TIERS Information Technology Journal
Vol.4, No.1, June 2023, pp. 73~82
ISSN: 2723-4533 / E-ISSN: 2723-4541
DOI: 10.38043/tiers.v4i1.4380

Pneumonia Classification Utilizing VGG-16 Architecture and Convolutional Neural Network Algorithm for Imbalanced Datasets

Mohammad Idhom[1], Dwi Arman Prasetya[2], Prismahardi Aji Riyantoko[3], Tresna Maulana Fahrudin[4], Anggraini Puspita Sari[5]
idhom@upnjatim.ac.id[1], arman.prasetya.sada@upnjatim.ac.id[2], prismahardi.aji.ds@upnjatim.ac.id[3], tresna.maulana.ds@upnjatim.ac.id[4], anggraini.puspita.if@upnjatim.ac.id[5]
[1,2,3,4,5] Faculty of Computer Science, Universitas Pembangunan Nasional Veteran Jawa Timur, East Java

ABSTRACT

This research focuses on accurately classifying pneumonia in children under the age of 5 using X-ray images, considering the challenge of an imbalanced dataset. A modified VGG-16 CNN architecture is evaluated for pneumonia classification in Chest X-Ray Images. The study compares testing results with and without data augmentation techniques and explores the potential application of the model in an Android-based machine learning system for pneumonia diagnosis assistance. Using a dataset of 5,856 Chest X-Ray images categorized as normal or pneumonia, obtained from Kaggle, the research conducts two test scenarios: one without data augmentation and another with data augmentation techniques. The modified VGG-16 CNN algorithm's performance is evaluated using the accuracy metric. The results highlight the effectiveness of data augmentation in improving pneumonia classification accuracy. The augmented tests outperform the non-augmented ones, achieving an impressive 92% accuracy, indicating a significant 15% improvement over the non-augmented scenario. This improvement underscores the efficacy of data augmentation techniques in enhancing the CNN's ability to accurately classify pneumonia, particularly when faced with an imbalanced dataset. Furthermore, the research explores the potential integration of the trained model into an Android-based machine learning system for pneumonia diagnosis assistance. This integration would enable doctors to analyze X-ray images and identify potential pneumonia cases in patients. The integration of advanced machine learning systems in healthcare holds promise for improving patient care and the accuracy of pneumonia diagnoses. In summary, this research contributes to the accurate classification of pneumonia in children under 5 years old using X-ray images. It emphasizes the efficacy of data augmentation techniques in enhancing classification accuracy and explores the practical application of an Android-based machine learning system for pneumonia diagnosis assistance. These findings underscore the importance of advanced machine learning systems in healthcare and their potential to improve pneumonia diagnosis accuracy and enhance patient care.

Keywords: Machine Learning; Classification; CNN; Pneumonia; VGG-16.

Article Info

Accepted	:	18-05-2023
Revised	:	04-03-2023
Published Online	:	25-06-2023

Correspondence Author:
Mohammad Idhom,
Faculty of Computer Science,
Universitas Pembangunan Nasional "Veteran" Jawa Timur,
Rungkut Madya Street – Gununganyar Surabaya, East Java, 60294, Indonesia
Email: idhom@upnjatim.ac.id

1. INTRODUCTION

Over the past few decades, there has been a profound transformation in human work and lifestyle due to Artificial Intelligence (AI) [1][11]. AI involves the creation of computer systems capable of emulating human intelligence and behavior [10]. Machine Learning (ML), a specialized field of AI, is being actively explored in several industries, with a particular emphasis on healthcare [2]. In the healthcare sector, ML finds application in diverse areas such as predicting diseases, multi-disease prediction, assisting in surgical procedures, personalized medicine, and detecting medical images [3].

Pneumonia, also known as "wet lung," is a prevalent type of lung disease globally [4][5]. It refers to the inflammation of the alveoli, which can occur in either one or both lungs. The condition is characterized by the accumulation of fluid or pus in the affected alveoli, leading to symptoms such as coughing up phlegm, fever, and respiratory difficulties [5]. The primary causes of pneumonia are bacterial, viral, or fungal infections [6]. Approximately 450 million individuals worldwide contract pneumonia annually [7].

While pneumonia can affect individuals of all ages, certain groups are particularly vulnerable, such as children under the age of 5 and adults over the age of 65, especially those with pre-existing lung conditions [8]. Pneumonia, which is classified as an infectious disease, is the primary cause of death among children on a global scale [9]. In 2019, the World Health Organization (WHO) reported that pneumonia caused around 14% of deaths among children between the ages of 0 and 5, leading to an estimated 740,180 fatalities.

Pneumonia has emerged as a significant issue that demands effective management. Therefore, early detection of pneumonia is crucial. One method used for pneumonia detection is through X-ray imaging. X-ray images play a critical role in aiding doctors in diagnosing various diseases [12]. However, the manual diagnosis of X-ray images by physicians can be time-consuming [13], necessitating the need for computer-assisted automatic diagnostics to assist doctors in analyzing numerous X-ray images of patients potentially afflicted with pneumonia.

The Convolutional Neural Network (CNN) is a remarkably efficient algorithm specifically designed for handling image data within Neural Networks (NN) [14]. The CNN architecture is specifically crafted to capture significant features from images, making it ideal for tasks like image classification, object detection, and image recognition. CNNs employ convolutional layers to apply filters and identify patterns in images, followed by pooling layers to decrease spatial dimensions, and fully connected layers for tasks like classification or regression [15]. CNNs have exhibited exceptional performance across a wide range of computer vision tasks, establishing them as the preferred algorithm for image-related applications.

In the study conducted by Sharma et al. [5], the primary emphasis of the study was on utilizing Convolutional Neural Networks (CNNs) for feature extraction and classification of pneumonia cases. The researchers utilized two distinct CNN architectures: one incorporating a dropout layer and the other lacking a dropout layer. The testing process involved training both CNN architectures using both the original dataset and an augmented dataset. The findings revealed that the most favorable performance was observed in the testing scenario that employed the dropout layer and data augmentation, achieving an accuracy rate of 90.68%. This suggests that the inclusion of dropout layers and data augmentation techniques contributed to improved classification accuracy in pneumonia detection using CNNs.

In a similar study conducted by Maysanjaya [12], a CNN architecture have two convolutional layers were utilized, each having a filter size of 32. After the convolutional layers, the model included two max-pooling layers and three dense layers. The Rectified Linear Unit (ReLU) activation function was utilized throughout the network, and the Adam optimization algorithm was employed. The testing phase was conducted from the 25th to the 200th epoch, with observations taken at intervals of every multiple of 25 epochs. The study reported an average accuracy value of 89.58% for the classification task. These findings suggest that the CNN architecture employed by Maysanjaya demonstrated good performance in pneumonia detection, achieving a high level of accuracy in the classification task.

In their research, Nugroho et al. [16] conducted an experiment on pneumonia classification using two algorithms: CNN (Convolutional Neural Network) and CNN-ELM (Convolutional Neural Network-Extreme Learning Machine). The study explored different input image sizes to assess their impact on algorithm performance. The results revealed that both algorithms achieved their highest performance when the input image size was set at 200x200 pixels. For the CNN algorithm, the accuracy reached 84.78%, while the CNN-

.

ELM algorithm achieved an accuracy of 93.59%. The findings indicate that the CNN-ELM algorithm demonstrated superior performance compared to the CNN algorithm in pneumonia classification. Moreover, the results also underscore the significant impact of input image size on the classification performance.

This study focuses on the detection and classification of pneumonia using Convolutional Neural Networks (CNNs). Pneumonia is a prevalent lung disease with significant global health implica tions, especially for vulnerable populations like children and the elderly. Early detection is crucial for effective management and timely intervention. To assist healthcare professionals in diagnosis and treatment, the writer aims to investigate the performance of CNN algorithms in detecting and classifying pneumonia from X-ray images. This research seeks to provide insights that can improve the accuracy and efficiency of pneumonia diagnosis using CNN algorithms and medical imaging. The plan is to explore different CNN architectures, evaluate the impact of techniques such as dropout layers and data augmentation, and assess the influence of input image size on classification accuracy [17]. By conducting this research, the writer aims to contribute to the development of reliable and efficient automated systems for pneumonia detection, which can improve patient outcomes and assist healthcare providers in their decision-making process.

In the mentioned study, the authors employed the CNN algorithm with a modified VGG-16 architecture as the top layer. The dataset utilized was the Chest X-Ray Images (Pneumonia), which was characterized by an imbalanced distribution of data [18], [19]. The testing phase consisted of two scenarios: one without data augmentation and the other with data augmentation. This research focuses on investigating the impact of data augmentation on the performance of the CNN algorithm with a modified VGG-16 architecture for pneumonia classification in an imbalanced dataset. Data augmentation techniques are used to address the issue of imbalanced data, where the pneumonia class has fewer samples compared to the non-pneumonia class. By artificially generating additional training samples, data augmentation enhances the model's ability to learn and generalize from the imbalanced data. This approach improves the model's performance and its capability to handle imbalanced data by addressing the challenge of limited pneumonia samples through dataset augmentation.

2. RESEARCH METHOD

Fig. 1 illustrates the stages of the research conducted, which began with a literature review and the collection of necessary datasets. The core process initiated with data preprocessing, followed by data augmentation, the design of the CNN algorithm utilizing the VGG-16 architecture, model training, and model testing. Subsequently, model evaluation was performed to assess the obtained model's performance. If the results were unsatisfactory, the process returned to the data preprocessing stage. Once the core process was completed, the final step involved deploying the model with the best performance that had been achieved. This figure demonstrates the systematic workflow followed in the research, ensuring thoroughness in addressing the problem and optimizing the model's performance.

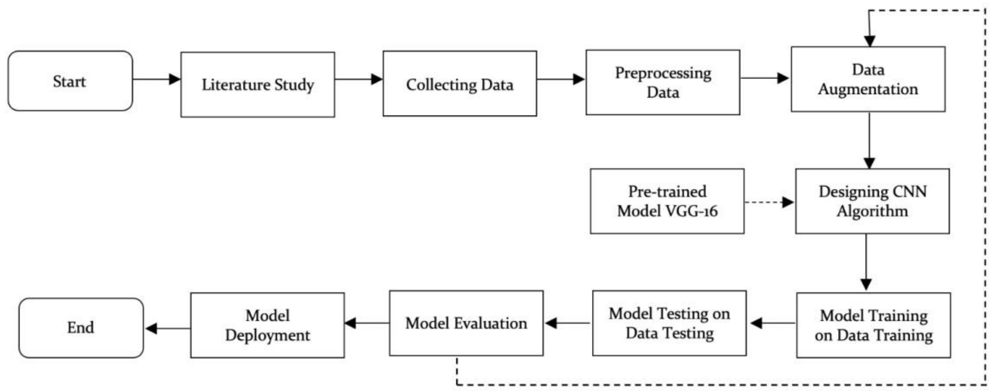

Fig. 1 Research Scenario Diagram

2.1. Literature Review

A literature study is a meticulous and comprehensive exploration of scientific literature, encompassing various sources and methodologies. Through this process, researchers gain valuable insights, develop a solid theoretical framework, and situate their own research within the broader academic landscape, ultimately contributing to the advancement of knowledge in their respective field.

2.2. Data Collecting

The researchers in this study employed the Chest X-Ray Images (Pneumonia) dataset sourced from Kaggle. This dataset comprises 5,856 images that were divided into three sets: training, validation, and testing. The dataset encompasses two categories of data, namely normal and pneumonia. This dataset should include images that are sufficiently diverse and cover many different cases of pneumonia. During the data collection process, several things were analyzed including data diversity, accurate data annotations, and image quality. The distribution of the datasets across the different sets is depicted in Table 1.

Table 1. The number of dataset distribution

Dataset	Normal	Pneumonia
Training	1.341	3.875
Validation	8	8
Testing	234	390
Total	**1.583**	**4.237**

2.3. Preprocessing Data

To ensure consistency in the dataset, the Chest X-Ray Images (Pneumonia) dataset contains images of varying sizes, with most images having dimensions above 1000×1000 pixels. As part of the data preprocessing stage, all the images are resized to a standardized size of 224×224 pixels. This resizing process allows for generalization and uniformity in the dataset, enabling further analysis and modeling tasks to be performed effectively [20].

2.4. Data Augmentation

Table 2 illustrates the different data augmentation techniques employed in this study to address the issue of an imbalanced dataset, as evident from Table 1. The uneven distribution of datasets in each class can potentially lead to overfitting of the model. To mitigate this problem, data augmentation techniques are utilized. These techniques involve applying transformations to the existing data to increase its diversity and balance the dataset [7]. The specific data augmentation techniques used in this study are presented in Table 2.

Table 2. Data Augmentation Technique

Augmentation Technique	Description	Value
Rescale	Rescale pixel values to fit within the range of 0 to 1.	1/255
Range of Rotation	Specify a random range within which to rotate the image.	20
Range of Zoom	Define the range of magnification for the image.	0.2
Range of Width Shift	Specify a broad range for shifting the image.	0.2
Range of Height Shift	Specify a large range for shifting the image.	0.2
Range of Shear	Randomly apply a shift transformation to the image.	0.2
Flip of Horizontal	Horizontally flip the image.	True
Mode of Fill	Fill the empty space in the image.	Nearest

2.5. CNN Algorithm Design

The CNN algorithm, a widely used neural network approach, excels at processing and extracting data features [21]. It is particularly suitable for tasks like computer vision and natural language processing. By employing convolutional layers, the algorithm effectively learns intricate patterns and spatial relationships within data. Its hierarchical feature extraction enables it to handle variations in scale and rotation [22]. Additionally, parameter sharing and weight tying optimize memory usage and computation. The algorithm's ability to learn spatial hierarchies of features enhances interpretability. Overall, the CNN algorithm's versatility, efficiency, and hierarchical learning capabilities have revolutionized data analysis and opened up new possibilities in various fields Image classification is extensively used to analyze images and identify objects accurately and efficiently. It has revolutionized fields like autonomous vehicles, medical imaging, and surveillance systems. In autonomous vehicles, it helps recognize pedestrians, traffic signs, and other vehicles for safe navigation [23].

Simonyan and Zisserman made significant progress in their research by introducing the VGG-16 architecture, which represents an enhancement over the previously developed AlexNet architecture. In 2014, their work was showcased at the ImageNet Large Scale Visual Recognition Competition (ILSRVC), where

VGG-16 achieved the second position, trailing behind GoogLeNet [14]. In this study opted to use the VGG-16 architecture but made some alterations to the top layer. They introduced a flatten layer after the max-pooling layer in block 5, followed by a fully connected layer with the ReLU activation function. The output layer employed the sigmoid activation function. Fig. 2 provides a visual representation of the VGG-16 architecture proposed in this study.

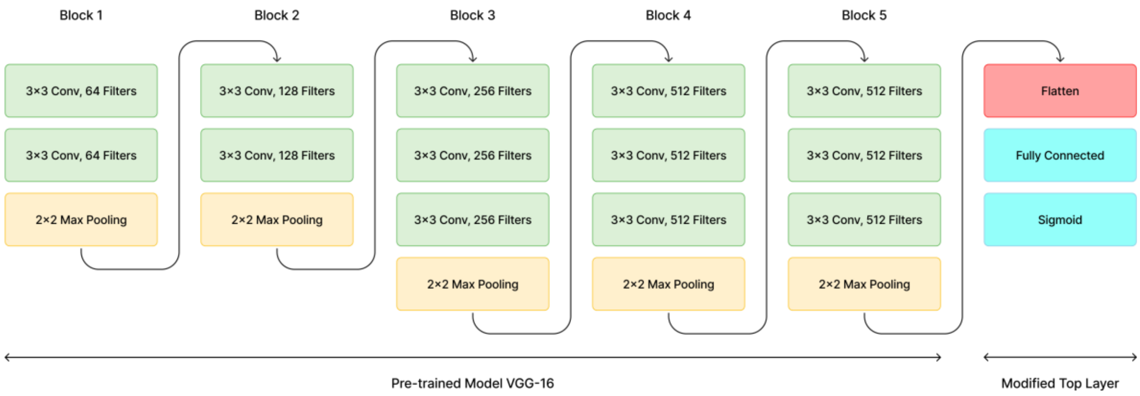

Fig. 2. VGG-16 Architecture with Top Layer Modification

2.6. Model Training

During this stage, the training process will commence using a set of 5,216 preprocessed and augmented images. These images have undergone the necessary preprocessing and augmentation steps. The modified VGG-16 architecture [24], as described in the previous stage, will be utilized for training. The model training will proceed iteratively, with each iteration referred to as an epoch. The training process will continue until a predetermined number of epochs is reached, which indicates the completion of the training stage.

2.7. Model Testing

Following the training stage, model testing was conducted using a separate set of 624 images from the test data. These images had already undergone preprocessing and were ready for evaluation. The model predicted values for each image were compared against the original label values present in the test data. By comparing these values, the number of correct predictions can be calculated. To assess the model's performance more comprehensively, a confusion matrix can be utilized. The confusion matrix offers a detailed summary of the model's predictions, enabling the calculation of metrics such as accuracy, precision, recall, and F1-score. By comparing the predicted and actual labels, the confusion matrix assists in evaluating the model's accuracy in correctly classifying images.

2.8. Model Evaluating

The confusion matrix is a tabular representation that summarizes the accuracy of classification results using evaluation metrics [5]. It provides a breakdown of the predicted and actual labels for a classification task. The key metrics defined within the confusion matrix and to evaluate the performance of the trained model, performance metric values can be calculated using the following equations:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{1}$$

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

$$F1 - Score = \frac{2 \times (Precision \times Recall)}{(Precision + Recall)} \tag{4}$$

2.9. Model Deployment

In the final stage, the top-performing model from the previous stage is deployed to establish a machine learning system, ensuring convenient access to its predictive capabilities through a user-friendly interface. This deployment process typically involves utilizing a server model, where the trained model is stored on a cloud provider like the Google Cloud Platform. By leveraging cloud infrastructure, the model can be accessed remotely, enabling users to interact with it and obtain predictions efficiently. This approach ensures that the model's functionality is readily available to users, facilitating seamless integration into various applications or systems. The deployed model is executed within a Docker container environment, which includes the Flask framework for building the API (Application Programming Interface) [25]. The API serves as the communication channel between users and the machine learning system. In this case, the REST (Representational State Transfer) API communication protocol is utilized for seamless interaction between users and the system, particularly in Android application services [26][27].

Fig. 3 provides a visual representation of the machine learning system design employed in this study. It illustrates the various components involved, including the cloud storage for the model, the Docker container environment, the Flask API framework, and the REST API for user communication. This design ensures the efficient and user-friendly deployment of the machine learning model, making it accessible and usable for various applications.

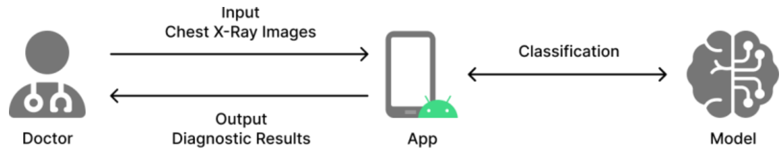

Fig. 3 Android-Based Machine Learning System

3. RESULTS AND DISCUSSION

In this research, we conducted a comparative analysis with a previous study that involved combining pneumonia data with CNN methodology. We examined the effects of using dropout layers and not using them, as well as applying hyperparameter tuning and varying the number of epochs from 25 to 200. Additionally, we explored the performance of both CNN and CNN-ELM approaches. The accuracy results obtained in our study ranged between 89% and93%.

Based on Table 1, we categorized the data into three sections: training, validation, and testing. The dataset consisted of categorical data, specifically divided into two classes: normal data with 1583 samples and pneumonia data with 4237 samples. As part of the preprocessing stage, we applied image rescaling from 1000 x 1000 pixels to 224 x 224 pixels to perform data augmentation. During the testing phase, we conducted two stages: CNN without augmentation and CNN with augmentation. The purpose of data augmentation was to address data imbalance, which is evident in Table 1. By augmenting the data, we aimed to achieve a more balanced distribution. Moreover, data augmentation helps minimize overfitting during the data modeling process.

We devised the CNN-VGG16 algorithm in this research. In our approach, we introduced modifications to the top layer. After the 5th block's max-pooling layer, we added a flatten layer, followed by a fully connected layer with ReLU activation. The output layer utilized the sigmoid activation function. To assess the algorithm's efficacy, we split the dataset into a training set comprising 5126 samples (90%) and a testing set of 624 samples (10%). Evaluating the algorithm's performance involved utilizing model evaluation metrics based on the confusion matrix, including accuracy, precision, recall, and F1-Score calculations.

.

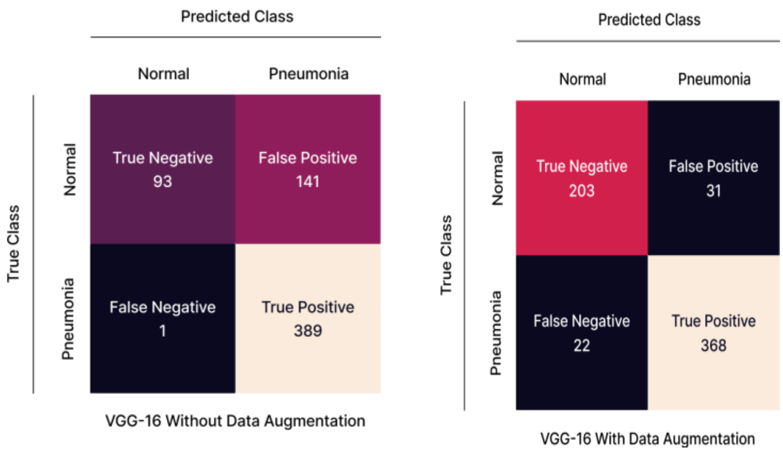

Fig. 4 Matrix Confusion

In this experiment, the objective was to assess the impact of data augmentation on the performance of the VGG-16 architecture. The outcomes of the trials were summarized in Figure 4, which presented the confusion matrix depicting the entire trial process. The confusion matrix provided a comprehensive overview of the model's performance, showing the distribution of predicted labels against the actual labels. The confusion matrix enabled the examination of true positives, true negatives, false positives, and false negatives, offering valuable insights into the model's accuracy, precision, recall, and overall effectiveness in classifying the provided data. Table 3 illustrated the outcomes of classifying normal pneumonia, with and without augmentation, highlighting discrepancies in accuracy, precision, recall, and F1-Score. Accuracy represented the overall correctness of the classification model, while precision gauged its capability to avoid false positives. Recall emphasized the model's proficiency in avoiding false negatives, while F1-Score provided a balanced evaluation, considering both precision and recall simultaneously.

Table 3. Classification Results

Scenario	Category	Accuracy	Precision	Recall	F1-Score
CNN-VGG16 without Augmentation	Normal	77%	0.99	0.40	0.57
	Pneumonia		0.73	1.00	0.85
CNN-VGG16 with Augmentation	Normal	92%	0.90	0.87	0.88
	Pneumonia		0.92	0.94	0.93
Improvement		**15%**			

The performance testing of the VGG-16 architecture without data augmentation resulted in an accuracy of 77%, while the testing with data augmentation achieved an accuracy of 92%. When comparing the evaluation results of the VGG-16 architecture from both testing scenarios, the best performance was observed in the second testing scenario. The difference is quite significant, with an improvement about 15% in accuracy between with-augmentation and without-augmentation accuracy of value.

Table 4. Measurements Calculating based on Formula (1)-(4)

Scenario	Accuracy	Precision	Recall	F1-Score
CNN-VGG16 without Augmentation	$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} = \frac{389+93}{389+141+93+1} = \frac{482}{624} = 0{,}77 \approx 77\%$	$Precision = \frac{389}{389+141} = \frac{389}{530} = \frac{389}{530} = 0{,}73$	$Recall = \frac{389}{389+1} = 0{,}99 \approx 1$	$F1-Score = \frac{2 \times (0{,}73 \times 1)}{(0{,}73+1)} = \frac{1{,}46}{1{,}73} = 0{,}85$
CNN-VGG16 with Augmentation	$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} = \frac{368+203}{368+31+308+22}$	$Precision = \frac{TP}{TP+FP} = \frac{368}{368+31}$	$Recall = \frac{TP}{TP+FN} = \frac{368}{368+22}$	$F1-Score = \frac{2 \times (Precision \times Recall)}{(Precision + Recall)} = \frac{2 \times (0{,}92 \times 0{,}94)}{(0{,}93+0{,}94)}$

$= \frac{571}{624}$ $= 0{,}92 \approx 92\%$	$= \frac{368}{399}$ $= 0{,}92$	$= \frac{368}{390}$ $= 0{,}94$	$= \frac{1{,}73}{1{,}87}$ $= 0{,}93$

Table 4 presents the evaluation of the CNN-VGG16 algorithm using accuracy, precision, recall, and F1-Score calculations. The appropriateness of the chosen formulas (1) to (4) was supported by computational calculations. The evaluation considered both the original dataset and an augmented dataset, enhancing the algorithm's ability to handle data variations. The results provided a comprehensive assessment of the algorithm's performance, covering overall correctness, positive prediction ratio, positive instance detection, and an overall effectiveness measure. This evaluation process offered valuable insights into the algorithm's classification and prediction capabilities.

Table 5. Comparing the Results with Previous Research

Scenario	Class of Data	Accuracy
CNN-VGG16 without Augmentation	Normal	77%
	Pneumonia	
CNN-VGG16 with Augmentation	Normal	92%
	Pneumonia	
CNN-dropout layer and Augmentation [5]	Pneumonia	90,68%
CNN-Hipertunning Parameter [12]	Pneumonia	89,58%
CNN	Pneumonia	84,78%
CNN-ELM [22]	Pneumonia	93,59%

Referring to Table 5, the comparison between our study and the previous study reveals that CNN-VGG16 with data augmentation performs better than other CNN algorithm scenarios. However, it is important to note that while CNN-ELM does show slightly higher accuracy, the difference is not significant and it approaches the level of accuracy achieved in our research. Basen on our analysis, CNN-VGG16 and CNN-ELM are two different approaches in image processing. CNN-VGG16 utilizes a deep and complex CNN architecture, requiring intricate preprocessing and data augmentation to achieve good results in object recognition tasks. On the other hand, CNN-ELM combines CNN's convolutional layers with the simpler Extreme Learning Machine (ELM), resulting in shorter training times and often yielding good results in image processing tasks with small datasets or limited resources. The choice between the two depends on the specific task, dataset size, and available resources. Nevertheless, this study's constraint is that it concentrated on a particular CNN architecture. Subsequent investigations should investigate how data augmentation influences different architectures and other tasks related to medical image analysis. By doing so, we can further verify the efficacy of data augmentation.

The CNN-VGG16 approach, with or without augmentation, utilizes numerical computations and is integrated into an application developed for the Android system in the healthcare field. We deployed the CNN-VGG16 model into the Android application and published it on the Android platform. The CNN-VGG16 algorithm runs as a background process on the Android system, achieving the accuracy levels we have developed in our research. The Android system we developed is user-friendly, allowing healthcare professionals to diagnose patients by inputting their X-ray image data along with their name label. The diagnostic results are then categorized as normal or indicating pneumonia. Overall, the integration of our advanced machine learning-based system in healthcare shows potential for improving patient care.

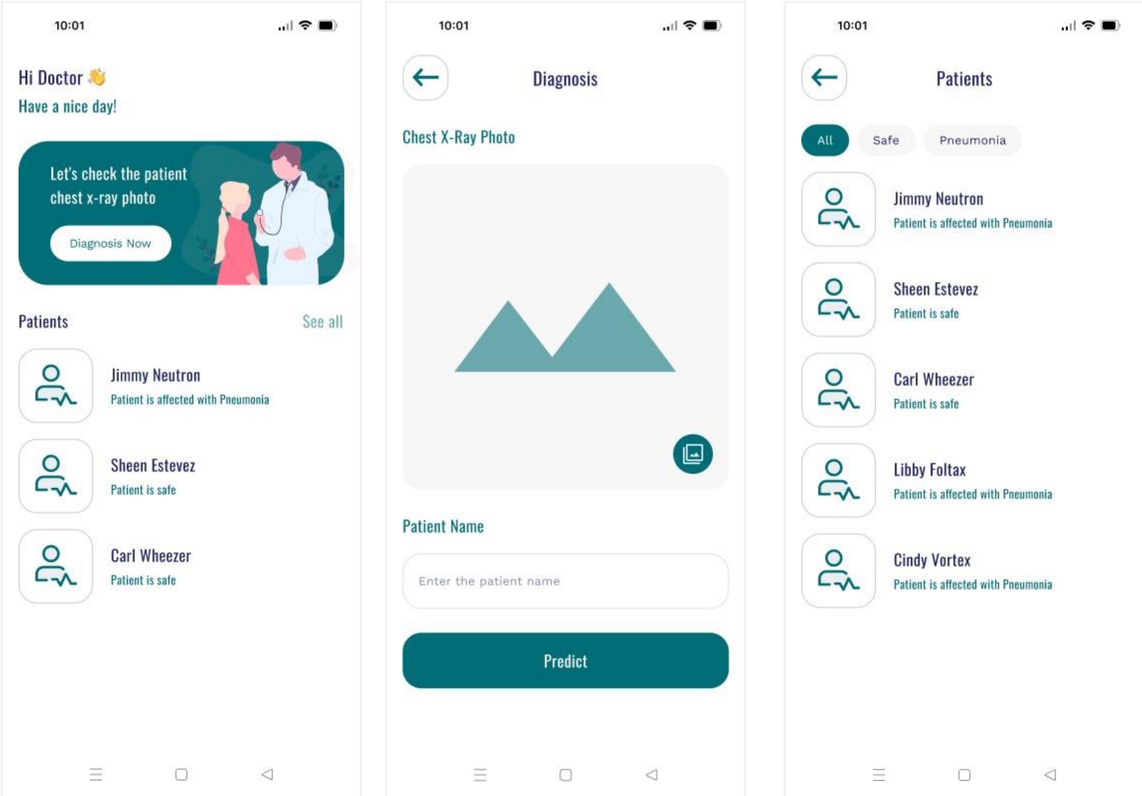

Fig. 5 Android-Based Machine Learning System User Interface

The operation of the application is fairly straightforward. To conduct a diagnosis, a doctor can access the diagnosis page and upload an x-ray image together with the patient's name. Subsequently, the application will analyze the uploaded image and provide a diagnosis, which will be displayed on the dashboard page. Furthermore, the diagnosis results will be automatically stored in the local storage of the Android smartphone being utilized. Additionally, doctors have the ability to review the diagnostic history on the patient page.

4. CONCLUSION

The study focused on using a modified VGG-16 architecture with a CNN algorithm to classify pneumonia. The research findings show that employing data augmentation techniques significantly improves the model's performance when dealing with imbalanced data classification. The study contributes by demonstrating the effectiveness of data augmentation in improving the accuracy and robustness of machine learning models for pneumonia diagnosis. The research has important implications for assisting doctors in analyzing lung images and improving healthcare outcomes. The implications of the study are significant for the field of healthcare and medical image analysis. By employing data augmentation techniques, the researchers were able to enhance the performance of a modified VGG-16 architecture in classifying pneumonia. This finding holds the potential to improve the accuracy and reliability of pneumonia diagnosis, ultimately leading to better healthcare outcomes. The study also highlights the effectiveness of data augmentation in addressing the challenge of imbalanced data classification, which is a common issue in medical image analysis. By mitigating this problem, healthcare professionals can rely on more robust and accurate machine learning models. Additionally, the study suggests that the benefits of data augmentation may extend beyond the specific CNN architecture used, prompting further exploration of its impact on other architectures and medical imaging tasks. Overall, these implications underscore the promise of integrating advanced machine learning systems into healthcare, offering the potential to enhance patient care and revolutionize medical diagnosis and treatment. However, the study's limitation lies in its focus on a specific CNN architecture, and future research should explore the impact of data augmentation on other architectures and medical image analysis tasks to further validate its effectiveness. Overall, integrating advanced machine learning systems in healthcare shows promise for enhancing patient care.

REFERENCES

[1] Shao, Z., et al. 2020. “AI 2000: A Decade of Artificial Intelligence.” *12th ACM Conference on Web Science, Proceeding* 2020 - July: 345-354. https://doi.org/10.1145/3394231.3397925

[2] Pallathadka, H., et al. 2021. “IMPACT OF MACHINE learning ON Management, healthcare AND AGRICULTURE.” *Mater Today Proceeding* 2021-July,80: 2803-2806. 10.1016/j.matpr.2021.07.042

[3] Destiawan, D, et al. 2018. “Implementasi Fuzzy Logic Pada Short Range Radar Untuk Pengamanan BT (Basis Tempur) Tingkat Regu.” *Jurnal Teknik Elektro dan Komputer TRIAC* 5 (2). https://doi.org/10.21107/triac.v5i2.4062

[4] Bharati, S., et al. 2020. “Hybrid deep learning for detecting lung diseases from X-ray images.” *Inform Med Unlocked*. 20: 100391. https://doi.org/10.1016/j.imu.2020.100391

[5] Sharma, H., et al. 2020. “Feature Extraction and Classification of Chest X-Ray Images Using CNN to Detect Pneumonia,” *2020 10th International Conf. on Cloud Computing, Data Science & Engineering, Proceeding* 2020 - January: 227–231. https://doi.org/10.1109/Confluence47617.2020.9057809

[6] Al Haromainy, M. M., et al. 2022. “Classification Of Javanese Script Using Convolutional Neural Network With Data Augmentation.” *2022 IEEE 8th Information Technology International Seminar (ITIS), Proceeding* 2023-January: 288-291. https://doi.org/10.1109/ITIS57155.2022.10010262

[7] Ksibi, A., et al. 2022. “Improved Analysis of COVID-19 Influenced Pneumonia from the Chest X-Rays Using Fine-Tuned Residual Networks,” *Computer Intelligence Neuroscience*: 1-13. https://doi.org/10.1155%2F2022%2F9414567

[8] Mardlijah., et al. 2018. “OPTIMAL CONTROL OF NEUTRAL LIPIDS IN MICROALGAE PRODUCTION WITH NUTRIENT LIMITATION,” *MATTER: International Journal of Science and Technology*. 4 (1): 61-72. https://doi.org/10.20319/mijst.2018.41.6172

[9] Sari, A. P., et al. 2022. “Forecasting Model of Wind Speed and Direction by Convolutional Neural Network - Deep Convolutional Long Short Term Memory.” *2022 IEEE 8th Information Technology International Seminar (ITIS), Proceeding* 2023-January: 200-205. https://doi.org/10.1109/ITIS57155.2022.10010005

[10] Riyantoko, P. A and Muhaimin, A. 2023. “A simple data sentiment analysis using Bjorka phenomenon on Twitter.” *Nusantara Science and Technology Proceedings* 33: 330-336. https://doi.org/10.11594/nstp.2023.3353

[11] Fahrudin, T. M., et al. 2021. " KEBI 1.0: Indonesian Spelling Error Detection System for Scientific Papers using Dictionary Lookup and Peter Norvig Spelling Corrector." Lontar Komputer : Jurnal Ilmiah Teknologi Informasi 12 (2): 78-90. https://doi.org/10.24843/LKJITI.2021.v12.i02.p02

[12] Maysanjaya, I. Md. D. 2020. “Klasifikasi Pneumonia pada Citra X-rays Paru-paru dengan Convolutional Neural Network.” *Jurnal Nasional Teknik Elektro dan Teknologi Informasi*. 9 (2): 190-195. https://doi.org/10.22146/jnteti.v9i2.66

[13] Sari, A.P., Hiroshi Suzuki, Takahiro Kitajima, Takashi Yasuno, Dwi Arman Prasetya, and Rahman Arifuddin. (2022). “Short-Term Wind Speed and Direction Forecasting by 3DCNN and Deep Convolutional LSTM.” *IEEJ Transactions on Electrical and Electronic Engineering*, 17 (11):1620-1628. https://doi.org/10.1002/tee.23669

[14] Sari A.P., D.A. Prasetya, M. M. Al Haromainy, F.P. Aditiawan, A. N. Sihananto, W. S. Saputra. 2022. Analisis Faktor Kesuksesan Penggunaan eBelajar Menggunakan Metode Hot-Fit di STIKI Malang. *Seminar Nasional Sains Data (SENADA) 2022, prosiding* 2022-November: 92-102. https://prosiding-senada.upnjatim.ac.id/index.php/senada/article/view/54/16

[15] Yamashita, R., et al. 2018. “Convolutional neural networks: an overview and application in radiology.” *Insights Imaging*, 9 (4): 611–629. https://doi.org/10.1007/s13244-018-0639-9

[16] Sari, A.P., et al. 2020. “Prediction Model of Wind Speed and Direction using Convolutional Neural Network-Long Short Term Memory.” *2020 IEEE International Conference on Power and Energy (PECon) Proceeding* 2020-December: 358-363. https://doi.org/10.1109/PECon48942.2020.9314474

[17] Tu, X. et al. 2019. “Enhance the Motion Cues for Face Anti-Spoofing using CNN-LSTM Architecture.” Arxiv: Computer Vision and Pattern Recognition, 1-20. https://doi.org/10.48550/arXiv.1901.05635

[18] Sari, A.P., Hiroshi Suzuki, Takahiro Kitajima, Takashi Yasuno, Dwi Arman Prasetya, and Abd. Rabi'. 2021. “Deep convolutional long short-term memory for forecasting wind speed and direction.” *SICE Journal of Control, Measurement, and System Integration* 14(2): 30-38. https://doi.org/10.1080/18824889.2021.1894878

.

[19] Naseer, I., et al. 2022. "Performance Analysis of State-of-the-Art CNN Architectures for LUNA16." *Sensors*. 22 (12): 4426. https://doi.org/10.3390/s22124426
[20] Putrama Alkhairi and A. P. Windarto. 2023. "Classification Analysis of Back propagation-Optimized CNN Performance in Image Processing," *Journal of Systems Engineering and Information Technology (JOSEIT)* 2 (1): 8-15. https://doi.org/10.29207/joseit.v2i1.5015
[21] Purwono, P., et al. 2023. "Understanding of Convolutional Neural Network (CNN): A Review." *International Journal of Robotics and Control Systems* 2 (4): 739-748. https://doi.org/10.31763/ijrcs.v2i4.888
[22] Nugroho , B. and E. Y. Puspaningrum. 2021. "Kinerja Metode CNN untuk Klasifikasi Pneumonia dengan Variasi Ukuran Citra Input." *Jurnal Teknologi Informasi dan Ilmu Komputer* 8 (3): 533-538. http://dx.doi.org/10.25126/jtiik.2021834515
[23] Saputra, W. S. J., et al. 2022. "Car Classification Based on Image Using Transfer Learning Convolutional Neural Network." *2022 IEEE 8th Information Technology International Seminar (ITIS), Proceeding* 2023-January: 324-327. https://doi.org/10.1109/ITIS57155.2022.10010073
[24] Nugraha, G. S., et al. 2023. "Comparison of CNN's Architecture GoogleNet, AlexNet, VGG-16, Lenet-5, Resnet-50 in Arabic Handwriting Pattern Recognition." *Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control*. http://dx.doi.org/10.3390/electronics11111780
[25] N. Wiranda and A. E. Putra. 2022. "Mobile-based Primate Image Recognition using CNN," *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*. 16 (2): 149-158. https://doi.org/10.22146/ijccs.65640
[26] M. M. Al Haromainy, D. A. . Prasetya, and A. P. Sari, "Improving Performance of RNN-Based Models With Genetic Algorithm Optimization For Time Series Data", TIERS, vol. 4, no. 1, pp. 16-24, Jun. 2023.
[27] F. D. Marleny and M. Mambang, "Predictive Modeling Classification of Post-Flood and Abrasion Effects With Deep Learning Approach", TIERS, vol. 3, no. 1, pp. 1-10, Jun. 2022.