JOIV : Int. J. Inform. Visualization, 7(3) - September 2023 651-658

INTERNATIONAL JOURNAL ON INFORMATICS VISUALIZATION

journal homepage : www.joiv.org/index.php/joiv

Chest X-ray Image Classification to Identify Lung Diseases Using Convolutional Neural Network and Convolutional Block Attention Module

Chandra Halim [a,*], Nathanael Geordie Eka Putra [a], Nico Ardian Nugroho [a], Derwin Suhartono [a]

[a] *Computer Science Department, School of Computer Science, Bina Nusantara University, Palmerah, Jakarta 11480, Indonesia*

Corresponding author: [*]chandra.halim001@binus.ac.id

Abstract—**Image classification, the process of categorizing and labeling groups of pixels or vectors within an image based on specific rules, is continuously developed by many researchers in the world to solve many problems. One of those problems is x-ray image classification to determine lung diseases. This research tries to solve the problem of classifying COVID-19, pneumonia, and healthy lungs using x-ray images. The image datasets were collected from several sources. This research aims to build a reliable and robust Convolutional Neural Network (CNN) enhanced with Convolutional Block Attention Module (CBAM) mechanism. CNN is used to do the feature extraction and the classification, whereas CBAM is used to improve the performance of the CNN by focusing on the important features in given data. Research methods are done through extensive data selection, preprocessing, and parameter tuning to achieve a well-performing model. While there is still a lack of research on x-ray classification using the attention mechanism, this research proposes it as the main method. This research also does a further experiment on the effect of the imbalanced dataset on the model. The evaluation is done using a cross-validation method. This research results reach 97.74% of accuracy, precision, recall, and f1-score. This research concludes that CBAM increases the performance of a CNN module. Using a larger dataset can be beneficial in this kind of research as well as evaluation by radiologists.**

Keywords— **Chest X-ray; COVID-19; pneumonia; healthy lungs; CNN; CBAM.**

Manuscript received 23 Aug. 2022; revised 30 Oct. 2022; accepted 26 Jan. 2023. Date of publication 10 Sep. 2023.

I. INTRODUCTION

Lungs are vital organs in the human body that do respiration and support human life. Many diseases could infect the lungs. So, it is necessary to have an effective yet efficient method to detect those diseases. This research sought a reliable method to classify recent viral lung diseases, COVID-19, and pneumonia.

On December 31, 2019, World Health Organization (WHO) office in China got a statement from Wuhan Municipal Health Commission regarding viral pneumonia and unknown cause pneumonia cases in Wuhan city [1]. As we know, after this report, COVID-19 caused pandemic caused by SARS-CoV-2 virus, and it has been ongoing for approximately 2.5 years since this research was written. This pandemic changes how people live globally in every aspect of life. As of April 8, 2022, WHO confirmed 494,587,638 cases of COVID-19, with total death reaching 6,170,283 individuals [2]. This disease infects humans and causes respiratory symptoms resembling flu. WHO summarized some symptoms of COVID-19 involving common symptoms like cold, cough, fatigue, losing sense of taste and smell, and in some cases include sore throat, headache, pains, diarrhea, rash on skin, and irritated eyes [3], [4].

Until today, RT-PCR was considered as the golden standard of COVID-19 detection. RT-PCR is done by taking the sample from *orofaring* and *nasofaring* using swab method [5]. Unfortunately, there were concerns regarding sensitivity of RT-PCR, which was considered low [6], [7]. Alternative screening method that has also been utilized for COVID-19 screening has been radiography examination, where chest radiography imaging (e.g., chest X-ray (CXR) or computed tomography (CT) imaging) is conducted and analyzed by radiologists to look for visual indicators associated with SARS-CoV-2 viral infection [8] [9] [10]. Sathi et al [11] stated that there are some patterns on the COVID-19 CXR.

On the other side, Pneumonia also has similar symptom comparing to COVID-19 such as cough, shortness of breath, pounding heart, cold, sweating or shivering, loss of appetite, and chest pain [12]. Rahman [13] states that the most effective way to identify pneumonia is using CXR. Many researches have been done to develop method to identify COVID-19 and pneumonia. Research by Ibrahim [14] developed a model using pre-trained AlexNet architecture to do binary classification and multiclass classification. This research showed that the main model achieves 93.42% accuracy on multiclass classification scenarios. Hammoudi [15] compared pre-trained architectures such as ResNet34, ResNet50, and DenseNet169, and a combination of Inception ResNetV2 & RNN. The final result showed that the best algorithm is DenseNet169, with 92.57% accuracy.

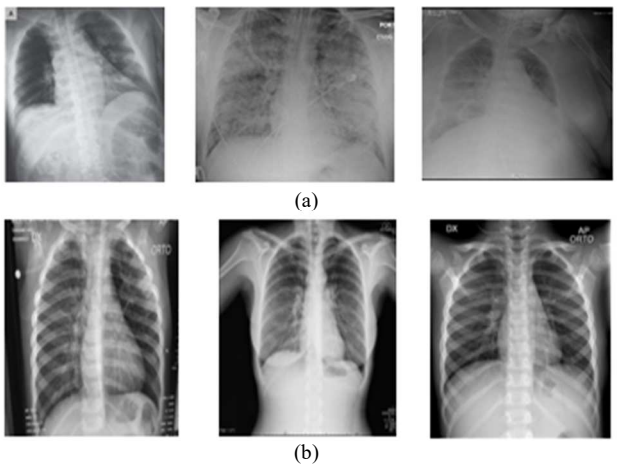
Another research adopted an attention mechanism in its model. AVNC research developed a VGG-style network and Convolutional Block Attention Module (CBAM) [16]. This research gave a result of 96.78% of the f1-score. Further research used Darknet as a backbone to identify COVID-19, pneumonia, and normal lungs [17]. The results that were achieved on binary classification were 98.08% of accuracy and 87.02% of accuracy on multiclass classification.

This study aims to develop a novel model to identify COVID-19, pneumonia, and normal lungs and improve it using the CBAM mechanism proposed by Woo [18]. Another purpose of this study is to test the effect of the under sampling method on the dataset on the model's performance.

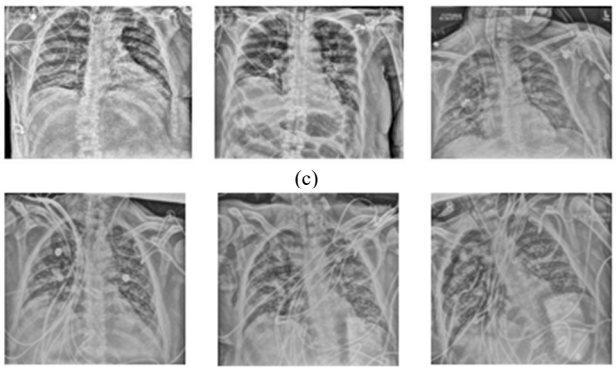
II. Material and Method

A. Dataset

This research utilizes an x-ray image dataset from two sources. COVID-19 dataset is collected from *Società Italiana di Radiologia Medica e Interventistica* (SIRM), consisting of 943 positive COVID-19 CXR used in [19]. These images are gathered in jpg image format. Although there were three more data sources curated [19], data from SIRM were chosen because it provides cleaner data compared to other available research sources. The initial experiment in this research also shows that combining data from multiple sources hinders the model's ability to detect patterns in the data. Samples from SIRM and other sources are provided below to compare data clarity.



(a) (b)

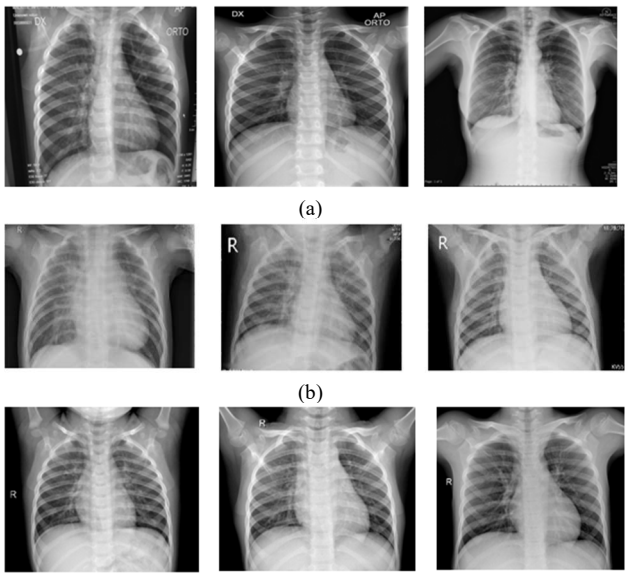


(c) (d)

Fig. 1 COVID-19 CXR dataset Sample from: a) Cohen. b) SIRM. c) Ricord. d) Stonybrook

Figure 1 shows that data from Cohen are mostly blurry, whereas Ricord and Stonybrook data seem off in contrast and contain numerous noises, such as medical tubes in the samples above. Of the four sources [19], SIRM clearly provides the best dataset for this research.

The normal lungs and pneumonia datasets are collected from another research which consists of 3,875 pneumonia CXRs and 1,341 normal lungs CXRs [15]. This dataset is divided into train and test datasets using cross-validation method.



(a) (b) (c)

Fig. 2 Dataset sample from each class: a) COVID-19, b) Pneumonia, c) healthy lungs

Dataset distribution can be summarized in Table I.

TABLE I
DATASET DISTRIBUTION

	COVID-19	Pneumonia	Normal
Total Images	943	3,875	1,341

B. Data Preprocessing

Data on machine learning cannot always be input for the model. This happens because the machine cannot understand data like free text, pictures, or videos. Using a defective dataset can mislead the whole training process[20]. Chicco [21] states that data preprocessing is the most important process before the model is trained. Data preprocessing is needed to convert those data into an understandable format for the machine. This research did three data preprocessing methods: resizing, dataset balancing, and gray scaling. The CXR datasets are resized to 256x256, so they can be fitted into the CNN model and reduce computation costs. This technique is proven not to alter spatial information contained in the data, as stated by Tang [22]. The input data size does not significantly affect the model's performance. The amount of the dataset is also balanced using random under sampling. The purpose of this method is to subtract the amount of the majority class until it is the same as the amount of the minority class. The imbalance dataset can be defined as group data in which the interclass ratio is imbalanced [23]. An imbalanced class ratio in each subset could lead to leaning bias toward the class with the highest ratio [24].

TABLE II
DATA ALLOCATION AFTER UNDER SAMPLING

	Covid	Pneumonia	Normal
Train	754	755	754
Test	189	188	189
Total	943	943	943

Even though this research uses CXR, which is already in gray, it is still necessary to grayscale it so the model doesn't treat the dataset as an RGB. If this research uses RGB as the input data, then three channels on the input layer will affect the model's performance. Utilization of this method in machine learning algorithms has been proven to give better recognition ability rather than using data in RGB form and also to increase model's efficiency in the training process [25].

C. Experiments

Our experiment begins with using DarkCovidNet as the initial CNN model as it was proven to perform well and be light enough despite lacking training resources. Then, to achieve better results, this research does several parameter adjustments within the Convolutional Neural Network:

- The sum of convolutional layer (17,10,5)
- Activation function (LeakyReLU & ReLU)
- Epoch (100,50,25)

Each time the experiment gives a better result, it is adjusted using other parameters until it becomes the optimal model for the dataset. Details on how each parameter is adjusted are described thoroughly in section III.

CBAM was incorporated into the model to optimize the model even more. CBAM consists of Channel Attention Module (CAM) and a Spatial Attention Module (SAM) which is placed between each convolutional layer in CNN to filter features in data both channel and spatial-wise.

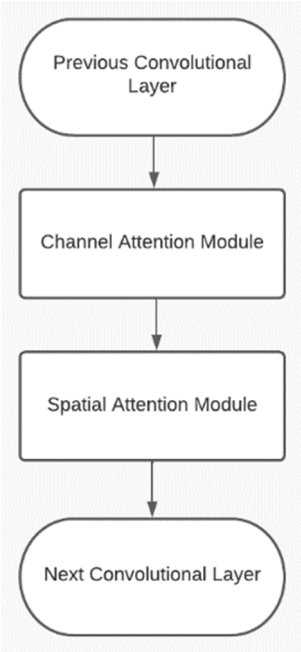


Fig. 3 CAM and SAM placement within the convolutional layer

CAM and SAM are placed sequentially to identify whether the features are important or not. CAM is used to identify the important part of the given feature, and SAM is used to locate the important feature. CBAM's implementation can be summarized as:

$$F' = Mc(F) \otimes F \quad (1)$$

$$F'' = Ms(F') \otimes F' \quad (2)$$

Where $\otimes$ denotes element-wise multiplication. During multiplication, the attention values are broadcasted evenly: channel attention values are broadcasted along the spatial dimension, and vice versa. F'' is the final refined output. Fig. 3 portrays the computation process of each attention module. The following describes the details of each attention module.

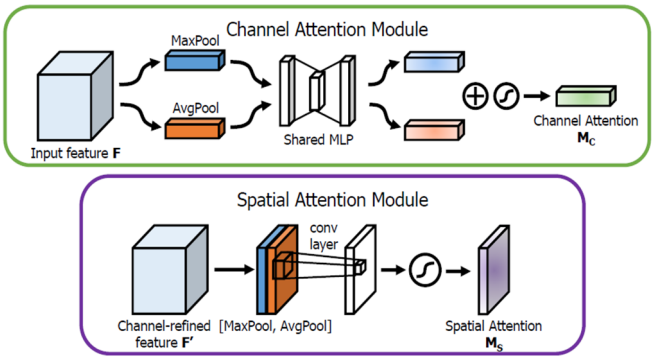


Fig. 4 Diagram of each Attention Module

The algorithms start with activating the channel attention. Woo [18] has proven the usage of max pooling and average pooling to the feature map simultaneously, increasing the representation ability of the model rather than using them separately. Next, both pooling results became the input to one hidden layer multi layers perceptron before being added to each other per matrix element and activated using the sigmoid

function. The channel attention map can be summarized as follows:

$$M_c(F) = \sigma(MLP(AvgPool(F)) + MLP(MaxPool(F))) \quad (3)$$

Where σ denotes the sigmoid function and F is the feature map from the previous convolutional layer. Operation done in the channel attention module is forwarded to the spatial attention module, which aggregates using max pooling and average pooling to the feature map generated by the channel attention map. This process creates two new 2D feature maps to be convoluted using a 7x7 size kernel. Channel attention module can be summarized as:

$$M_s(F) = \sigma(f^{7x7}([AvgPool(F); MaxPool(F)])) \quad (4)$$

Where F is the feature map from the channel attention module and σ denotes the sigmoid function. The preliminary experiment suggests the best CNN model to solve this problem achieved using 5 convolutional layers. Therefore, such a model was selected to be incorporated with CBAM. The number of filters used in this model increases over the layers to capture in-depth features in each data. Each layer uses 3x3 kernel size with a Maxpooling operation between each layer. Dropout layers were placed in several layers to reduce overfitting. Dropout is a regularization technique to avoid overfitting on the model by randomly eliminating several neurons and the connection to the other layers[26]. Details on each layer of the CNN architecture can be observed in Table III.

TABLE III
CNN ARCHITECTURE

Layer	Sum of *Filters*	*Kernel Size*	*Stride*
1	32	3x3	1
2	64	3x3	1
3	64	3x3	1
4	128	3x3	1
5	256	3x3	1

Table IV provides detailed information of the hyperparameter used in the model.

TABLE IV
CNN HYPERPARAMETER

Batch Normalization	Per convolutional layer
Max pooling	**Kernel**: 2x2
	Stride: 2
Dropout	0.1 (2nd layer)
	0.2 (4th, 5th, dense)
Optimizer	Stochastic Gradient Descent (SGD)
Loss Function	Cross Entropy
Batch Size	32
Learning Rate	0.01

As this research aims to compare the model's performance with and without CBAM, the structure is slightly altered to facilitate the experiment. The structure is shown in Fig. 5.

D. Training

Some steps do the training process. First, the dataset is loaded by an image data generator and fed into the model. After the convolutional proses, the result comes as feature maps. The feature maps are flattened into a 1-dimensional array, so it fits into the neural network. Finally, the neural network gives the classified class from the highest probability of the classes. CBAM is also applied to the model between each convolutional layer to improve the model's ability to classify some classes so it can be more precise. When the feature maps reached the CBAM layer, it would be fed into the channel module first to extract the channel feature, and then the channel features were fed into the spatial module to extract the spatial features. Finally, spatial features were added to the feature maps from the previous convolutional layer. This research also trains using balanced and imbalance datasets to investigate the effect of imbalance and balanced datasets on the model's final performance.

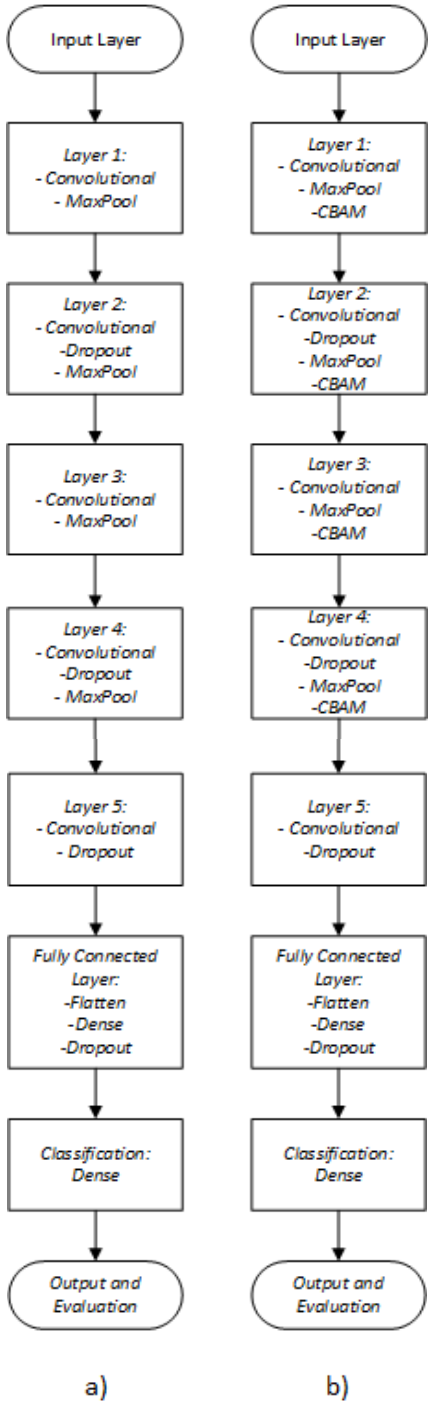


Fig. 5 Model Structure a) without CBAM, and b) using CBAM

E. Evaluation

Evaluation is done using a cross-validation method. To be more detailed, this research uses a stratified k-fold method to split the dataset into train and test datasets equally. A stratified K-fold can ensure that each class's ratios are the same in every fold [27]. This experiment uses 5-fold as it was indicated to have a better result on various test cases [28]. [29] Stated that the most optimal evaluation method for classification is using the confusion metric. To assess the model's performance, this research uses evaluation metrics such as accuracy, f1-score, precision, and recall which are calculated based on the given confusion metric. These evaluation metrics are recorded in each fold. In the end, the mean of the evaluation metric is calculated and serves as the final performance result. Fig. 4 depicts how the evaluation mechanism has been done in this research.

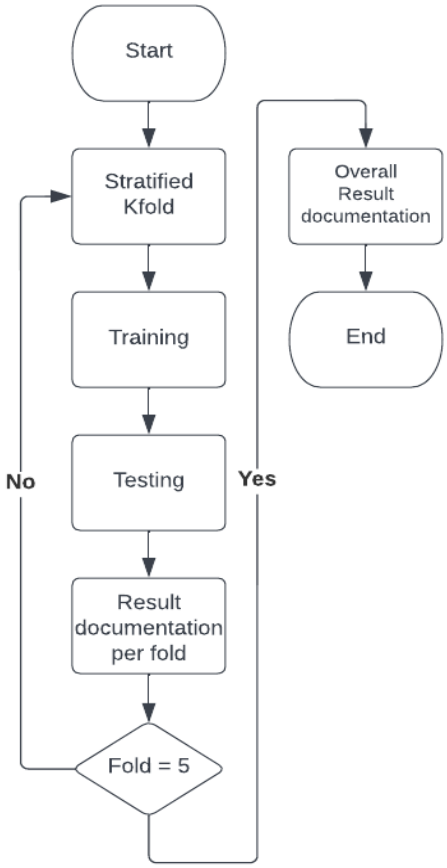


Fig. 6 Evaluation mechanism in this research

The flow starts with doing the stratified k-fold to split the dataset into training and testing datasets. The algorithm splits the dataset into 80:20 ratio of training and testing datasets. This ratio is chosen because this ratio does not produce high bias and variance [30]. A stratified k-fold also splits the classes evenly. Next is training the CBAM model. Each experiment uses different numbers of epochs in order to find the optimal one. Every time the model finishes training, it is tested using the testing dataset, and the result is recorded into the evaluation metrics per fold. These steps are repeated until the fold reaches 5-fold, and finally, the final result is generated by calculating the mean of each evaluation metric in each fold.

III. Result and Discussion

As described in section I, this research aims to develop a novel model to identify COVID-19, pneumonia, and normal lungs. So, this research does some research to find an optimal configuration for the model, starting from the amount of convolutional layer, activation functions, and the number of the epoch being used. At the beginning of the experiment, this research used the CNN architecture from DarkCOVIDNet with 17 convolutional layers, Leaky ReLU as the activation function, and 100 numbers of epochs. This experiment gives insight that the model is overfitted.

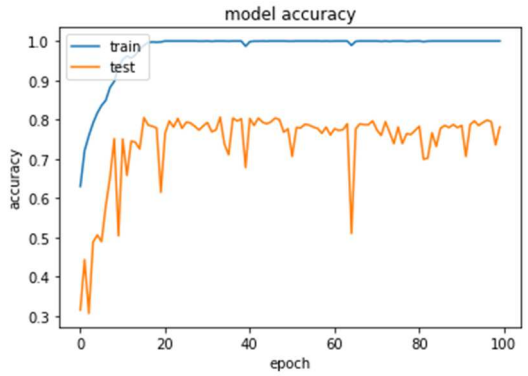


Fig. 7 Train to test accuracy comparison of the first scenario

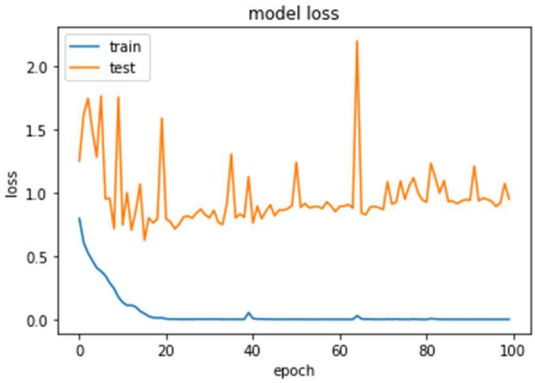


Fig. 8 Train to test loss comparison of the first scenario

TABLE V
CONFUSION MATRIX OF FIRST SCENARIO

		Predicted		
		COVID-19	**Normal**	**Pneumonia**
Class	**COVID-19**	310	4	11
	Normal	10	173	167
	Pneumonia	43	62	222

The next experiment reduced the number of the convolutional layer to 10 and also reduced the number of epochs to 50 epochs. The purpose of this experiment is to cut the computational cost of this research regarding to lack of computational resource. This experiment gives insights that reducing does reduced overfit as well and in the case of reducing number epochs, the model still got the optimal performance and cut the computational cost. This experiment is still overfit, because of the dataset set that was used. The dataset consists of thousands of CXR from many sources in each class and it's hard to split it evenly into train and test dataset. Even though it was split evenly, it could still overfit to certain data source for each class because of the large gap

of numbers of data from each class. So, this research used the dataset as it is described in section two.

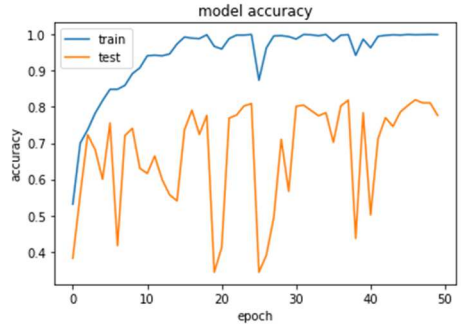


Fig. 9 Train to test accuracy comparison of the second scenario

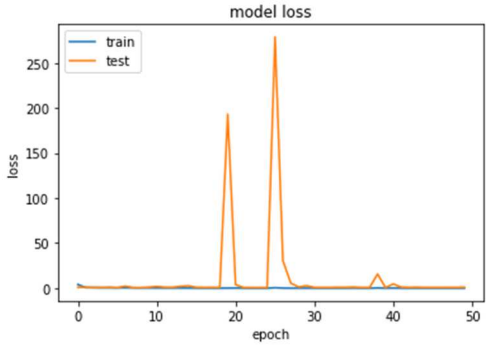


Fig. 10 Train to test accuracy comparison of the second scenario

TABLE VI
CONFUSION MATRIX OF SECOND SCENARIO

		Predicted		
		COVID-19	**Normal**	**Pneumonia**
Class	**COVID-19**	279	13	33
	Normal	1	203	146
	Pneumonia	0	35	292

The next experiment reduced the convolutional layers to 5 layers and tried to use ReLU activation function because dying neuron problem rarely happened in this dataset. It is rare because the dataset is in image form, so the pixel is rarely valued as zero. This experiment also reorders the dataset used in training where it seems clear from the confusion matrix that using a mixed-source dataset affects the model from learning the pattern contained in the dataset. These changes result in optimal architecture to process the dataset. The final performance metrics show a good performance.

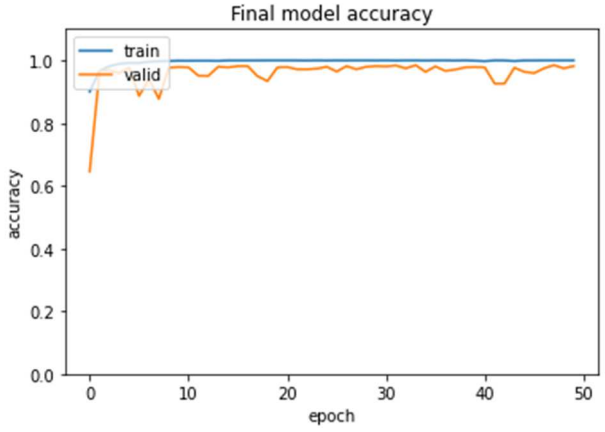


Fig. 11 Train to test accuracy comparison of the third scenario

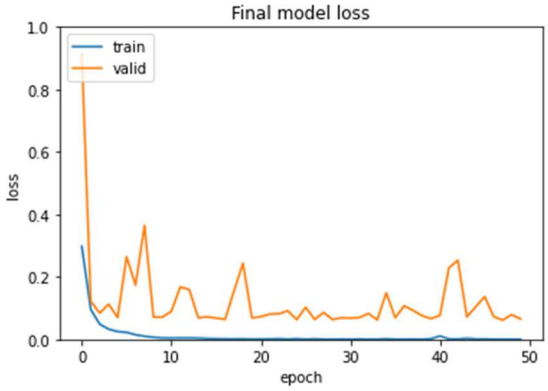


Fig. 12 Train to test loss comparison of the third scenario

TABLE VII
CONFUSION MATRIX OF THIRD SCENARIO

		Predicted		
		COVID-19	**Normal**	**Pneumonia**
Class	**COVID-19**	184	0	4
	Normal	3	251	14
	Pneumonia	3	7	765

After some experiments to create an optimal model, this research did further experiments to assess the ability of CBAM to improve the model's performance and the effect of using an imbalanced dataset and a balanced dataset. This experiment implements stratified k-fold cross-validation to verify the model's performance. Final performance metrics are calculated by finding the mean of the performance metrics for each fold. This experiment shows that data balancing affects the model performance and CBAM also affects the model performance. Both of these techniques increase the accuracy of the model to classify CXR into the three classes.

TABLE VIIII
EFFECT ON UNDER SAMPLING AND CBAM ON TEST SCENARIOS

no	***Under-sampling***		***CBAM***		***% Acc.***	***% Prec.***	***% Rec.***	***% F1***
	w/	**wo/**	**w/**	**wo/**				
1		·		·	95.06	95.32	92.80	92.78
2	·			·	96.68	96.68	96.68	96.68
3		·	·		97.56	97.61	96.89	97.20
4	·		·		97.74	97.74	97.74	97.74

* w/ = with
** wo/ = without

Furthermore, the effect of CBAM implementation can be seen in the heatmap below where the model identifies lung area in the chest to be more important than the other area captured in the x-ray. Feature extraction of the data is more focused on the lungs rather than the area of the heart, rib cage, etc. where COVID-19 or pneumonia mass can be found.

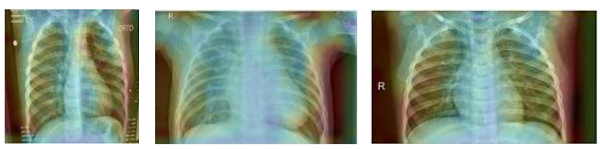


Fig. 13a Heatmap of the model without CBAM Implementation

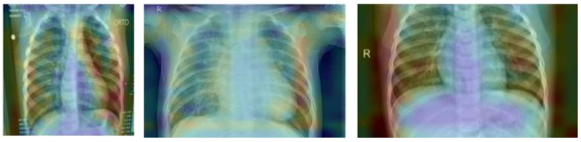


Fig. 13b Heatmap of the model with CBAM Implementation

Compared to the previous experiment mentioned on Section I, our model achieves the best performance across various metrics.

TABLE IIIX
RESULT COMPARISON

Reference	Result		
	% Accuracy	% Recall	% F1-score
[10]	94	91.3	-
[11]	95.72	-	-
[12]	-	-	96.87
[13]	87.02	-	-
This Research	**97.74**	**97.74**	**97.74**

IV. CONCLUSIONS

This study implements the use of CNN and CBAM and the under-sampling method for classifying CXR into three classes: COVID-19, pneumonia, and healthy lungs. Our novel model achieved 97.74% in all accuracy, precision, recall, and f1-score metric as the final result. Due to overfitting and lack of resources, our model uses fewer layers than its source from the DarkCOVIDNet model and has proven to generalize the specified problem.

Experiments that have been conducted discover that the use of CBAM and under sampling method are effective in boosting model performance, either when used separately or combined. CBAM were proven useful to increase model performance as it can find important features and their location within a spatial dimension. On the other hand, under sampling were used to make sure the model generalized all class equally, satisfying the model's generalization capability.

Furthermore, this research has discovered that using data from multiple resources can be a disadvantage as different data characteristics may hinder the model from extracting recognizable patterns within the data. Further research using a larger dataset may also be beneficial to this subject and radiology experts' evaluation on this research and the methods used. A comparative study using various models from similar research might also be necessary to create a better-performing model.

REFERENCES

[1] W. H. Organization, "WHO coronavirus (COVID-19) dashboard." 2022. [Online]. Available: https://covid19.who.int/

[2] W. H. Organization, "Listings of WHO's response to COVID-19." 2020. [Online]. Available: https://www.who.int/news/item/29-06-2020-covidtimeline

[3] W. H. Organization, "Coronavirus." 2020. [Online]. Available: https://www.who.int/health-topics/coronavirus#tab=tab_3

[4] Y. Shi *et al.*, "An overview of COVID-19," *Journal of Zhejiang University-SCIENCE B*, vol. 21, no. 5, pp. 343–360, May 2020, doi: 10.1631/jzus.B2000083.

[5] B. Singh, B. Datta, A. Ashish, and G. Dutta, "A comprehensive review on current COVID-19 detection methods: From lab care to point of care diagnosis," *Sensors International*, vol. 2, p. 100119, 2021, doi: 10.1016/j.sintl.2021.100119.

[6] Y. Fang *et al.*, "Sensitivity of Chest CT for COVID-19: Comparison to RT-PCR," *Radiology*, vol. 296, no. 2, pp. E115–E117, Sep. 2020, doi: 10.1148/radiol.2020200432.

[7] M. Dramé *et al.*, "Should RT-PCR be considered a gold standard in the diagnosis of COVID-19?," *Journal of Medical Virology*, vol. 92, no. 11. John Wiley and Sons Inc, pp. 2312–2313, November 1, 2020. doi: 10.1002/jmv.25996.

[8] W. Guan *et al.*, "Clinical Characteristics of Coronavirus Disease 2019 in China," *New England Journal of Medicine*, vol. 382, no. 18, pp. 1708–1720, Sep. 2020, doi: 10.1056/NEJMoa2002032.

[9] C. Huang *et al.*, "Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China," *The Lancet*, vol. 395, no. 10223, pp. 497–506, Sep. 2020, doi: 10.1016/S0140-6736(20)30183-5.

[10] M.-Y. Ng *et al.*, "Imaging Profile of the COVID-19 Infection: Radiologic Findings and Literature Review," *Radiol Cardiothorac Imaging*, vol. 2, no. 1, p. e200034, Sep. 2020, doi: 10.1148/ryct.2020200034.

[11] S. Sathi *et al.*, "Role of Chest X-Ray in Coronavirus Disease and Correlation of Radiological Features with Clinical Outcomes in Indian Patients," *Canadian Journal of Infectious Diseases and Medical Microbiology*, vol. 2021, pp. 1–8, Oct. 2021, doi: 10.1155/2021/6326947.

[12] N. H. Service, "Pneumonia." 2017. [Online]. Available: https://www.nhs.uk/conditions/pneumonia/

[13] T. Rahman *et al.*, "Transfer Learning with Deep Convolutional Neural Network (CNN) for Pneumonia Detection Using Chest X-ray," *Applied Sciences*, vol. 10, no. 9, p. 3233, Sep. 2020, doi: 10.3390/app10093233.

[14] A. U. Ibrahim, M. Ozsoz, S. Serte, F. Al-Turjman, and P. S. Yakoi, "Pneumonia Classification Using Deep Learning from Chest X-ray Images During COVID-19," *Cognit Comput*, Sep. 2021, doi: 10.1007/s12559-020-09787-5.

[15] K. Hammoudi *et al.*, "Deep Learning on Chest X-ray Images to Detect and Evaluate Pneumonia Cases at the Era of COVID-19," *J Med Syst*, vol. 45, no. 7, p. 75, Sep. 2021, doi: 10.1007/s10916-021-01745-4.

[16] S.-H. Wang, S. L. Fernandes, Z. Zhu, and Y.-D. Zhang, "AVNC: Attention-Based VGG-Style Network for COVID-19 Diagnosis by CBAM," *IEEE Sens J*, vol. 22, no. 18, pp. 17431–17438, Sep. 2022, doi: 10.1109/JSEN.2021.3062442.

[17] T. Ozturk, M. Talo, E. A. Yildirim, U. B. Baloglu, O. Yildirim, and U. R. Acharya, "Automated detection of COVID-19 cases using deep neural networks with X-ray images," *Comput Biol Med*, vol. 121, p. 103792, Sep. 2020, doi: 10.1016/j.compbiomed.2020.103792.

[18] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, "CBAM: Convolutional Block Attention Module." pp. 3–19, 2018. doi: 10.1007/978-3-030-01234-2_1.

[19] L. Wang, Z. Q. Lin, and A. Wong, "COVID-Net: a tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images," *Sci Rep*, vol. 10, no. 1, p. 19549, Sep. 2020, doi: 10.1038/s41598-020-76550-z.

[20] S. García, J. Luengo, and F. Herrera, "Intelligent Systems Reference Library 72 Data Preprocessing in Data Mining." [Online]. Available: http://www.springer.com/series/8578.

[21] D. Chicco, "Ten quick tips for machine learning in computational biology," *BioData Min*, vol. 10, no. 1, p. 35, Sep. 2017, doi: 10.1186/s13040-017-0155-3.

[22] Y.-X. Tang *et al.*, "Automated abnormality classification of chest radiographs using deep convolutional neural networks," *NPJ Digit Med*, vol. 3, no. 1, pp. 1–8, 2020.

[23] H. Kaur, H. S. Pannu, and A. K. Malhi, "A Systematic Review on Imbalanced Data Challenges in Machine Learning," *ACM Comput Surv*, vol. 52, no. 4, pp. 1–36, Jul. 2020, doi: 10.1145/3343440.

[24] J. M. Johnson and T. M. Khoshgoftaar, "Survey on deep learning with class imbalance," *J Big Data*, vol. 6, no. 1, p. 27, Dec. 2019, doi: 10.1186/s40537-019-0192-5.

[25] H. M. Bui, M. Lech, E. Cheng, K. Neville, and I. S. Burnett, "Using grayscale images for object recognition with convolutional-recursive neural network," in *2016 IEEE Sixth International Conference on Communications and Electronics (ICCE)*, Sep. 2016, pp. 321–325. doi: 10.1109/CCE.2016.7562656.

[26] N. Srivastava, G. Hinton, A. Krizhevsky, and R. Salakhutdinov, "Dropout: A Simple Way to Prevent Neural Networks from Overfitting," 2014.

[27] D. Berrar, "Cross-Validation," *Encyclopedia of Bioinformatics and Computational Biology*. Elsevier, pp. 542–545, 2019. doi: 10.1016/B978-0-12-809633-8.20349-X.

[28] D. Anguita, A. Ghio, S. Ridella, and D. Sterpi, "K-Fold Cross Validation for Error Rate Estimate in Support Vector Machines.," in *DMIN*, 2009, pp. 291–297.
[29] D. Laflly, Toward an Open Resource Using Services Cloud Computing for Environmental Data, vol. 1. Wiley, 2020.
[30] G. James, D. Witten, T. Hastie, and R. Tibshirani, "An Introduction to Statistical Learning with Applications in R Second Edition," 2021.