employ deep learning identification tools to acquire novel insights on data and employ this knowledge to improve their effectiveness of assessment. Therefore, engines are considered machines of intelligence since they employ a feedback loop to continuously enhance their capabilities. Assessing vast clinical data can be a formidable task. The application of analyzing information, artificial intelligence (AI), and algorithms for learning mechanism

in enhanced computer-based detection provides substantial benefits in the detection of various illnesses and medical issues.

The coronavirus (COVID-19) pandemic originated and rapidly disseminated worldwide during the latter months of 2019, giving rise to a concerning scenario. The illness initially manifested in the Chinese town of Wuhan around the last month of 2019. It was officially designated as an international health crisis by the World Health Organization (WHO) at the beginning of 2020. As of the third month of 2020, the WHO formally classified it as an International pandemic [3]. A few of the symptoms caused by the Coronavirus include respiratory infections, persistent coughing, higher temperatures, and fatigue. The reverse transcription-polymerase chain reaction (RT-PCR) technique is employed to identify individuals who have contracted the virus. However, the procedure for using this evaluation technique can be time-consuming, typically requiring many hours or even days to yield results [4]. RT-PCRs are known for their high labor requirements and high processing costs. Therefore, experts face significant challenges in developing solutions that utilize identification technology. Current applications of AI include the computerization of disease identification. The current evidence indicates that AI achieves exceptional outcomes in the automated classification of photographs by utilizing various deep learning techniques. The method of identification through computer imagery relies on the classification features obtained from the Chest X-ray image (CXR), as shown in Fig. 1. Furthermore, artificial intelligence refers to programs that can acquire knowledge and make decisions by analyzing massive data sets.

(a) COVID-19 lung image (b) Normal lung image (c) Pneumonia viral lung image

Fig. 1. X-ray image for some example lung diseases

The concept of deep learning was initially developed in the past decade under the umbrella of AlextNet [5]. AlexNet gained immense popularity as a deep learning approach owing to its remarkable accomplishments in scientific programs particularly in the area of visual retrieval. AlexNet consists of convolutional neural network (CNN) models. Following the development of AlexNet, several deep learning approaches emerged, such as VGG-16, GoogleNet, RestNet, DenseNet, MobileNet, and others. Typically, layers consisting of a few convolutional and layers with maximum pooling are employed, with full connection and SoftMax methods applied at the conclusion of the stacking process. Recurrent neural networks (RNN), which consist of long short-term memory (LSTM), gated recurrent units (GRU), and auto-encoder (AE), constitute another deep learning paradigm [6]. Previous studies have demonstrated the exceptional performance of deep learning methods in the domains of cyber security detection algorithms [7], [8] and recommender system for e-commerce business [9]–[15], image processing for medical domain [16], [17], Lung cancer detection using hybrid deep CNN model [18], Covid-19 detection based on GAN (Generative Adversarial Network) [19]. Nevertheless, most of the models using deep learning mentioned before have a large number of variables. Moreover, majority of automatically lung diseases detection faced the problem inaccurate classification.

Deep learning is the predominant artificial intelligence method for recognizing lung diseases included COVID-19, lung pneumonia, and viral virus. Among the more widely used types of algorithms for deep learning is the VGG (Visual Geometry Group). The VGG architecture comprises distinct variations,

namely VGG-16, VGG-19, and VGG-32. The structure of the VGG is formed by combining two convolutional techniques, which incorporate the ReLU (Rectified Linear Unit) activation function. Following the ReLU processing, the data undergoes maximum pooling and is connected to several higher layers. The entire procedure utilizes the ReLU activation function, while the final phase employs the Softmax function to generate the classification results. A research study used VGG-16 for the automated detection of COVID-19 [20]. This research examined four distinct categories of datasets: standard, COVID-19, bacterial, and pneumonia images. Based on the evaluation criteria, the VGG-16 model attained an accuracy of 87.49%.

The CNN architecture known as DenseNet was introduced by Gao Huang *et al*. in 2017 [21]. More precisely, researchers employed complete output stages that were connected with each of the following nodes situated inside a thick layer. Hence, a network demonstrates a significant degree of multilayer connectivity, so validating its classification as DenseNet. The use of the idea showcases effectiveness in the particular setting of embracing traits. This approach accomplished a substantial decrease in the total number of networks. The structure of DenseNet consists of several solid components and transforming blocks that are deliberately distributed across two sequential solid-block stages. In terms of COVID-19 categorization, the DenseN*et al*gorithm demonstrates exceptional performance [22]. A different study applied a DenseN*et al*gorithm to X-ray images of lungs with normal, COVID-19, or pneumonia conditions. The research result demonstrated that this DenseN*et al*gorithm attained an accuracy rate of 96.25%.

The adoption of deep learning algorithms faced the essential problem in multi-class lung disease detection, including COVID-19. The majority of them result in inaccurate detection for a multi-class classification task. In this study, we attempted to reduce misclassification with several novel scenarios, including:

- Enhance the quality of XCR image with the augmentation process and adoption of contrast enhancement using CLAHE.
- Enhance the classifier algorithm using a hybrid Inception-ResNet and Transformer mechanism. InceptionResNet is responsible for feature extraction of CXR, while Transformer handles the feature selection task.
- Incorporated the augmentation process, contrast enhancement process for the CXR image into the InceptionResNet and the Transformer algorithm

2. Related Works

CT (Chest Computed Tomography) and X-ray images are frequently used in various countries as a viable method for detecting COVID-19. Nevertheless, the procedure for identifying COVID-19 is complex and requires the use of laboratory radiography on individuals. Cancer of the breast is a significant cause of death in people. An expedited identification has the potential to enhance human mortality rates [23]. The application of methods such as image processing and machine learning has demonstrated significant potential in the field of pulmonary cancer detection. This part provides a comprehensive assessment of neural network algorithms used for tuberculosis (TB), COVID-19, pulmonary cancer, and pneumonia.

Applying VGG-16, ResNet-50, and InceptionV3, which are machine learning technologies, to medical images of pulmonary conditions and COVID-19 produced encouraging outcomes. Pneumonia has been identified as one of the most common signs of COVID-19. Transfer learning facilitates the identification of a common viral agent responsible for both pneumonia and COVID-19. A research investigation reveals that the expertise gained from an experiment in recognizing viral pneumonia can be applied to identify COVID-19. Therefore, Haralick characteristics may be utilized to simplify the process of extracting characteristics. This methodology entails conducting a statistical analysis that specifically targets a specific aspect of COVID-19 diagnosis. Transfer learning has consistently demonstrated substantial results compared to conventional classifiers [20].

Several investigations have created and examined a fully computerized system for detecting COVID-19 using CXR. The COVID-19 artificial intelligence method was utilized to gather visual information from 3D CT scans of the chest to aid in the diagnosis of COVID-19. The results demonstrate that the technique outperforms previous research. This linked research utilized model training of CNN structures, including Inception-ResNetV2, ResNet152, ResNet50, InceptionV3, and ResNet101, to detect COVID-19 pneumonia by analyzing CXR pictures. Out of all the available models, ResNet50 had the highest level of accuracy in classifying results [24].

The GSA-DenseNet121-COVID-19 is a distinctive CNN structure that combines DenseNet121 with an enhancement approach known as GSA (gravitational searching). The DenseNet121-COVID-19 model demonstrates superior performance in identifying COVID-19 compared to other DenseNet121 models, which achieved a diagnostic accuracy of 94% on the benchmark set. The proposed approach was compared to an Inception-V3 CNN framework, which was performed by hand to calculate hyperparameter estimations. The GSA-DenseNet121 algorithm surpassed the benchmark method, which was just able to classify 95% of the specimens in the dataset [25].

Model training using EfficientNet has been improved through kernel PCA (Principal Component Analysis). Subsequently, these characteristics were combined through an attribute fusion method. Ultimately, stacking ensemble meta-classifiers were developed to categorize the algorithm in two separate phases [26]. At the starting point, projections were generated utilizing an SVM and a random forest approach. These projections were subsequently combined and used in the subsequent phase. In a similar study, scientists employed ResNet32 and the deep transfer learning method to classify individuals who received an infection with COVID-19, and the findings were subsequently reported. When analyzing the COVID-19 classification against prior supervised training designs, empirical evidence showed the algorithm achieved better results than the prior algorithms [27].

Immediate identification of lung cancer significantly improves the chances of survival, increasing the range from 14% to 49%. While CT techniques are generally more precise than X-rays, an authoritative diagnosis sometimes requires the use of many different kinds of imaging. A computerized DNN (Deep Neural Network) has the ability to accurately detect lung cancer in CT images. Thus, research has suggested the utilization of an adaptive boosting method with a Dense-Net to accurately categorize lung images as either normal or cancerous. The database for training consists of 201 chest photos, of which 85 percent are used to learn and the other 15 percent are applied to evaluate identification. The suggested method has demonstrated an assessment performance of 90%. Based on the results of the research, the classifier based on MLP and CNN with Adaboost had a superior performance of 90.85% compared to the other [28].

The CNN, DNN, and dense auto-encoder algorithms have been applied for the purpose of detecting cancerous calcium. The classification systems were used to classify CT images of both malignant and benign nodules in the lungs. The LIDC dataset analyzed the connections with an accuracy rating of 84.15%, a sensitivity of 83.96%, and a specificity of 84.32% [24]. The network of CNN demonstrated the highest level of precision among the three networks.

A study employed a machine learning ensemble classification algorithm, SVM, and KNN to categorize and differentiate between COVID-19 and pneumonia, addressing the obstacles faced by doctors in this task. Nevertheless, an RNN (Recurrent Neural Network) equipped with LSTM (Long Short-Term Memory) was suggested as a deep learning framework for the detection of lung diseases. The tests showed that the proposed approach was both resilient and efficient [29]. A further investigation employs an ensemble of InceptionResNetV2, ResNet50, and MobileNetV2 models for the purpose of classification. The results showed that the ResNet50, MobileNetV2, and InceptionResNetV2 methods achieved an F1 score of 94.84%, which was greater than the results of the other designs [30].

A study used optimal deep neural network, shorted ODNN, and LDA (Linear Discriminant Analysis) were utilized to assess CT chest pictures by reducing the amount of information of deep characteristics. The ODNN, when paired with CT images and enhanced utilizing the Gravitational Search Method, is

employed for the classification of lung disease. This approach yields a sensitivity of 96.2%, specificity of 94.2%, and accuracy of 94.56% [31]. Furthermore, transfer learning methods are employed to classify COVID-19, pneumonia, and normal people by utilizing a CNN with weights that have been trained. The database successfully classified individuals with both circulating SARS-CoV-2 and pneumonia, indicating a significant finding in the study [32]. A separate investigation explored the possibility of utilizing machine learning, specifically Retina Net and Mask R-CNN, as a combined approach to accurately identify and locate pneumonia in CXR photographs. This approach achieved a recall rate of 0.793 for a substantial database [33]. The transfer methods were employed to gather pictures of CXR and CT scans for various chest illnesses. Due to the similarities between COVID-19 and pneumonia, the identification of COVID-19 is difficult. Aim to handle the problem requires a comprehensive analysis of an Individual's medical Images. The objective is achieved by employing an innovative structure that has been trained to recognize the illness caused by viruses, specifically to detect COVID-19. The results of learning transferred are significantly distinct from those of typical classifications

3. Method

In this study, we explore several approaches designed to improve the classification task for lung disease detection. Our algorithm duplicated LUNGINFORMER, where the main algorithm consists of Inception and Transformer. The details of several processes and the implementation of primary materials are shown in the subsection.

3.1. Datasets of Lung Diseases

In this study, we considered implementing real lung disease datasets. 1823 pictures of an identified poster anterior (PA) perspective of CXR imagery make up the dataset utilized for this investigation. Three distinct datasets were utilized for COVID-19 cases, whereas three separate datasets were used for viral pneumonia, non-pneumonia, or normal cases. Annotated optical coherence tomography (OCT) and CXR photos were used for these conditions. The dataset comprises 536 images of COVID-19, 619 images of viral pneumonia, and 668 images of normal patients. The dataset's COVID-19 cases range in age from 18 to 75 years; Table 1 shows the characteristic image specifications used in the dataset. COVID-19 images exhibit significant height and width variance, with maximum height and maximum width in comparison to other image classes. Sample images of COVID-19 patients, viral pneumonia, and normal people are demonstrated in Fig. 1 where a normal CXR image shows clear lungs devoid of any abnormal opacification or pattern; COVID-19 (extreme right) shows ground-glass opacification and consolidation in the right upper lobe and left lower lobe; and viral pneumonia (middle) shows a more diffuse interstitial pattern in both lungs. In our experiment scenario, the dataset was split in an 80-20 ratio for the purpose of training our deep learning model, with 80% of the dataset being utilized for training and 20% for testing. Table 1 demonstrates the distribution of images in the training and test sets. In order to create effective deep learning models, all CXR pictures were first inspected for quality control by removing any scans that were of poor quality or unreadable.

Table 1. Characteristic of lung disease datasets

Type Classes	Training (80%)	Testing (20%)	Total/class (100%)
Normal	495	123	619
Covid-19	428	107	536
Virus	534	133	668
Total	363	1457	1823

3.2. Feature Extraction with InceptionResNet

The InceptionResNet is a sophisticated CNN design that integrates the Inception component with the residual network. This combination results in a robust and effective architecture specifically designed for visual recognition applications. The Inception component partitions the data being supplied across several individual parallel streams, applies specialized adjustments to each route, and then combines the results. The residual link enables the learning network to acquire knowledge of the residual relationship

among the inputs and the results, therefore facilitating the efficient learning of extremely complicated structures.

The adoption residual mechanism, InceptionResNet, builds upon the achievements of earlier Inception designs, including GoogleNet. Previous studies have demonstrated that residual relationships may improve the efficiency and effectiveness in developing sophisticated neural networks by addressing the issue of vanishing gradients [34]. In this study, we address this issue. We adopted InceptionResNet as a feature extraction task. The transformer mechanism is responsible for the feature selection task.

3.3. Feature Selection with Transformer

The Transformer method is an innovative framework throughout the domain of AI. The algorithm has generated considerable interest in the recent past. In its essence, the Transformer is a neural network design that exclusively depends on the attention technique, eliminating the conventional recurrent and convolutional components. The Attention mechanism, an essential feature of the Transformer, facilitates the model's acquisition of knowledge about the connections among various data points in a sequence, therefore enabling it to efficiently capture long-range dependencies. This stands in sharp opposition to recurrent networks, which iteratively process sequence components and are limited to providing Attention to short-term context [2]. The complete model of our model, named LUNGINFORMER, is shown in Fig. 2.

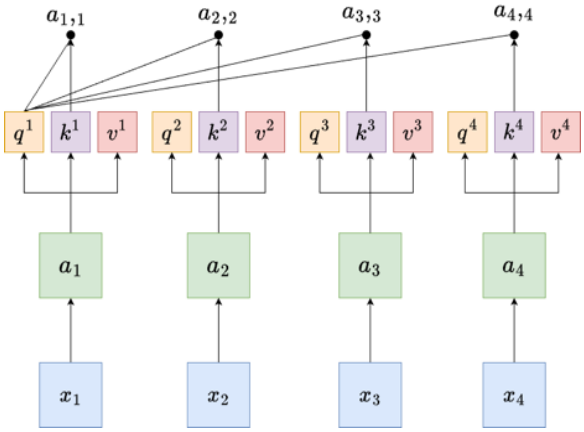


Fig. 2. Self-attention module architecture

The Self-Attention applied to Transformer neural networks is a standardized dots in attention method. Fig. 2 displays the characteristic vector of input x_i, denoted as ai, which is then mapped to the matching q_i, k_i as well as v_i data. The Self-Attention technique executes matrix operations for every q as well as k, in other words, dot multiplication processes. Prior to therefore, it is necessary to normalize each input mapping in order to obtain the attention value vector. The complete Transformer module show on Fig. 3.

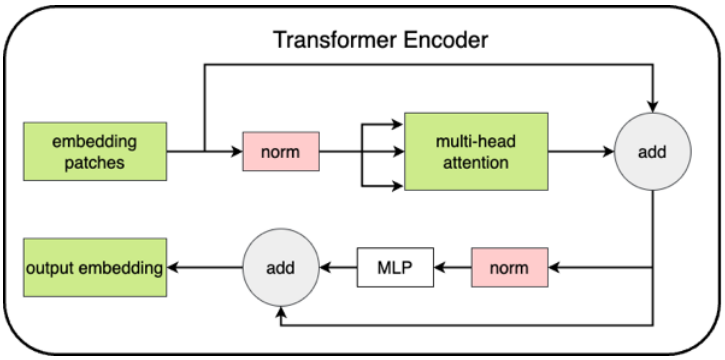


Fig. 3. Transformer module

Suppose the vector of query dimension of key is d_k, while number of vector mapping in the term of d_v, the value multiplied calculation of every key factor was computed by $\sqrt{d_k}$. Where the weight factor was computed using Softmax mechanism.

$$Attention\ model(Q, K_i, V_i) = softmax\left(\frac{Q^T K_i}{\sqrt{d_k}}\right) V_i \quad (1)$$

Where the value Q, V and K can be computed with formula as follow:

$$Attention\ model(Q, K, V) = softmax\left(\frac{Q^T K}{\sqrt{d_k}}\right) V \quad (2)$$

In the encoder section contain Add layer and Norm module. Both can be computed with formula as follow:

$$Norm(x + head - multi - attention(x)) \quad (3)$$

$$Norm(x + feed - forward(x)) \quad (4)$$

Where x symbolizes feed-forward and head multi-attention representation, in terms of the feed-forward layer, they are linked by a fully connected layer. They implemented the ReLU activation function.

In this study, we implemented a hybrid model of Inception-ResNet and Transformer. Where InceptionResNet is responsible for creating feature extraction of the x-ray. A transformer is accountable for creating feature selection for a classification task. The output of the transformer consisted of three classes: Normal, COVID-19, and Virus. The details of hybridization are shown in Fig. 4.

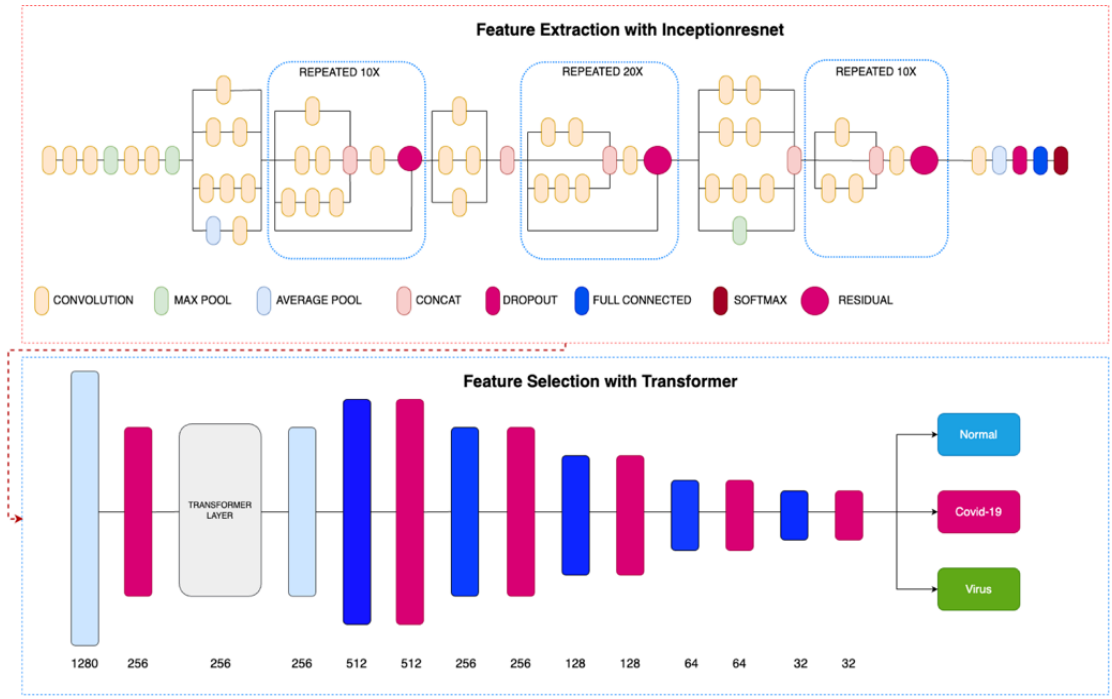


Fig. 4. LUNGINFORMER: Novel deep learning framework for Covid-19 detection using hybridization of InceptionResNet and Transformer

3.4. Contrast enhancement with CLAHE

The clarity and characteristic extraction associated with digital photos can be dramatically improved by enhancing contrast, a crucial step in image processing. Histogram equalization is a frequently employed method for enhancing contrast. Its objective is to modify the distribution of the brightness of

pixels for the purpose of providing an additional consistent histogram. Nevertheless, conventional techniques for histogram equalization may occasionally result in undesired distortions, including excessive amplification or changes in brightness. An enhanced variant of the traditional histogram equalization method, CLAHE, was developed to overcome these constraints [3].

The CLAHE algorithm operates by partitioning the image into discrete, contextual areas and applying Histogram Equalization to each area separately. The utilization of this local technique serves to alleviate the excessive magnification of noise, which may arise from global Histogram Equalization. Simultaneously, it offers optimal contrast enhancement by adapting its brightness values according to any specific attributes of every local region [4].

3.5. Evaluation Metrics

Evaluation metrics are used for classification cases in machine learning. This evaluation has various formulations. In this study, 5 types of evaluation metrics were used with the following formulation details.

$$Accuracy = \frac{(TP+TN)}{(P+N)}$$

$$Recall = \frac{(TN)}{(FP+TN)}$$

$$Precision = \frac{(TP)}{(TP+FP)}$$

$$F1 = \frac{2TP}{(2TP+FP+FN)}$$

where TP = true positive, TN = true negative, P = positive, N = negative, FP = false positive, TN = true negative.

Confusion matrix test: this method can detect the correct class predicted correctly [TP] and the wrong class predicted incorrectly [TN] in each class. Details on each class prediction will make it easier to analyze the causes of detection errors in each class.

AUC test: AUC (Area Under the Curve) is a method to evaluate the accuracy level of classification performance. The range of values between (0, 0) and (1, 1), where values near 1 represent that the model performs excellently. A result below 0.5 indicates that the model performed poorly. The model with a result below 0.5 is not suitable for adoption.

4. Results and Discussion

In this study, we divided our experiment into three scenarios. In the first scenario, our model only adopted the InceptionResNet method with Normal pre-processing. In the second experiment, we considered implementing CLAHE for contrast enhancement. We observed the impact of CLAHE adoption on reducing the misclassification task. In the last scenario, we observed the impact of CLAHE adoption and the Transformer mechanism on the InceptionResNet classification task. The details of the experiment report are shown in the following subsections.

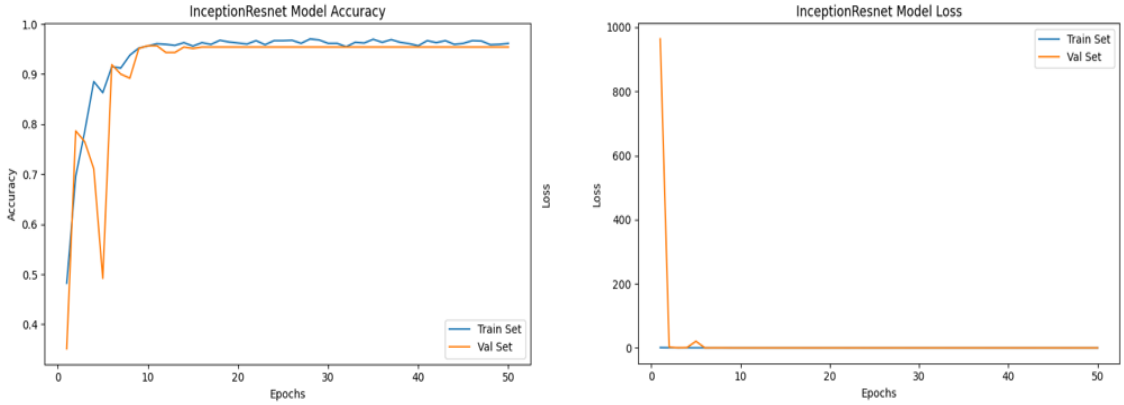
4.1. Result of the original InceptionResNet test

Our experiment considers implementing several hyperparameter scenarios, including 50 epochs, 32 batch sizes, and an optimizer using Adam, with a minimum learning rate of 0.000000000001. The result of the experiment with evaluation metrics is shown in Table 2. The performance of InceptionResNet achieved an accuracy of 0.94 for the three classes of lung diseases. While the precision, recall, and F1 scores achieved lower performance for the Normal class. The Covid-19 class achieved the highest result over the Normal and Virus classes. We considered calculating an accuracy test on average. This is needed to compare the results with previous research written in the final session of this article.

Table 2. Evaluation metric of InceptionResNet

Classes	Accuracy	Precision	Recall	F1
Normal	0.94	0.92	0.93	0.92
Covid		0.96	0.97	0.97
Virus		0.94	0.92	0.93

Figs. 5a and 5b illustrate the process by which the algorithm learns from data. The process is running normally. There is no issue with overfitting or underfitting. The training process achieved convergence after 50 epochs. The accuracy and validation tests yielded nearly identical values. That means the algorithm has successfully learn the data training. The loss test demonstrated excellent performance, where the value reached almost 0. There is no difference between the lost and validation tests.

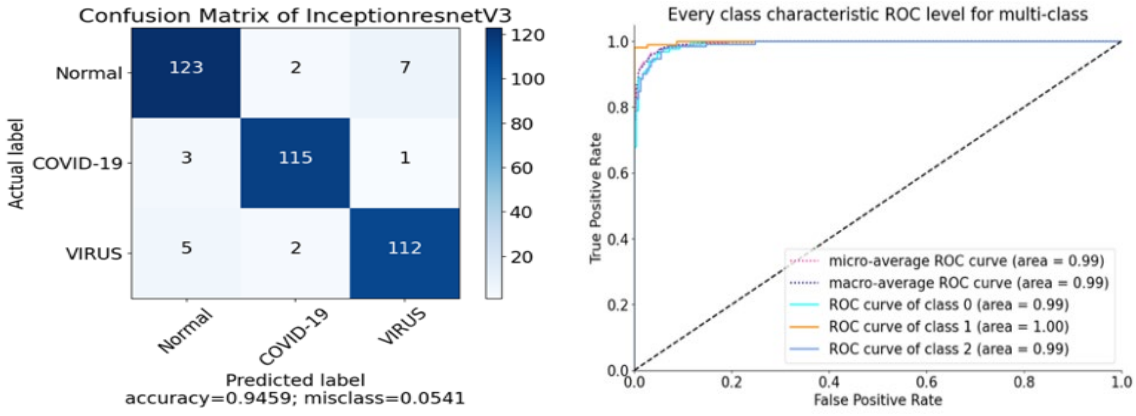


(a) (b)

Fig. 5. Accuracy and Loss test for InceptionResNet

The confusion matrix test is shown in Fig. 6a. The confusion matrix demonstrated that the LUNGINFORMER algorithm achieved 0.9459 in accuracy, and the misclassification was 0.0541. The majority of errors detected by the Normal class are classified as Virus, and the majority of errors detected by the Virus class are classified as Normal. The Virus class and Normal class indicated that they were facing biased data. There is a similar pattern between the Normal and Virus datasets.

The AUC test is illustrated in Fig. 6b. The AUC results demonstrate that every class achieved significant performance in classifying lung diseases. The class Normal with blue light color achieved 0.99, the Covid-19 class with orange color achieved 1.00, and the Virus class with dark blue color achieved 0.99. The overall class of the AUC test shows an accuracy of almost 1.00. The model is suitable for classification model in medical application

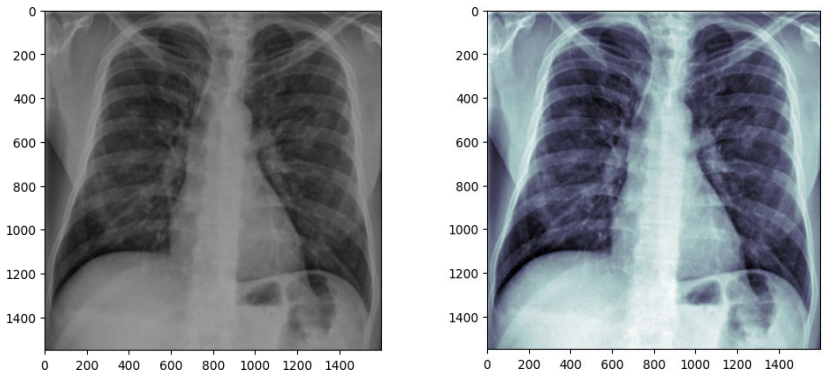


(a) (b)

Fig. 6. Confusion matrix and AUC test of Inceptionresnet

4.2. Impact of CLAHE adoption on Inceptionresnet performance

The second experiment adopted the CLAHE technique to improve the standard quality of the CXR image. We used the CV2 library to implement CLAHE for the whole dataset. The example or original X-ray datasets are shown in Fig. 7a. The CXR image appears blurry and has low contrast.



(a) (b)

Fig. 7. XCR image transformation after contrast enhancement with CLAHE

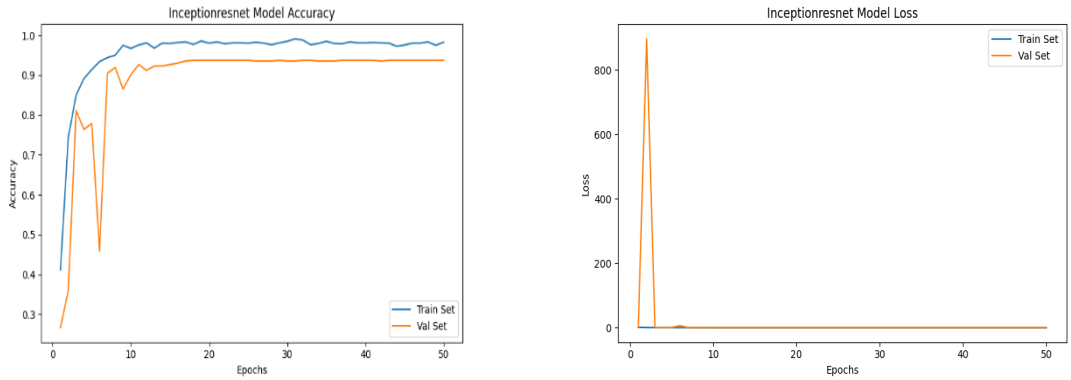
The transformation of the original image into an enhanced image with contrast using CLAHE is shown in Fig. 7b. The CXR with CLAHE demonstrates an image with high contrast and white balance. The contrast enhancement is expected to improve the performance of machine learning in classifying lung diseases.

The experimental results of LUNGINFORMER using CLAHE are shown in Table 3. The issue arises in the Normal class, where the precision value is 0.93. The performance increased about 1% over traditional InceptionResNet without CLAHE. Better precision was achieved for the COVID-19 and Virus classes. On the other hand, the Normal class indicated facing mis-detection. There are several challenges ahead for the Normal class.

Table 3. Evaluation of InceptionResNet with CLAHE

Classes	Accuracy	Precision	Recall	F1
Normal	0.96	0.93	0.96	0.95
Covid		0.99	0.98	0.98
Virus		0.96	0.94	0.95

Figs. 8a and 8b represent the accuracy test and loss test. The accuracy test show the model achieved convergence in 50 epoch. There is no over fitting and under fitting issue on the training model. While, in the loss test show that the loss value achieved good perform. It was indicated the loss and validation test almost reach in 0.

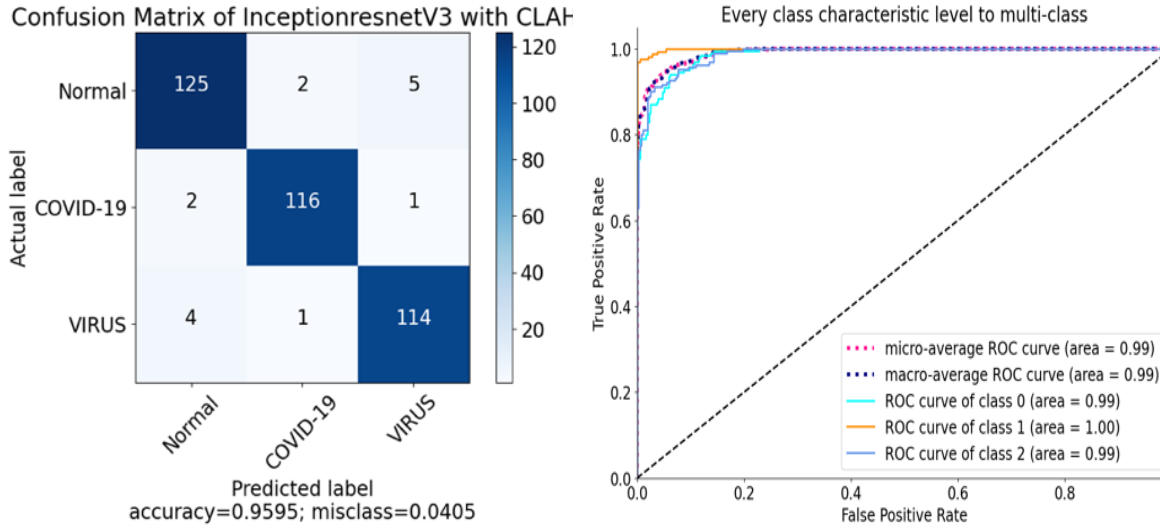


(a) (b)

Fig. 8. Accuracy and Loss test of Inceptionresnet with CLAHE

The adoption of CLAHE achieved good performance in the confusion matrix test, where the true detection achieved 0.9595 and the false detection was 0.04. The adoption of CLAHE increased the performance of the confusion matrix by 1.3%. The adoption of CLAHE reduces misclassification in the Normal class and the virus class. In comparison, the accurate detection for the Normal and virus classes significantly increases performance.

The AUC test demonstrated that the Normal class achieved a value of 0.99 (light blue color), the Covid-19 class achieved a value of 1.00 (orange color), and the virus class achieved a value of 0.99. On the other hand, there is some classification error in the Normal and Virus classes. At the same time, there is no mistake detection for the COVID-19 class. The detailed AUC test is shown in Fig. 9. According to the AUC test, the algorithm achieved the best performance in the classification task. It was indicated that the AUC test reaches almost 1.00.



(a) (b)

Fig. 9. Confusion matrix of InceptionResNet with CLAHE

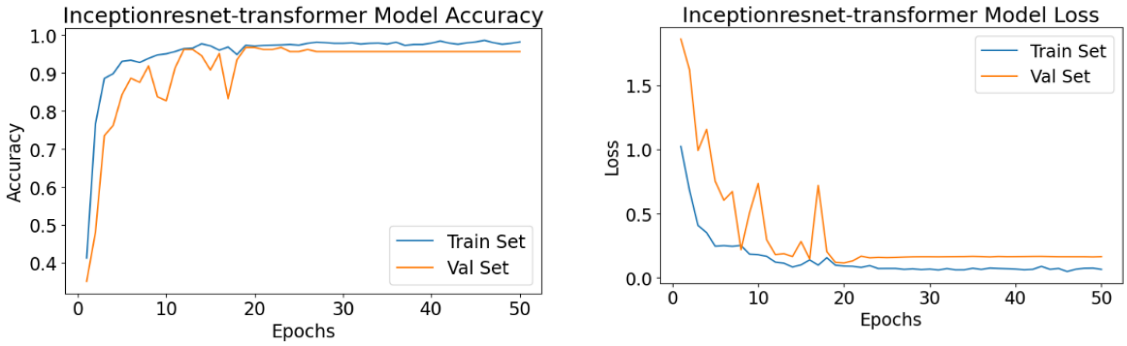
4.3. Impact of CLAHE and Transformer for InceptionResNet Performance

The third experiment investigated the use of CLAHE and Transformer for the feature selection task in the InceptionResNet model. Following the experiment report in Table 4, LUNGINFORMER achieved an average accuracy of 0.98. The misdetection for normal Dan virus class success can be reduced. Indeed, the normal class still found misclassification detection. It can be seen that the precision test achieved 0.96. The precision result indicated an error dominant in the Normal class.

Table 4. Evaluation metric of InceptionResNet and Transformer with CLAHE

Classes	Accuracy	Precision	Recall	F1
Normal	0.98	0.96	0.97	0.97
Covid		0.99	0.99	0.99
Virus		0.97	0.98	0.98

The adoption of the Transformer for InceptionResNet exhibits a distinct pattern compared to the original InceptionResNet in the previous experiment. The adoption of the Transformer required a lot of time to achieve convergence. The accuracy and loss tests are shown in Figs. 10a and 10b, respectively. There is no issue about over-fitting and under-fitting for LUNGINFORMER, where the training process requires time consumption due to involving two processes, including InceptionResNet and the Transformer algorithm.

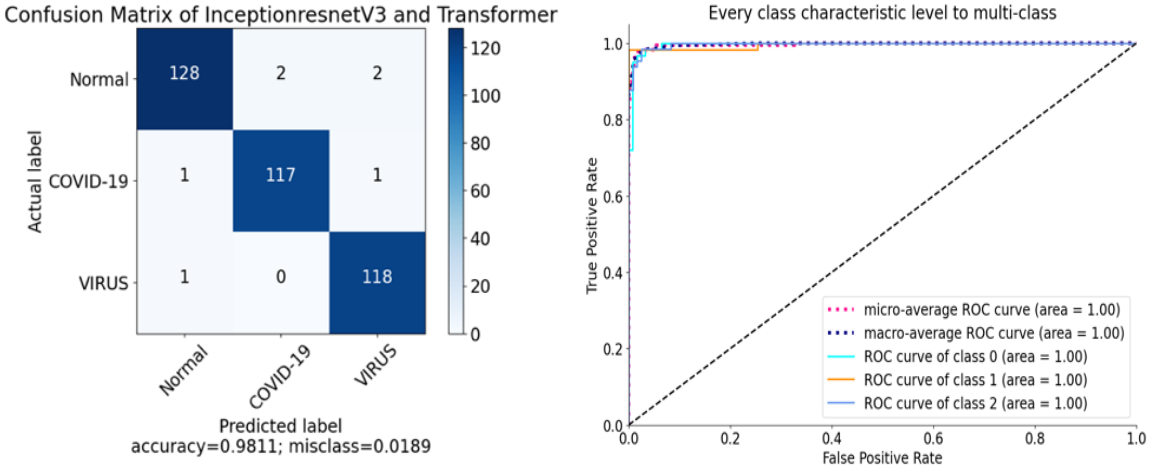


(a) (b)

Fig. 10. Accuracy and loss test of InceptionResNet and Transformer with CLAHE

The confusion matrix test demonstrated that the adoption of the Transformer significantly influenced the effectiveness of lung disease detection, with a reduction in misclassification of 1.89% and an increase in correct classification to 98.1%. The dominant error detection between the Normal class and the virus class can be solved. The details of the confusion matrix result for LUNGINFORMER are shown in Fig. 11a.

The AUC test for LUNGINFORMER is shown in Fig. 11b. The result of the AUC test shows that the LUNGINFORMER model achieved an excellent score with 1.00 for every class. The sign of Normal (light blue), Covid-19 (orange), Virus (dark blue) indicated perfect line.



(a) (b)

Fig. 11. Confusion matrix of InceptionResNet and Transformer with CLAHE

4.4. Comparison with state-of-the-art

Lung disease detection faces several issues. Several strategies have been made to handle inaccurate detection. In this study, our proposed model outperforms previous work. The complete previous work is shown in Table 5. Several models tried to enhance the performance with ensemble learning, such as [5], [6]. Their ensemble model, which utilized VGG, RestNet, and DenseNet, achieved an accuracy of 0.98. The model proves to be the most effective compared to other deep learning models. The shortcoming of ensemble learning is high computation cost. The hybrid deep learning model is a major strategy for enhancing the classification task. Another proposed model, such as [7], utilizes a hybrid approach combining VGG and ResNet, achieving a 0.98 accuracy. The model considers implementing a serial model. A hybrid model [8] using Restnet101, Resnet152, Inception, and InceptionResNet, and a hybrid model with VGG16, Resnet, and DenseNet. Both models achieved good performance with 0.961

and 0.968. In the term of binary classification point of view, they consider to applied binary class. The binary class easier to handle over multi-class classification.

Table 5. Comparison results than previous work

Ref.	Algorithm	Class Datasets	Accuracy
[7]	VGG+Resnet	121 Covid-19, 122 Normal	0.98
[5]	Ensemble model with VGG, RestNet, DenseNet	2373 Covid-19, 2890 Pneumonia, 3193 TBC, 3038 Normal	0.98
[8]	Restnet101, Resnet152, Inception, Inceptionresnet	Bacteria, Covid-19,Normal, Virus	0.961
[9]	Densenet121	99 Covid-19, 88 Pneumonia	0.983
[10]	VGG16, Resnet, DenseNet	3883 Pneumonia, 1349 Normal	0.968
[11]	CNN, Adaboost	100 Normal, 101 Abnormal	0.908
[12]	RNN, LSTM	1341 Normal, 2890 Pneumonia, 1483 Viral	0.954
[6]	(ensemble)Inception, Resnet, MobileNet	1493 Viral, 2780 Bacterial, 1583 Normal, 231 Covid-19	0.948
[13]	CNN, LDA, MGSA	Normal, Benign, Malignant	0.945
[14]	Modified Xception	229 Covid-19, 1079 Normal, 3106 Pneumonia	0.910
[15]	Hybrid Retina net, Region-CNN	26658 (4 classes)	0.758
Proposed Model	Inceptionresnet and Transformer, CLAHE	See Table 1	0.980

The majority of studies considered adopting a deep learning platform such as RestNet [5]–[8], [10] and VGG. Their model has become a popular option for developing a hybrid model [5], [7], [10]. Both deep learning platforms achieved tremendous results compared to other models such as Inception, Xception, and DenseNet [8], [9], [14]. The existing deep learning model consists of a complicated network. Millions of networks influence time consumption and computation cost. Some of them consider applying a traditional deep learning model, such as CNN, LSTM, or CNN [11], [13], [15]. Unfortunately, the adoption of generic deep learning achieved low performance in accuracy. The benefit of a traditional Deep learning model is that the computation cost is very cheap compared to a complex deep learning model. A traditional deep learning model is a suitable option for low-computation devices.

Similar to LUNGINFORMER, where they applied an optimization schema using Adaboost [11]. The Adaboost mechanism is considered for implementing ensemble learning for the feature selection task. The model achieved a low performance of 0.90. Indeed, the low performance was caused by several factors, including biased data of images and the quality of medical images. Different from previous work with traditional CNN, A CNN model with data optimization using LDA and MGSA succeeded in achieving better performance in accuracy with 0.945 [13]. The data optimization plays an important role in increasing the performance of classification. Another classical deep learning model proposed by researchers for multi-class lung disease [12]. Their work achieved an accuracy of 0.954. Contrast with LUNGINFORMER, which adopted CLAHE, their model implements image improvement using ROI (Region of Interest), aiming to enhance the quality of the image. The results show that the adoption of ROI success increases RNN and LSTM.

Some of the previous work involves an ensemble learning mechanism [5]. The scenario to combine some deep learning models in ensemble learning proved better performance over generic deep learning. Unfortunately, ensemble learning requires a computational cost for the algorithm to train lung image datasets. The application of ensemble learning for rich regions or countries is not a problem because they have sufficient budgets, but this will be a serious issue for regions or areas with poor infrastructure, as they often lack the necessary funds. Even if they have a budget, it can be challenging to reach because an internet connection for cloud-based computing needs does not support it. For areas with minimal budget or infrastructure for computing, the use of traditional machine learning methods, such as decision trees, SVM, and KNN, tends to make sense.

Finally, the comparison results show that the adoption of enhanced image quality and the hybridization of deep learning have influenced the effectiveness of the classification task. Similar to LUNGINFORMER, several image enhancement and hybrid deep learning models proved better performance over traditional deep learning models. Indeed, the comparison scenario in this study is not fair enough due to some differences. For instance, the comparison scenario faces different numbers of datasets, classifications, image quality, and hyperparameter schemas. However, it cannot be denied that the use of transformers has a significant impact on producing classifications of lung diseases. This is proven by the experiments conducted in this study.

4.5. Promising and challenging for medical application

In this study, we have implemented a deep learning algorithm. We combined InceptionResNet and a transformer. In the recent 5 years, deep learning has emerged as a powerful tool in the detection and diagnosis of lung diseases, offering valuable support to practical medical analysis, by leveraging complex neural networks, especially CNN, as a basic mechanism to create advanced deep learning. Our proposed model, called LUNGIFORMER, can automatically learn patterns from large volumes of medical imaging data such as chest X-rays and CT scans. LUNGINFORMER is capable of detecting lung abnormalities, including pneumonia, tuberculosis, chronic obstructive pulmonary disease (COPD), lung nodules, and even early signs of lung cancer with high accuracy.

From a clinical practice point of view, LUNGINFORMER can enhance radiological analysis by acting as an intelligent assistant. It can rapidly scan and interpret images, highlight suspicious areas, and provide probability scores for different diseases, allowing radiologists to make more informed decisions. For instance, early-stage lung cancer is often difficult to detect due to subtle imaging features; deep learning models trained on thousands of annotated images can recognize such patterns, supporting earlier intervention and potentially improving patient outcomes.

Moreover, LUNGINFORMER helps address variability and fatigue in human interpretation. Automated systems can process large volumes of images consistently, reducing oversight and improving efficiency. This is particularly beneficial in high-demand settings or regions with limited access to trained specialists.

In addition to image analysis, LUNGINFORMER can integrate data from electronic health records (EHRs), lab tests, and clinical notes to provide a more comprehensive view of a patient's condition. This supports differential diagnosis and personalized treatment planning. Despite challenges such as the need for large, annotated datasets, regulatory approval, and concerns over interpretability, deep learning continues to evolve rapidly. With ongoing research and improved model transparency, its integration into medical imaging workflows holds great promise for enhancing diagnostic accuracy, reducing workload, and ultimately improving patient care in the detection and management of lung diseases.

5. Conclusion

Automatic detection for lung diseases with CXR images faces a serious problem. The problem arises due to some aspects, including poor quality of XCR image and biased image intersection classes. This problem caused the detection result to be inaccurate. In this study, we tried to handle the problem with 2 essential approaches with the adoption of CLAHE and a hybrid deep learning technique. Our approach aims to enhance the performance of the classification task. CLAHE is responsible for improving the quality of the image, while the hybrid of Inception-ResNet and Transformer is responsible for enhancing classification. LUNGIFORMER achieved better performance over previous work for lung pneumonia detection. Our datasets showed biased image characteristics between viral pneumonia and COVID-19. There is a similarity between pneumonia and COVID-19 image patterns. Adoption of CLAHE and Transformer improved the inaccurate result of the traditional deep learning approach. In the future study, the scenario hybridization with various deep learning platforms with the Transformer algorithm for feature selection will yield promising results. Several scenarios with various multi-class datasets will be challenges for the future study.

Acknowledgment

I'm glad for the finalization of our research. Thanks to Universitas Amikom Yogyakarta for supporting this research.

Declarations

Author contribution. All authors contributed equally to the main contributor to this paper. All authors read and approved the final paper.
Funding statement. Funding statement. Universitas Amikom Yogyakarta supported the study with grant number RP-1730998448.
Conflict of interest. The authors declare no conflict of interest.
Additional information. No additional information is available for this paper.

Data and Software Availability Statements

The complete dataset in this experiment can be accessed with the URL as follows: https://www.kaggle.com/datasets/sid321axn/covid-cxr-image-dataset-research

References

[1] F. Shi *et al.*, "Review of Artificial Intelligence Techniques in Imaging Data Acquisition, Segmentation, and Diagnosis for COVID-19," *IEEE Rev. Biomed. Eng.*, vol. 14, pp. 4–15, 2021, doi: 10.1109/RBME.2020.2987975.

[2] R. Vaishya, M. Javaid, I. H. Khan, and A. Haleem, "Artificial Intelligence (AI) applications for COVID-19 pandemic," *Diabetes Metab. Syndr. Clin. Res. Rev.*, vol. 14, no. 4, pp. 337–339, Jul. 2020, doi: 10.1016/j.dsx.2020.04.012.

[3] "COVID-19 cases," *WHO*, 2025. [Online]. Available at: https://data.who.int/dashboards/covid19/cases?n=c.

[4] A. Alimadadi, S. Aryal, I. Manandhar, P. B. Munroe, B. Joe, and X. Cheng, "Artificial intelligence and machine learning to fight COVID-19," *Physiol. Genomics*, vol. 52, no. 4, pp. 200–202, Apr. 2020, doi: 10.1152/physiolgenomics.00029.2020.

[5] M. Z. Alom *et al.*, "The History Began from AlexNet: A Comprehensive Survey on Deep Learning Approaches," *arXiv*, pp. 1–39, Mar. 2018. [Online]. Available at: https://arxiv.org/abs/1803.01164.

[6] L. Deng and D. Yu, "Deep Learning: Methods and Applications," *Found. Trends® Signal Process.*, vol. 7, no. 3–4, pp. 197–387, 2014, doi: 10.1561/2000000039.

[7] A. Sunyoto and Hanafi, "Enhance Intrusion Detection (IDS) System Using Deep SDAE to Increase Effectiveness of Dimensional Reduction in Machine Learning and Deep Learning," *Int. J. Intell. Eng. Syst.*, vol. 15, no. 4, p. 2022, Aug. 2022, doi: 10.22266/ijies2022.0831.13.

[8] H. Hanafi, A. Pranolo, Y. Mao, T. Hariguna, L. Hernandez, and N. F. Kurniawan, "IDSX-Attention: Intrusion detection system (IDS) based hybrid MADE-SDAE and LSTM-Attention mechanism," *Int. J. Adv. Intell. Informatics*, vol. 9, no. 1, p. 121, Mar. 2023, doi: 10.26555/ijain.v9i1.942.

[9] Hanafi, R. Widyawati, and A. S. Widowati, "Effect of service quality and online servicescape toward customer satisfaction and loyalty mediated by perceived value," *IOP Conf. Ser. Earth Environ. Sci.*, vol. 704, no. 1, p. 012011, Mar. 2021, doi: 10.1088/1755-1315/704/1/012011.

[10] Hanafi, N. Suryana, and A. S. H. Basari, "Deep Contextual of Document Using Deep LSTM Meet Matrix Factorization to Handle Sparse Data: Proposed Model," *J. Phys. Conf. Ser.*, vol. 1577, no. 1, p. 012002, Jul. 2020, doi: 10.1088/1742-6596/1577/1/012002.

[11] Hanafi *et al.*, "Improvement of E-commerce Recommender System Using Hybridization of Bert, Matrix Factorization and Attention Mechanism," *Int. J. Intell. Eng. Syst.*, vol. 17, no. 5, pp. 725–740, Oct. 2024, doi: 10.22266/ijies2024.1031.55.

[12] Hanafi, N. Suryana, and A. S. H. Basari, "Generate Contextual Insight of Product Review Using Deep LSTM and Word Embedding," *J. Phys. Conf. Ser.*, vol. 1577, no. 1, p. 012006, Jul. 2020, doi: 10.1088/1742-6596/1577/1/012006.

[13] Hanafi, N. Suryana, and A. S. B. H. Basari, "Convolutional-NN and Word Embedding for Making an Effective Product Recommendation Based on Enhanced Contextual Understanding of a Product Review," *Int. J. Adv. Sci. Eng. Inf. Technol.*, vol. 9, no. 3, pp. 1063–1070, Jun. 2019, doi: 10.18517/IJASEIT.9.3.8843.

[14] Hanafi and B. Mohd Aboobaider, "Word Sequential Using Deep LSTM and Matrix Factorization to Handle Rating Sparse Data for E-Commerce Recommender System," *Comput. Intell. Neurosci.*, vol. 2021, no. 1, p. 8751173, Jan. 2021, doi: 10.1155/2021/8751173.

[15] H. Hanafi, N. Suryana, and A. S. H. Basari, "Dynamic convolutional neural network for eliminating item sparse data on recommender system," *Int. J. Adv. Intell. Informatics*, vol. 4, no. 3, p. 226, Nov. 2018, doi: 10.26555/ijain.v4i3.291.

[16] Hanafi, A. Pranolo, and Y. Mao, "Cae-covidx: Automatic covid-19 disease detection based on x-ray images using enhanced deep convolutional and autoencoder," *Int. J. Adv. Intell. Informatics*, vol. 7, no. 1, pp. 49–62, 2021, doi: 10.26555/ijain.v7i1.577.

[17] M. Hanafi, "AMIKOMNET: Novel Structure for a Deep Learning Model to Enhance COVID-19 Classification Task Performance," *Big Data Cogn. Comput.*, vol. 8, no. 7, p. 77, Jul. 2024, doi: 10.3390/bdcc8070077.

[18] B. Mostafa, M. Sakr, and A. Keshk, "Employing the Capabilities of LSTM and Bi-LSTM for Lung Cancer Detection and Classification," *Int. J. Intell. Eng. Syst.*, vol. 17, no. 5, pp. 412–423, Oct. 2024, doi: 10.22266/ijies2024.1031.32.

[19] D. N. Vinod, N. Kapileswar, J. Simon, P. P. Kumar, and M. Saraswathy, "Prognosis of COVID-19 Using Ultrasound Scans Augmented by Generative Adversarial Networks," *Int. J. Intell. Eng. Syst.*, vol. 17, no. 4, pp. 598–610, Aug. 2024, doi: 10.22266/ijies2024.0831.45.

[20] K. Srinivas, R. Gagana Sri, K. Pravallika, K. Nishitha, and S. R. Polamuri, "COVID-19 prediction based on hybrid Inception V3 with VGG16 using chest X-ray images," *Multimed. Tools Appl.*, vol. 83, no. 12, pp. 36665–36682, Jun. 2023, doi: 10.1007/s11042-023-15903-y.

[21] D. Singh, V. Kumar, and M. Kaur, "Densely connected convolutional networks-based COVID-19 screening model," *Appl. Intell.*, vol. 51, no. 5, pp. 3044–3051, May 2021, doi: 10.1007/s10489-020-02149-6.

[22] A. Jaiswal, N. Gianchandani, D. Singh, V. Kumar, and M. Kaur, "Classification of the COVID-19 infected patients using DenseNet201 based deep transfer learning," *J. Biomol. Struct. Dyn.*, vol. 39, no. 15, pp. 5682–5689, Oct. 2021, doi: 10.1080/07391102.2020.1788642.

[23] A. K. Jaiswal, P. Tiwari, S. Kumar, D. Gupta, A. Khanna, and J. J. P. C. Rodrigues, "Identifying pneumonia in chest X-rays: A deep learning approach," *Measurement*, vol. 145, pp. 511–518, Oct. 2019, doi: 10.1016/j.measurement.2019.05.076.

[24] A. Narin, C. Kaya, and Z. Pamuk, "Automatic detection of coronavirus disease (COVID-19) using X-ray images and deep convolutional neural networks," *Pattern Anal. Appl.*, vol. 24, no. 3, pp. 1207–1220, Aug. 2021, doi: 10.1007/s10044-021-00984-y.

[25] D. Ezzat, A. E. Hassanien, and H. A. Ella, "An optimized deep learning architecture for the diagnosis of COVID-19 disease based on gravitational search optimization," *Appl. Soft Comput.*, vol. 98, p. 106742, Jan. 2021, doi: 10.1016/j.asoc.2020.106742.

[26] V. Ravi, H. Narasimhan, C. Chakraborty, and T. D. Pham, "Deep learning-based meta-classifier approach for COVID-19 classification using CT scan and chest X-ray images," *Multimed. Syst.*, vol. 28, no. 4, pp. 1401–1415, Aug. 2022, doi: 10.1007/s00530-021-00826-1.

[27] J. Luján-García, C. Yáñez-Márquez, Y. Villuendas-Rey, and O. Camacho-Nieto, "A Transfer Learning Method for Pneumonia Classification and Visualization," *Appl. Sci.*, vol. 10, no. 8, p. 2908, Apr. 2020, doi: 10.3390/app10082908.

[28] N. Kalaivani, N. Manimaran, D. S. Sophia, and D. D Devi, "Deep Learning Based Lung Cancer Detection and Classification," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 994, no. 1, p. 012026, Dec. 2020, doi: 10.1088/1757-899X/994/1/012026.

[29] S. Goyal and R. Singh, "Detection and classification of lung diseases for pneumonia and Covid-19 using machine and deep learning techniques," *J. Ambient Intell. Humaniz. Comput.*, vol. 14, no. 4, pp. 3239–3259, Apr. 2023, doi: 10.1007/s12652-021-03464-7.

[30] K. El Asnaoui, "Design ensemble deep learning model for pneumonia disease classification," *Int. J. Multimed. Inf. Retr.*, vol. 10, no. 1, pp. 55–68, Mar. 2021, doi: 10.1007/s13735-021-00204-7.

[31] L. S.K., S. N. Mohanty, S. K., A. N., and G. Ramirez, "Optimal deep learning model for classification of lung cancer on CT images," *Futur. Gener. Comput. Syst.*, vol. 92, pp. 374–382, Mar. 2019, doi: 10.1016/j.future.2018.10.009.

[32] J. E. Luján-García, M. A. Moreno-Ibarra, Y. Villuendas-Rey, and C. Yáñez-Márquez, "Fast COVID-19 and Pneumonia Classification Using Chest X-ray Images," *Mathematics*, vol. 8, no. 9, p. 1423, Aug. 2020, doi: 10.3390/math8091423.

[33] I. Sirazitdinov, M. Kholiavchenko, T. Mustafaev, Y. Yixuan, R. Kuleev, and B. Ibragimov, "Deep neural network ensemble for pneumonia localization from a large-scale chest x-ray database," *Comput. Electr. Eng.*, vol. 78, pp. 388–399, Sep. 2019, doi: 10.1016/j.compeleceng.2019.08.004.

[34] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. A. Alemi, "Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning," *Proc. AAAI Conf. Artif. Intell.*, vol. 31, no. 1, pp. 4278–4284, Feb. 2017, doi: 10.1609/AAAI.V31I1.11231.