Journal of Robotics and Control (JRC)
Volume 5, Issue 3, 2024
ISSN: 2715-5072, DOI: 10.18196/jrc.v5i3.21494

Effectiveness of CNN Architectures and SMOTE to Overcome Imbalanced X-Ray Data in Childhood Pneumonia Detection

Yuri Pamungkas [1*], Muhammad Rifqi Nur Ramadani [2], Edwin Nugroho Njoto [3]
[1] Department of Medical Technology, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia
[2] Department of Biology, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia
[3] Department of Medicine, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia
Email: [1] yuri@its.ac.id, [2] rifqinurramadani05@gmail.com, [3] edwin_n_njoto@its.ac.id
*Corresponding Author

***Abstract*—Pneumonia is a disease that causes high mortality worldwide in children and adults. Pneumonia is caused by swelling of the lungs, and to ensure that the lungs are swollen, a chest X-ray can be done. The doctor will then analyze the X-ray results. However, doctors sometimes have difficulty confirming pneumonia from the results of chest X-ray observations. Therefore, we propose the combination of SMOTE and several CNN architectures be implemented in a chest X-ray image-based pneumonia detection system to help the process of diagnosing pneumonia quickly and accurately. The chest X-ray data used in this study were obtained from the Kermany dataset (5216 images). Several stages of pre-processing (grayscaling and normalization) and data augmentation (shifting, zooming, and adjusting the brightness) are carried out before deep learning is carried out. It ensures that the input data for deep learning is not mixed with noise and is according to needs. Then, the output data from the augmentation results are used as input for several CNN deep learning architectures. The augmented data will also utilize SMOTE to overcome data class disparities before entering the CNN algorithm. Based on the test results, the VGG16 architecture shows the best level of performance compared to other architectures. In system testing using SMOTE+CNN Architectures (VGG16, VGG19, Xception, Inception-ResNet v2, and DenseNet 201), the optimum accuracy level reached 93.75%, 89.10%, 91.67%, 86.54% and 91.99% respectively. SMOTE provides a performance increase of up to 4% for all CNN architectures used in predicting pneumonia.**

Keywords—Pneumonia; X-Ray Images; SMOTE; CNN Architectures.

I. INTRODUCTION

In 2019, pneumonia resulted in the death of 800,000 children worldwide. So that every 39 seconds, a child dies from this disease. Most children with pneumonia die under two, and nearly 153,000 child deaths occur in the first month of life [1]. Apart from affecting children, pneumonia can also affect adults with more or less the same symptoms/effect [2]. Pneumonia is swelling that occurs in the lungs and is usually caused by an infection [3]. People suffering from Pneumonia can experience various symptoms, from mild to severe [4]. These symptoms include shortness of breath, fever to chills, and coughing up phlegm (yellow or green sputum) [5]. So that Pneumonia can also be referred to as a wet lung disease [6]. When an infection occurs, the air sacs/alveoli (in one or both parts of the lung) will become inflamed and filled with pus. The pus will cause pneumonia sufferers to have difficulty breathing [7]. Fungal, viral and bacterial infections can cause Pneumonia. One type of virus that can cause Pneumonia is the SARS Corona Virus [8]. In addition, the emergence of pneumonia can also be accompanied by other lung diseases, such as tuberculosis (TB) [9]. Respiratory Syncytial Virus, SARS Corona Virus, and Influenza Virus are several viruses that can cause pneumonia in humans [10]. In contrast, Streptococcus Pneumonia is a type of bacteria that causes pneumonia. As previously explained, the symptoms of pneumonia sufferers can vary [11], such as chest pain, nausea/vomiting, no appetite, cough accompanied by phlegm, shortness of breath, fever and chills so that the treatment of pneumonia is also adjusted to the patient's condition and the severity of the disease [12]. For example, if a bacterial infection causes pneumonia, it is enough to give antibiotics [13]. The doctor will also give other medicines to reduce fever, relieve coughs, and relieve pain in patients. Vaccinating and maintaining personal hygiene can prevent pneumonia [14]. Moreover, not having direct contact with pneumonia patients can prevent other people from getting pneumonia. Considering that pneumonia can be a severe condition and requires immediate medical attention, understanding the symptoms, risk factors, and efforts to prevent and treat pneumonia is crucial (especially if the individual is in a high-risk group) [15].

One of the ways used to diagnose pneumonia is with a chest X-ray. A chest X-ray aims to determine the condition of the lungs and the area of the lung that is infected or inflamed [16]. However, there are some limitations in using chest X-rays for pneumonia detection. The limitations include problems with the availability of radiographic facilities in most health services in the developing world, varying amounts and quality of clinical information available, difficulty in interpreting chest radiographs from young children, the inability of radiographs to reveal early changes in pneumonia, and the inability of chest radiography to reliably distinguish viral from bacterial pneumonia [84]. Moreover, detecting lung consolidation using chest X-Ray becomes challenging when it measures less than 1.0 cm [85]. This difficulty arises from the nature of chest radiographs, which are two-dimensional representations combining normal and abnormal lobes, thereby complicating the identification of small lesions [86]. Additionally, X-ray

image data of pneumonia sufferers sometimes has weaknesses, such as unbalanced data. Data imbalance in the X-ray image dataset of pneumonia patients can significantly impact pneumonia detection results. When the number of images showing pneumonia (positive class) is much less than the images that do not show pneumonia (negative class), the detection model tends to prioritize the majority class and may ignore the minority class. As a result, the model may be less sensitive to actual pneumonia cases, resulting in a high error rate, especially when detecting positive cases. Data imbalance can also cause the model to provide highly false pessimistic predictions, where the model considers X-ray images showing signs of pneumonia as unfavorable. It can hurt patient management, resulting in incorrect diagnoses or delays in necessary treatment. Therefore, it is essential to address data imbalance using techniques such as class weighting, data augmentation, resampling, threshold tuning, or remodeling to improve model performance in detecting pneumonia cases, especially those that are a minority class in the dataset. Another issue related to this disease is that the characteristics indicating its presence often overlap with symptoms of other diseases. This can make it challenging for radiologists to diagnose it [87]. These limitations can be minimized by the development of computer-aided diagnosis systems such as the application of deep learning techniques in radiology [88]. With the help of a detection system based on deep learning, the doctor's decision-making process in diagnosing pneumonia is assisted or easier.

Research conducted by Bharati et al. [17] succeeded in proposing a lung disease detection system based on X-ray images using the Hybrid Deep Learning method. Several lung diseases, such as Pneumonia, Tuberculosis, Asthma, and Fibrosis, were successfully differentiated using medical image processing and deep learning for their classification. Patient X-ray data was obtained from the Kaggle repository and classified according to the type of disease the patient suffers from with several classifiers such as Vanilla Neural Network, CNN, Visual Neural Network, and Capsule Neural Network. The results obtained an accuracy value of 73% (for classification using Vanilla Neural Network), 69% (for classification using CNN), 69.5% (for classification using Visual Neural Network), and 63.8% (for classification using Capsule Neural Network). Research conducted by Jain et al. [18] also succeeded in creating a pneumonia detection system based on chest X-ray images using the CNN and Transfer Learning methods. Chest X-ray data are classified into two categories, infected with pneumonia and non-pneumonia, with several variations related to the convolutional layers, parameters, and hyperparameters used. Based on the system that has been made, accuracy values are obtained of 87.28% (using VGG16), 88.46% (using VGG19), 77.56% (using ResNet50), and 70.99% (using Inception v3). In addition, Ayan et al. [19] research also succeeded in creating a pneumonia diagnosis system using the Deep Learning method. The CNN method with the Xception and VGG16 models was chosen to classify patient chest X-ray data. In the data training process, fine-tuning and transfer learning are selected. Based on the system that has been created, an accuracy rate of 87% is obtained for classification using VGG16 and 82% for classification using the Xception network. Based on the background of the problem and previous research related to pneumonia detection, this research will propose a chest X-ray-based pneumonia detection system using several CNN models. In addition, the SMOTE technique will be combined with CNN to overcome the inequality in data classes for healthy patients and pneumonia sufferers. Chest X-rays were obtained from the Kermany Dataset (Kaggle repository), which contains a data set of X-rays of children aged 1 to 5 years [20], which everyone can access. Meanwhile, the CNN models used in this research include Xception, Inception-RestNet v2, VGG16, VGG19, and DenseNet201, considered the best-performance architectures. With several CNN models at once, it is hoped that an increase in the optimal evaluation matrix value of the model can be obtained, especially regarding the training process and image data classification.

II. METHODOLOGY

In this study, several steps were taken to predict pneumonia: data acquisition, pre-processing (data resize, rotating, horizontal flip, and augmentation), implementation SMOTE and class weighting, training & data testing based on the CNN model, and analysis of the results. For more details, here is a block diagram as in Fig. 1 related to the steps carried out in the Pneumonia prediction process using several types of Convolutional Neural Network models/architectures.

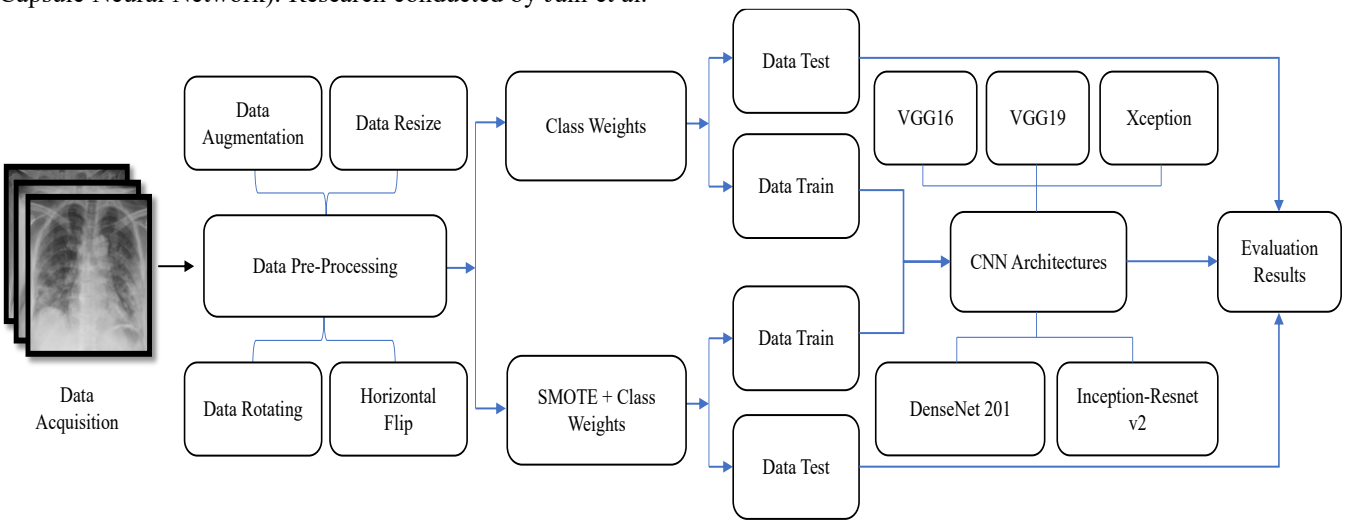

Fig. 1. Methodology research

A. Data Acquisition

This study used a lung X-ray image dataset from the Kermany Dataset (which can be obtained from the Internet) [20]. There are 5216 X-ray images used as training data for the CNN algorithm-based Pneumonia prediction process. From these data, 3875 X-rays were from patients with pneumonia, and 1341 X-rays were from people whose lungs were healthy. Meanwhile, there are 624 X-rays of the lungs for testing data, with details of 390 X-rays from Pneumonia sufferers and 234 X-rays from healthy people. The Fig. 2 is an example of X-ray images of Pneumonia patients and healthy people.

From these data, 3875 X-rays were from patients with pneumonia, and 1341 X-rays were from people whose lungs were healthy. Meanwhile, there are 624 X-rays of the lungs for testing data, with details of 390 X-rays from Pneumonia sufferers and 234 X-rays from healthy people. The following is an example of X-ray images of Pneumonia patients and healthy people.

B. Data Pre-Processing

The pre-processing stage consists of grayscaling and normalization of X-ray images. Grayscaling converts an image to a grey colour range, where each pixel has only a single brightness level without colour information [21]. In a grey image, each pixel is represented by a single brightness channel whose value ranges from 0 (represents black) to 255 (represents white) [22]. Grayscaling is used to reduce data dimensions, increase processing efficiency, and focus attention on image brightness information, especially in medical image processing and image analysis [23]. In addition, grayscale allows better focus on anatomical structures and abnormalities that may be present in the X-ray image [24]. In this study, the primary purpose of grayscaling is to make X-ray images of the lungs into a simple form (one-layer matrix). After the grayscaling process, the next step is image normalisation.

Data normalisation in X-ray image processing is changing the range of pixel values in an image to a more appropriate or standard scale [25]. Normalisation aims to change the range of pixel values into certain specified intervals to make data easier to process or analyse [26]. Normalization can help optimize the visibility of anatomical and pathological structures in the context of X-ray images [27]. A suitable intensity range ensures that all critical information in the image can be seen clearly without losing detail or relevant information [28]. With proper normalization, X-ray images can be prepared for advanced analysis processes, such as segmentation, lesion detection, or machine learning techniques [29]. In this study, the Min-Max Normalization method was used. This method changes the image's pixel values range to a specific range of values (from -1 to 1). The mechanism is to reduce the pixel value with the minimum pixel value in the image and divide it by the difference between the maximum and minimum pixel values [30]. The result is an image whose pixel values are within a predetermined range. The following is the Min-Max Normalization formula.

$$X_{norm} = \frac{x' - min(x)}{max(x) - min\ (x)} \times y \quad (1)$$

$$y = \left(new_{max}(x) - new_{min}(x)\right) + new_{max}(x) \quad (2)$$

where: x = the data attribute; $min\ (x)$ & $max\ (x)$ = the minimum and maximum absolute values of x; x' = the old value of each input data; $new_{min}\ (x)$ & $new_{max}\ (x)$ = the minimum and maximum values of the range.

Augmentation is a method in image processing that transforms or modifies pre-existing data into a new or novel form [31]. In medical image processing, data augmentation creates more data variations [32]. This wide variety of data will increase the ability of machine learning models to recognize variations in X-ray images of the lungs [33]. In this study, image data augmentation was carried out by shifting, zooming, and adjusting the brightness of existing lung X-ray data. The shifting function is to position the X-ray image of the lungs in a predetermined direction so that the machine-learning process can run optimally [34]. Zoom and brightness settings aim to adjust the clarity and contrast of the image [35]. In this study, shifting was carried out at 0.05 (width) and 0.05 (height). Meanwhile, the zoom is set at a value of 0.05, and the brightness level is in the range of 0.95 to 1.05.

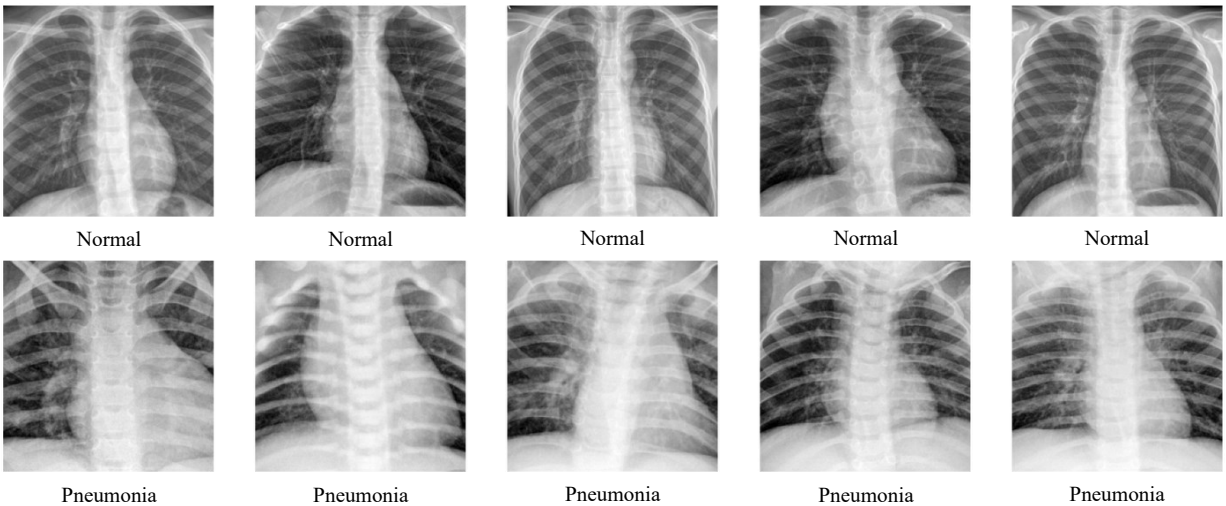

Fig. 2. X-ray image of pneumonia patient and healthy children

C. Synthetic Minority Oversampling Technique (SMOTE)

SMOTE is a technique used in data processing to handle class imbalance problems. Class imbalance occurs when the number of instances in one class is much less than that of other classes in the dataset, leading to biased classification model performance towards the majority class and ignoring the minority class. The main goal of SMOTE is to improve minority class representation by creating new synthetic samples based on existing samples in the minority class, thereby creating a more balanced distribution of classes in the dataset. The working mechanism of SMOTE involves creating new synthetic samples by taking existing minority samples and creating new samples among the nearest neighbors in the feature space. This process is carried out by randomly selecting a minority sample and finding its nearest neighbors. After that, one of those neighbors is randomly selected, and a new synthetic sample is created by combining the features of the initial minority sample and the selected neighbor along a line connecting the two. In this way, SMOTE not only expands existing minority samples but also creates new variations in the dataset that can help improve the classification model's ability to understand and differentiate minority classes. In use, SMOTE has proven effective in improving the performance of classification models, especially in cases of significant class imbalance. By improving the representation of minority classes, SMOTE helps reduce possible bias in classification models and improves the model's ability to recognize minority classes accurately. However, using SMOTE also requires careful consideration regarding the potential for overfitting in the resulting dataset and needs to be considered together with other techniques in dealing with class imbalance. The Fig. 3 is an illustration of the SMOTE.

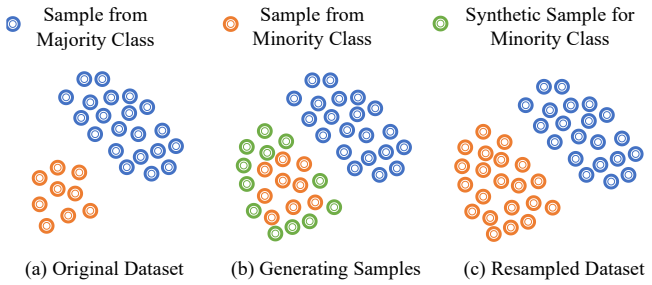

Fig. 3. Illustration of SMOTE

In general, the function of this algorithm is written as "*SMOTE (X, N, k)*," where X is the minority data, N is the percentage of the number of instances to be created, and k is the number of closest instances. The k value can be found using the Euclidean distance formula.

$$E\ Dist = \sqrt{(X_1 - Y_1)^2 + (X_2 - Y_2)^2 + \cdots + (X_n - Y_n)^2} \quad (3)$$

Then, data replication is carried out from the nearest instance using the following equation.

$$X_{syn} = X_i + (X_{knn} - X_i) \times \sigma \quad (4)$$

Where X_{syn} is the synthesized data from the replication, X_i is the i^{th} data in the minority class, and X_{knn} is the minority class data with the closest distance to the X_i data. Meanwhile, σ is a random number between 0-1.

If the numerical data is measured by its proximity to the Euclidean distance, then categorical data is measured based on its mode value. Calculating the distance between minor classes whose variables are on a categorical scale is done using the Value Difference Metric (VDM) formula.

$$\Delta\,(X,Y) = w_x w_y \sum_{i=1}^{N} \delta\,(x_i, y_i)^r \quad (4)$$

Where, $\Delta\,(X,Y)$ is the distance between X and Y. w_x, w_y is weights of observation. N is the number of explanatory variables. R has a value of 1 (Manhattan distance) or 2 (Euclidean distance). $\delta(x_i, y_i)^r$ is the distance between categories.

$$\delta\,(V_1, V_2) = \sum_{i=1}^{n} \left| \frac{C_{11}}{C_1} - \frac{C_{21}}{C_2} \right|^k \quad (5)$$

Where, $\delta\,(V_1, V_2)$ is the distance between values of V_1 and V_2. C_{1i} is the number of V_1 included class i. C_{2i} is the number of V_2 included class i. i is the number of classes, $i = 1, 2, \ldots, n$. C_1 is the number of occurrences of value 1. C_2 is the number of occurrences of value 2. n is the number of categories. k is constant value (usually 1).

D. Prediction using CNN

This study applies several types of CNN architectures in the pneumonia prediction process. These architectures include VGG16, VGG19, Resnet v2, Xception, and DenseNet201. The following is a detailed description of some of the CNN architectures.

1) VGG16 Architecture

VGG16 is a CNN architecture often used for image processing and recognition [36]. As its name implies, VGG16 has 16 layers, such as a convolution layer, an activation layer (the ReLU function), and a pooling layer for processing and image recognition. VGG16 usually utilizes small convolutions (3×3 size) to gain an increased understanding of image features [37]. A pooling layer is applied to reduce spatial dimensions and retain the extracted essential features. Combining these layers forms a deep structure to capture the hierarchical representation of the image [38]. This architecture is also quite effective for solving image processing tasks such as image classification using the ImageNet dataset [39]. In image processing based on the ImageNet dataset, the resulting model performs well in recognizing various existing objects [40]. Besides these advantages, the VGG16 architecture has several drawbacks, such as the long training and model testing time [41]. In addition, this architecture tends to be less efficient than the newer CNN architectures, such as Resnet and Inception. Fig. 4 is a block diagram of the VGG16 architecture.

The following is a detailed explanation regarding several components of the VGG16 architecture:

- Input: Based on the above architecture, VGG16 has an input image of 224×224 [42]. So to obtain this size, the

input image will be cropped starting from the centre of the image with a size of 224×224 pixels.

- Convolution Layer: This layer is responsible for extracting visual features from the input data [43]. VGG16 uses a convolution filter with a minor receptive field 3×3. In addition, a 1×1 convolution filter is used as input for the linear transformation [44].
- ReLU Activation Function: The activation function is commonly used in artificial neural networks, especially in convolution and hidden layers [45]. The ReLU function has superficial mathematical characteristics. For any input value greater than zero, this function will return the value itself. Meanwhile, when the input value is less than or equal to zero, this function will return a zero value [46]. The ReLU activation function can shorten the model training time in the CNN architecture [47].
- Hidden Layer: In neural networks, hidden layers are computational locations when the neural network processes input data to understand patterns or features in that data [48]. In the VGG16 architecture, ReLU is often used in the hidden layer to shorten the training process and reduce memory allocation during computation [49].
- Layer Pooling: Layer pooling is one type of layer in a neural network to reduce the spatial dimension of the feature representation generated by the previous convolution layer [50]. The main goal of layer pooling is to reduce computational complexity and avoid overfitting by reducing the number of parameters the model must learn [51]. This layer also helps maintain translation invariance against small shifts in the image so that the model is more resistant to variations in object position [52].

Fully connected layers: Fully connected layers are generally used at the end of a neural network to connect extracted feature representations from the previous layer to outputs suited to a specific task, such as classification or regression [53]. It is also often referred to as the output layer.

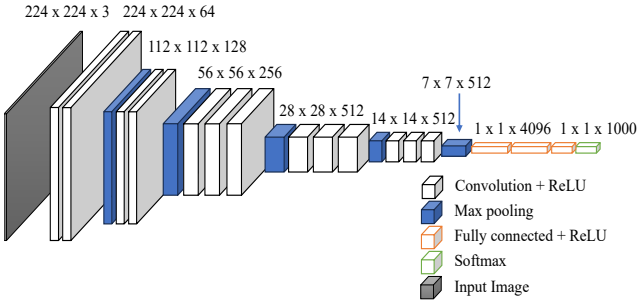

Fig. 4. VGG16 Architecture

2) *VGG19 Architecture*

VGG19 has a structure similar to VGG16 but with more convolution and fully connected layers [54]. Repetitive 3×3 convolution layers and 2×2 pooling layers form a deep structure that allows the model to capture increasingly complex feature hierarchies. Fully connected layers at the end of the architecture are responsible for generating class predictions [55]. VGG19 often uses pre-trained weights trained on large datasets, such as ImageNet, to leverage knowledge already gained in general image classification tasks. Adding this layer gives VGG19 a greater capacity to extract features from images better and more accurately. On the other hand, the computation time for data training is longer due to the addition of this layer [56]. The change in VGG19 architecture compared to VGG16 is that there are four additional convolution layers in VGG19, bringing the total convolution layers to 16. In addition, there are two additional fully connected layers before the output layer, which brings the total number of fully connected layers to 3 [57]. VGG19 maintains the philosophy of VGG16 in terms of simple structure and greater depth. This architecture is renowned for its ability to produce excellent and accurate feature representations of images. So, it is useful for object recognition and classification tasks [58]. Fig. 5 is a block diagram of the VGG19 architecture.

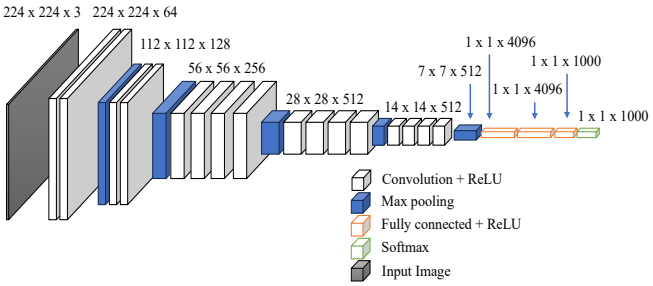

Fig. 5. VGG19 Architecture

3) *Xception Architecture*

Xception (Extreme Inception) is a convolutional neural network (CNN) architecture that develops the Inception architecture, initially introduced by the Google research team in the Inception-v3 workspace [59]. Xception takes the concept of Inception even further by replacing the standard convolution layer with a more extreme convolution operation, namely separable convolution. A separable convolution is an approach where spatial convolution (depthwise convolution) and channel convolution (pointwise convolution) are performed separately [60]. Spatial convolution is applied to each input channel separately, followed by pointwise convolution, which aims to combine the results of spatial convolution [61]. This approach reduces the number of parameters to learn and increases computational efficiency [62].

Xception aims to overcome some of the constraints in the original Inception architecture, mainly related to complex computations and the high number of parameters [63]. Using separable convolution, Xception can reduce computational complexity and speed up model training while maintaining or increasing performance in image recognition tasks [64]. The Xception architecture has proven successful in various image recognition tasks, including object recognition, image classification, and object segmentation [65]. Its success proves that understanding the features of a convolutional neural network can be achieved by optimizing the use of a more efficient convolution [66]. Fig. 6 is a block diagram of the Xception architecture.

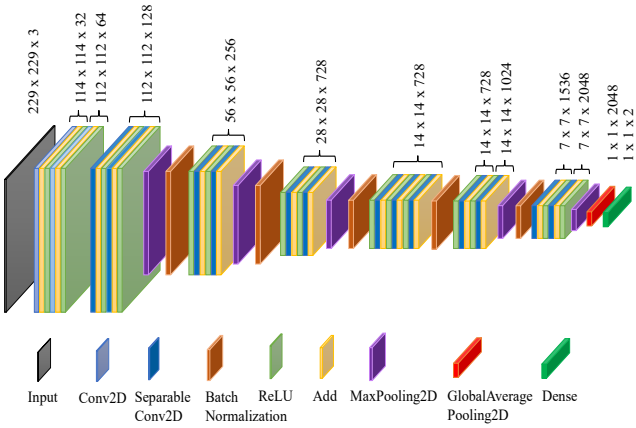

Fig. 6. Xception Architecture

4) Inception-ResNet Architecture

ResNet is a convolutional neural network architecture that has overcome the challenge of training intense networks by introducing the concept of residual blocks [67]. One of the challenges in training intense networks is the problem of vanishing gradients when doing iterative learning through multiple layers. The deeper the network layer, the more difficult it is for gradients (derivatives) to propagate back up through the initial layers [68]. It hinders the model's ability to learn and hinders performance gains. Residual blocks in ResNet have shortcut connections (also known as skip connections), allowing gradients to jump through multiple layers simultaneously [69]. In the residual block, the output of a layer is not only multiplied by the weight but also added by the original input before adjustment. It allows the original information to pass through without significant changes, which helps overcome gradient loss issues and enables intense network training [70].

The main advantage of ResNet is its ability to train intense networks (in some cases, more than 100 layers) more effectively and get better performance [71]. ResNet has performed excellently in various image recognition tasks, including object recognition, object segmentation, and object detection in competitions such as the ImageNet Large Scale Visual Recognition Challenge [72]. ResNet-v2 (Residual Network version 2) further develops the original ResNet architecture designed to fix several problems associated with training neural networks at scale [73]. One of the main changes in ResNet-v2 is using an "identity shortcut" operation, which allows the original signal to flow through the residual block without modification [74]. However, in ResNet-v2, the residual block focuses on preserving as much of the original information as possible by minimizing the transformations performed on the signal. It also helps in maintaining the gradient during training [75]. Fig. 7 is a block diagram of the Inception-Resnet v2 architecture.

5) DenseNet 201 Architecture

DenseNet-201 is a convolutional neural network architecture part of the DenseNet architecture family. Densely Connected Convolutional Networks (DenseNet) is a concept that combines layers in a network in a very tight and robust way, where each layer receives input from all previous layers in the chain [76]. DenseNet aims to address the problem of intense network training by preventing gradient loss issues and stimulating better feature learning [77]. Connecting each layer to the previous layer allows information to flow more efficiently through the network. In DenseNet-201, "201" refers to this architecture's total number of layers. It includes convolution, batch normalisation, activation, and pooling layers [78]. The more significant number of layers allows the model to learn complex feature representations from images but also requires more significant computational resources for training [79]. DenseNet-201 and other DenseNet architectures have noted excellent performance in various image recognition tasks, including image classification, object segmentation, and object detection [80]. The DenseNet architecture also has the advantage of more efficient parameter use compared to deeper architectures, such as ResNet [81].

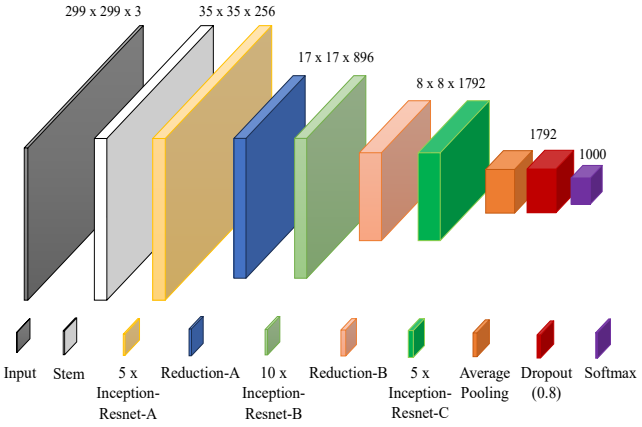

Fig. 7. Inception-Resnet v2 Architecture

III. RESULTS AND DISCUSSIONS

This study aims to obtain an optimal model in the prediction process of pneumonia based on X-ray images. The method used is CNN with several architectural variations such as VGG16, VGG19, Xception, Inception-Resnet v2, and DenseNet201. Several evaluation parameters as shown in Fig. 8 to Fig. 12, such as accuracy, sensitivity, specificity, precision, and F1-score of the confusion matrix, are analyzed to find out the optimal model that can be used for the prediction process. The confusion matrix is very useful in identifying patterns of prediction or classification errors made by the model to assist in taking action to improve model performance. Table I and Table II is the confusion matrix data from the model testing that was carried out in this study.

From the system testing results, evaluation parameter values are obtained, which can be used to describe the performance of each CNN architecture in the pneumonia prediction process. The accuracy parameter results show that the VGG16 model (91.51%) demonstrate better performance compared to other CNN models such as VGG19 (88.14%), Xception (89.42%), Inception-Resnet v2 (85.26%), and DenseNet 201 (90.87%). Interestingly, the output data from processing with SMOTE combined with CNN models indicate an increase in accuracy values for each model. The best performance in the SMOTE combination remains with the VGG16 model (93.75%), followed by other CNN models such as DenseNet 201 (91.99%), Xception (91.67%), Inception-Resnet v2 (86.54%), and VGG19 (89.10%). So, the

VGG16 model is suitable for understanding complex image features, which is useful in object recognition and other tasks. However, the VGG16 computation process takes longer than other CNN models, such as Xception, Inception-Resnet v2, and DenseNet 201.

TABLE I. MODEL TESTING RESULTS (CNN)

Parameter	Models				
	VGG16	VGG19	Xception	Inception-Resnet v2	DenseNet 201
TP	201	165	182	146	183
TN	370	385	376	386	384
FP	33	69	52	88	51
FN	20	5	14	4	6
Total	624				

TABLE II. MODEL TESTING RESULTS (CNN + SMOTE)

Parameter	Models				
	SMOTE + VGG16	SMOTE + VGG19	SMOTE + Xception	SMOTE + Inception-Resnet v2	SMOTE + DenseNet 201
TP	207	169	191	152	188
TN	378	387	381	388	386
FP	27	65	43	82	46
FN	12	3	9	2	4
Total	624				

where: TP = True Positive, TN = True Negative, FP = False Positive, and FN = False Negative. Based on these results, the values of several parameters, such as the CNN architectures' accuracy, sensitivity, precision, specificity, and F1-score can be presented asFig. 8 to Fig. 12.

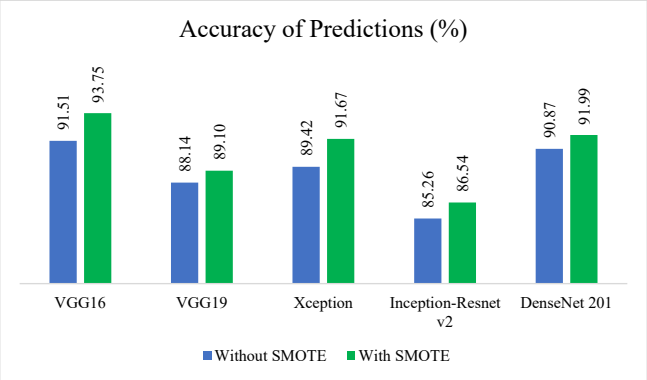

Fig. 8. The accuracy of each CNN architecture during the data testing

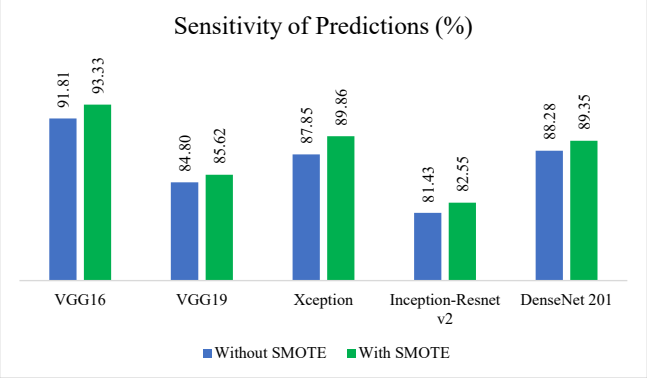

Fig. 9. The sensitivity of each CNN architecture during the data testing

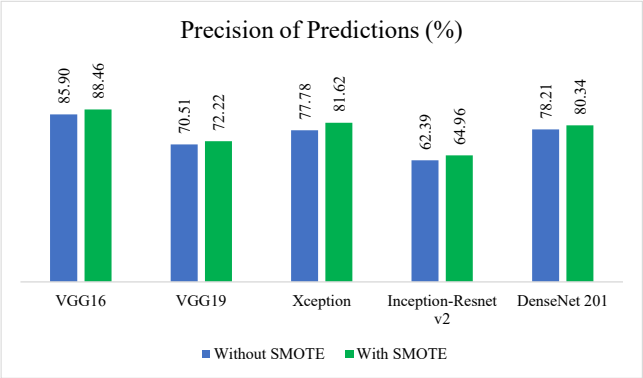

Fig. 10. The precision of each CNN architecture during the data testing

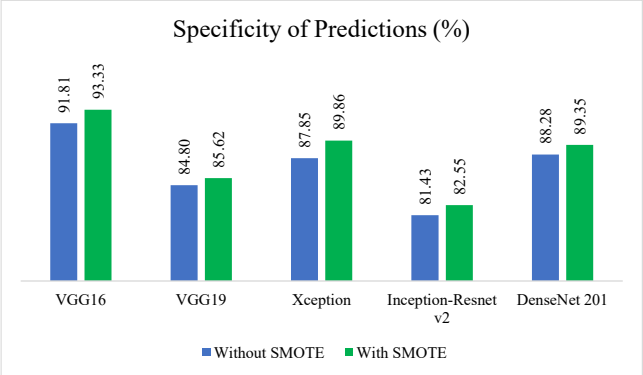

Fig. 11. The specificity of each CNN architecture during the data testing

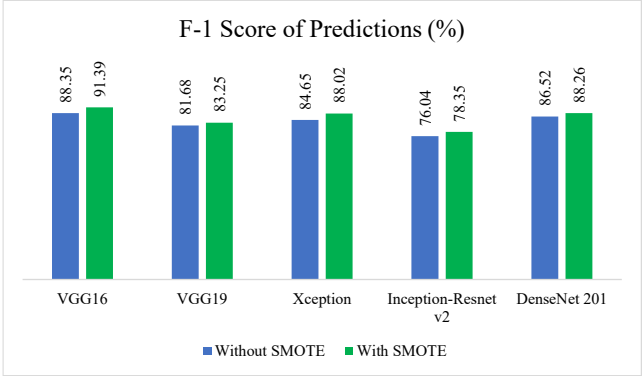

Fig. 12. The F1-score of each CNN architecture during the data testing

In addition to referring to the accuracy value, the researcher also observes the sensitivity value of each CNN architecture used. The highest sensitivity value is obtained when the testing system is conducted using the CNN Inception-Resnet v2 model in both non-SMOTE and SMOTE data (97.33% and 98.26% respectively). In contrast, VGG16 has the lowest sensitivity value compared to other CNN models (90.95%). The same pattern is also shown in the SMOTE data, with the highest sensitivity value obtained by the CNN Inception-Resnet v2 model (98.70%), while the lowest sensitivity value is held by the CNN VGG16 model (94.52%). This higher sensitivity value indicates that the model can better recognize data from the positive class, meaning that less positive data is missed or misclassified into the negative class [82]. However, an increase in sensitivity is often accompanied by an increase in the false positive rate (lower specificity), so it is important to find the right balance between sensitivity and specificity according to the needs of

the task and the impact of errors in each class [83]. Regarding the specificity value, The VGG16 model has the highest value, reaching 91.81% on non-SMOTE data and 93.33% on SMOTE data compared to other CNN models. In contrast, the Inception-Resnet v2 model has the lowest values on both non-SMOTE and SMOTE data, with 81.43% and 82.55% respectively. So in this study, the CNN VGG16 model tends to have a better balance of sensitivity and specificity both in non-SMOTE and SMOTE data with 90.95% and 91.81%, and 94.52% and 93.33% respectively, compared to Inception-Resnet v2 (balance of sensitivity and specificity value of 97.33% and 81.43% in non-SMOTE data, and 98.26% and 82.55% in SMOTE data).

Other evaluation parameters, such as precision and F1-score, also show the same trend for the VGG16 model. The parameter value for the non-SMOTE data, precision parameter reaches 85.90%, and the F1-score reaches 88.35%. Meanwhile, in the VGG19, Xception, Inception-Resnet v2, and DenseNet 201 models, the precision values were 70.51%, 77.78%, 62.39%, and 78.21%, respectively. Additionally, the F1 scores were 81.68%, 84.65%, 76.04%, and 86.52%, respectively, for VGG19, Xception, Inception-Resnet v2, and DenseNet 201. These results also have the same pattern with the SMOTE data applied on the CNN models with VGG16 having the highest values for precision and F-1 score of 88.46% and 91.39% respectively, compared with the other models. Researchers believe that the high precision and F1-score of VGG16 depends on the X-ray dataset used in research related to the detection of pneumonia. Not only that, the number of classes in the dataset (Normal and Pneumonia) and the threshold used to make classification decisions. These metrics will vary for each dataset and scenario.

From the generated data, it is found that the CNN VGG16 model exhibits the best performance among all tested CNN models against the Kermany dataset used in this study. We also tested data imbalance using the SMOTE method to generate synthetic data. The results showed that the use of the SMOTE method resulted in an increase of 1% to 3% in each testing parameter. Interestingly, VGG16 consistently remains the best-performing model in both scenarios (non-SMOTE and SMOTE datasets). VGG16 has been widely used in developing classification algorithms due to its applicability without limitations for small steps and smaller window sizes. It can embed 16 layers deep and perform better on larger datasets [89]. Moreover, VGG16 paired with a Neural Network (NN) classifier shows better results than VGG16 paired with other types of classifiers such as Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Random Forest (RF), and Naïve Bayes (NB), on the Kaggle [89] dataset. Interestingly, the same thing was found in this research that VGG16 has a high accuracy value compared to several CNN architectural methods. In several recent studies, VGG16 was seen to have higher accuracy scores compared to ResNet50v2+XceptionNet [90], ConvNet [91], and GoogleNet+ResNet+DensNet121 [92]. In line with these reports, our data shows that the F-1 score of the CNN model with VGG16 on the Kermany dataset is the highest compared to several other CNN architecture models used in this study (VGG19, Xception, Inception-Resnet v2, and DenseNet 201). Because the F-1 score determines the harmonic mean of precision and recall, VGG16 can be considered the model that performs best among the other models used. The CNN with VGG16 model achieved F-1 scores of 0.88 and 0.91 (for non-SMOTE and SMOTE data), which are the highest for the respective datasets (non-SMOTE and SMOTE datasets).

In this research, combining Convolutional Neural Networks (CNN) with Synthetic Minority Over-sampling Technique (SMOTE) can increase the effectiveness in predicting pneumonia based on X-ray images. CNNs have proven effective in learning complex features from medical images, including X-ray images, allowing them to identify patterns that may be difficult for conventional models to detect. However, CNN performance can be affected when data imbalance is faced, where the number of pneumonia images is minimal compared to non-pneumonia images. SMOTE is an oversampling method that synthetically produces new samples in the minority class by expanding the feature space. By applying SMOTE, we can increase the number of pneumonia samples in the dataset without duplicating the data, thereby reducing the effect of data imbalance. Thus, using a SMOTE-enhanced CNN gives the model a greater chance of learning relevant patterns from minority classes, improving its ability to detect actual pneumonia cases. This combination significantly increases the effectiveness of predicting pneumonia from X-ray images by reducing the error rate and increasing the sensitivity to positive cases.

IV. CONCLUSIONS

In this study, researchers implemented the CNN method with several architectural variations to detect/predict pneumonia from X-ray images of the lungs (Kermany dataset). Several types of CNN architectures, such as VGG16, VGG19, Xception, Inception-Resnet v2, and DenseNet 201, have been applied in this study to obtain an optimal model for the detection of pneumonia. To measure the performance level of each architecture, researchers use several evaluation parameters such as accuracy, sensitivity, precision, specificity, and F-1 score. Based on the test results, the VGG16 architecture shows the best level of performance compared to other architectures. In testing the system using VGG16, the optimal accuracy level reached 91.51%, sensitivity 90.95%, precision 85.90%, specificity 91.81%, and F-1 score 88.35%. The accuracy rates of other CNN architectures, such as VGG19, Xception, Inception-Resnet v2, and DenseNet 201, reach 88.14%, 89.42%, 85.26%, and 90.87%, respectively. The high accuracy of the VGG16 architecture and other CNN architectures proves the model performs well in extracting important features from images for object recognition processes. Additionally, using SMOTE combined with the CNN architecture increased the matrix evaluation value. For example, the accuracy value of VGG16 increases to 93.75%. Meanwhile, for other CNN architectures such as VGG19, Xception, Inception-ResNet v2, and DenseNet 201, it is 89.10%, 91.67%, 86.54%, and 91.99%, respectively. It proves that the balance of data classes in X-ray image-based pneumonia prediction is crucial and influences the prediction results. In future research, researchers will modify some of the hyperparameters of VGG16 so that the computation can run faster and the required memory allocation can be more efficient. The use of

SMOTE will also be maintained in future research to overcome inequality in data classes and to see the response of the prediction matrix evaluation if the x-ray dataset used is different from the one currently used.

ACKNOWLEDGMENT

The authors would like to acknowledge the Department of Medical Technology, Institut Teknologi Sepuluh Nopember, for the facilities and support in this research. The authors also gratefully acknowledge financial support from the Institut Teknologi Sepuluh Nopember for this work, under project scheme of the Publication Writing and IPR Incentive Program (PPHKI) 2024.

REFERENCES

[1] K. G. Rögnvaldsson, A. Bjarnason, I. S. Ólafsdóttir, K. O. Helgason, A. Guðmundsson, and M. G. Prof, "Adults with symptoms of pneumonia: a prospective comparison of patients with and without infiltrates on chest radiography," *Clinical Microbiology and Infection*, vol. 29, no. 1, p. 108, 2022.

[2] S. Quispe, I. Torres, and P. Shiguihara, "A Survey of Deep Learning Techniques Based on Computed Tomography Images for Detection of Pneumonia," *Engineering Proceedings*, vol. 42, no. 1, p. 5, 2023.

[3] V. Hespanhol and C. Bárbara, "Pneumonia mortality, comorbidities matter?," *Pulmonology*, vol. 26, no. 3, pp. 123–129, May 2020.

[4] F. M. de Benedictis, E. Kerem, A. B. Chang, A. A. Colin, H. J. Zar, and A. Bush, "Complicated pneumonia in children," *The Lancet*, vol. 396, no. 10253, pp. 786–798, Sep. 2020.

[5] D. Marangu and H. J. Zarb, "Childhood pneumonia in low-and-middle-income countries: An update," *Paediatric Respiratory Reviews*, vol. 32, pp. 3–9, Nov. 2019.

[6] M. C. B. Godoy, H. R. Ferreira Dalla Pria, M. T. Truong, G. S. Shroff, and E. M. Marom, "Invasive Fungal Pneumonia in Immunocompromised Patients," *Radiologic Clinics of North America*, vol. 60, no. 3, pp. 497–506, May 2022.

[7] C. W. Lanks, A. I. Musani, and D. W. Hsia, "Community-acquired Pneumonia and Hospital-acquired Pneumonia," *Medical Clinics of North America*, vol. 103, no. 3, pp. 487–501, May 2019.

[8] A. Zizza *et al*., "Factors Associated with Pneumonia in Patients Hospitalized with COVID-19 and the Role of Vaccination," *Vaccines*, vol. 11, no. 8, pp. 1342–1342, Aug. 2023.

[9] Y.-N. Liu *et al*., "Infection and co-infection patterns of community-acquired pneumonia in patients of different ages in China from 2009 to 2020: a national surveillance study," *The Lancet Microbe*, vol. 4, no. 5, pp. e330–e339, May 2023.

[10] J. Fachrel, A. A. Pravitasari, I. N. Yulita, M. N. Ardhisasmita, and F. Indrayatna, "Enhancing an Imbalanced Lung Disease X-ray Image Classification with the CNN-LSTM Model," *Applied sciences*, vol. 13, no. 14, pp. 8227–8227, Jul. 2023.

[11] N. Miyashita, "Atypical pneumonia: Pathophysiology, diagnosis, and treatment," *Respiratory Investigation*, vol. 60, no. 1, Nov. 2021.

[12] G. B. Nair and M. S. Niederman, "Updates on community acquired pneumonia management in the ICU," *Pharmacology & Therapeutics*, vol. 217, p. 107663, Jan. 2021.

[13] Z. V. Rueda *et al*., "Etiology and the challenge of diagnostic testing of community-acquired pneumonia in children and adolescents," *BMC Pediatrics*, vol. 22, no. 1, Mar. 2022.

[14] H. Guo *et al*., "Differentiate Clinical Characteristics Between Viral Pneumonia and Mycoplasma pneumoniae and Nomograms for Predicting Mycoplasma pneumoniae: A Retrospective Study in Primary Hospitals," *PubMed*, vol. 42, no. 12, pp. 1035–1040, Oct. 2023.

[15] E. S. Tanzarella *et al*., "An Observational Study to Develop a Predictive Model for Bacterial Pneumonia Diagnosis in Severe COVID-19 Patients—C19-PNEUMOSCORE," *Journal of Clinical Medicine*, vol. 12, no. 14, pp. 4688–4688, Jul. 2023.

[16] N. Jiang, Q. Long, Y. Zheng, and Z. C. Gao, "Advances in epidemiology, etiology, and treatment of community-acquired pneumonia," *PubMed*, vol. 57, no. 1, pp. 91–99, Jan. 2023.

[17] S. Bharati, P. Podder, and M. R. H. Mondal, "Hybrid deep learning for detecting lung diseases from X-ray images," *Informatics in Medicine Unlocked*, vol. 20, p. 100391, 2020.

[18] R. Jain, P. Nagrath, G. Kataria, V. Sirish Kaushik, and D. Jude Hemanth, "Pneumonia detection in chest X-ray images using convolutional neural networks and transfer learning," *Measurement*, vol. 165, p. 108046, Dec. 2020.

[19] E. Ayan and H. M. Ünver, "Diagnosis of Pneumonia from Chest X-Ray Images Using Deep Learning," *2019 Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT)*, pp. 1-5, 2019.

[20] D. Kermany, K. Zhang, and M. Goldbaum, "Labeled Optical Coherence Tomography (OCT) and Chest X-Ray Images for Classification," *Mendeley data*, vol. 2, no. 2, p. 651, 2018.

[21] A. Singhal, R. Shukla, P. K. Kankar, S. Dubey, S. Singh, and R. B. Pachori, "Comparing the capabilities of transfer learning models to detect skin lesion in humans," *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, vol. 234, no. 10, pp. 1083–1093, Jul. 2020.

[22] E. Hsu, I. Malagaris, Y.-F. Kuo, R. Sultana, and K. Roberts, "Deep learning-based NLP data pipeline for EHR-scanned document information extraction," *JAMIA Open*, vol. 5, no. 2, Apr. 2022.

[23] A. D. Wibawa, Y. Pamungkas, M. I. Perdana, and R. Rachmatika, "Iris Grid Image Classification using Naive Bayes for Human Biometric System," *2022 1st International Conference on Information System & Information Technology (ICISIT)*, pp. 55-60, 2022.

[24] A. Giełczyk, A. Marciniak, M. Tarczewska, and Z. Lutowski, "Pre-processing methods in chest X-ray image classification," *PloS one*, vol. 17, no. 4, p. e0265949, Jan. 2022.

[25] T. Tsuji *et al*., "Classification of chest X-ray images by incorporation of medical domain knowledge into operation branch networks," *BMC Medical Imaging*, vol. 23, no. 1, May 2023.

[26] M. Chetoui, M. A. Akhloufi, El Mostafa Bouattane, J. Abdulnour, S. Roux, and C. Bernard, "Explainable COVID-19 Detection Based on Chest X-rays Using an End-to-End RegNet Architecture," *Viruses*, vol. 15, no. 6, pp. 1327–1327, Jun. 2023.

[27] S. Liu, T. Cai, X. Tang, Y. Zhang, and C. Wang, "COVID-19 diagnosis via chest X-ray image classification based on multiscale class residual attention," *Computers in Biology and Medicine*, vol. 149, p. 106065, Oct. 2022.

[28] F. Li, X. Lu, and J. Yuan, "MHA-CoroCapsule: Multi-Head Attention Routing-based Capsule Network for COVID-19 chest X-ray image classification," *IEEE Transactions on Medical Imaging*, pp. 1–1, 2021.

[29] D. Ji, Z. Zhang, Y. Zhao, and Q. Zhao, "Research on Classification of COVID-19 Chest X-Ray Image Modal Feature Fusion Based on Deep Learning," *Journal of Healthcare Engineering*, vol. 2021, p. e6799202, Aug. 2021.

[30] I. M. Baltruschat, H. Nickisch, M. Grass, T. Knopp, and A. Saalbach, "Comparison of Deep Learning Approaches for Multi-Label Chest X-Ray Classification," *Scientific Reports*, vol. 9, no. 1, Apr. 2019.

[31] P. C. Motta, P. C. Cortez, B. R. S. Silva, G. Yang, and V. H. C. d. Albuquerque, "Automatic COVID-19 and Common-Acquired Pneumonia Diagnosis Using Chest CT Scans," *Bioengineering*, vol. 10, no. 5, p. 529, Apr. 2023.

[32] X. Qi, L. G. Brown, D. J. Foran, J. Nosher, and I. Hacihaliloglu, "Chest X-ray image phase features for improved diagnosis of COVID-19 using convolutional neural network," *Int. J. Comput. Assisted Radiol. Surgery*, vol. 16, no. 2, pp. 197–206, Jan. 2021.

[33] H. Louati, A. Louati, S. Bechikh, F. Masmoudi, A. Aldaej, and E. Kariri, "Topology optimization search of deep convolution neural networks for CT and X-ray image classification," *BMC Med. Imag*., vol. 22, no. 1, Jul. 2022.

[34] B. Chen, J. Li, G. Lu, H. Yu, and D. Zhang, "Label Co-Occurrence Learning with Graph Convolutional Networks for Multi-Label Chest X-Ray Image Classification," *IEEE J. Biomed. Health Inform*., vol. 24, no. 8, pp. 2292–2302, Aug. 2020.

[35] N.-A.-A. Alam, M. Ahsan, M. A. Based, J. Haider, and M. Kowalski, "COVID-19 Detection from Chest X-ray Images Using Feature Fusion and Deep Learning," *Sensors*, vol. 21, no. 4, p. 1480, Feb. 2021.

[36] W. Bakasa and S. Viriri, "VGG16 Feature Extractor with Extreme Gradient Boost Classifier for Pancreas Cancer Prediction," *J. Imag*., vol. 9, no. 7, p. 138, Jul. 2023.

[37] A. Huang, L. Jiang, J. Zhang, and Q. Wang, "Attention-VGG16-UNet: a novel deep learning approach for automatic segmentation of the median nerve in ultrasound images," *Quantitative Imag. Medicine Surgery*, vol. 12, no. 6, p. 3138, 2021.

[38] A. Alshammari, "Construction of VGG16 Convolution Neural Network (VGG16 CNN) Classifier with NestNet-Based Segmentation Paradigm for Brain Metastasis Classification," *Sensors*, vol. 22, no. 20, p. 8076, Oct. 2022.

[39] W. Pora, N. Kasamsumran, K. Tharawatcharasart, R. Ampol, P. Siriyasatien, and N. Jariyapan, "Enhancement of VGG16 model with multi-view and spatial dropout for classification of mosquito vectors," *PLOS ONE*, vol. 18, no. 7, Jul. 2023.

[40] D. Pandiar, S. Choudhari, and R. Poothakulath Krishnan, "Application of InceptionV3, SqueezeNet, and VGG16 Convoluted Neural Networks in the Image Classification of Oral Squamous Cell Carcinoma: A Cross-Sectional Study," *Cureus*, vol. 15, no. 11, 2023.

[41] S. Tena, R. Hartanto, and I. Ardiyanto, "Content-Based Image Retrieval for Traditional Indonesian Woven Fabric Images Using a Modified Convolutional Neural Network Method," *J. Imag*., vol. 9, no. 8, p. 165, Aug. 2023.

[42] M. Subramanian, N. P. Lv, and V. E. Sathishkumar, "Hyperparameter Optimization for Transfer Learning of VGG16 for Disease Identification in Corn Leaves Using Bayesian Optimization," *Big Data*, vol. 10, no. 3, pp. 215-229, 2021.

[43] G. Suganeshwari, R. Balakumar, K. Karuppanan, S. B. Prathiba, S. Anbalagan, and G. Raja, "DTBV: A Deep Transfer-Based Bone Cancer Diagnosis System Using VGG16 Feature Extraction," *Diagnostics*, vol. 13, no. 4, p. 757, Feb. 2023.

[44] 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," *Multimedia Tools Appl*., pp. 1-18, 2023.

[45] S. Montaha *et al*., "BreastNet18: A High Accuracy Fine-Tuned VGG16 Model Evaluated Using Ablation Study for Diagnosing Breast Cancer from Enhanced Mammography Images," *Biology*, vol. 10, no. 12, p. 1347, Dec. 2021.

[46] R. Mohan, K. Ganapathy, and A.R, "Brain tumour classification of Magnetic resonance images using a novel CNN based Medical Image Analysis and Detection network in comparison with VGG16," *J. Population Therapeutics Clin. Pharmacol*., vol. 28, no. 2, Jan. 2022.

[47] L.-Y. Ye, X.-Y. Miao, W.-S. Cai, and W.-J. Xu, "Medical image diagnosis of prostate tumor based on PSP-Net+VGG16 deep learning network," *Computer Methods and Programs in Biomedicine*, p. 106770, Mar. 2022.

[48] S. Srinivasan *et al*., "A Framework of Faster CRNN and VGG16-Enhanced Region Proposal Network for Detection and Grade Classification of Knee RA," *Diagnostics*, vol. 13, no. 8, pp. 1385–1385, Apr. 2023.

[49] Z. Omiotek and A. Kotyra, "Flame Image Processing and Classification Using a Pre-Trained VGG16 Model in Combustion Diagnosis," *Sensors*, vol. 21, no. 2, p. 500, Jan. 2021.

[50] Y. Zhang and L. Meng, "Study on Identification Method of Pulmonary Nodules: Improved Random Walk Pulmonary Parenchyma Segmentation and Fusion Multi-Feature VGG16 Nodule Classification," *Frontiers in Oncology*, vol. 12, Mar. 2022.

[51] Y. Wu, Y. He, and Y. Wang, "Multi-Class Weed Recognition Using Hybrid CNN-SVM Classifier," *Sensors*, vol. 23, no. 16, pp. 7153–7153, Aug. 2023.

[52] L. Kong and J. Cheng, "Classification and detection of COVID-19 X-Ray images based on DenseNet and VGG16 feature fusion," *Biomedical Signal Processing and Control*, vol. 77, p. 103772, Aug. 2022.

[53] S. Altun, A. Alkan, and İ. Altun, "LSS-VGG16: Diagnosis of Lumbar Spinal Stenosis With Deep Learning," *Clinical Spine Surgery*, vol. 36, no. 5, pp. E180-E190, 2023.

[54] M. S. Khan et al., "VGG19 Network Assisted Joint Segmentation and Classification of Lung Nodules in CT Images," *Diagnostics*, vol. 11, no. 12, pp. 2208–2208, Nov. 2021.

[55] L. Li, Z. Yang, X. Yang, J. Li, and Q. Zhou, "PV resource evaluation based on Xception and VGG19 two-layer network algorithm," *Heliyon*, vol. 9, no. 11, pp. e21450–e21450, Nov. 2023.

[56] M. Bansal, M. Kumar, M. Sachdeva, and A. Mittal, "Transfer learning for image classification using VGG19: Caltech-101 image data set," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1-12, 2021.

[57] A. Karacı, "VGGCOV19-NET: automatic detection of COVID-19 cases from X-ray images using modified VGG19 CNN architecture and YOLO algorithm," *Neural Computing and Applications*, vol. 34, no. 10, pp. 8253-8274, 2022.

[58] N. Madusanka, P. Jayalath, D. Fernando, L. Yasakethu, and B.-I. Lee, "Impact of H&E Stain Normalization on Deep Learning Models in Cancer Image Classification: Performance, Complexity, and Trade-Offs," *Cancers*, vol. 15, no. 16, pp. 4144–4144, Aug. 2023.

[59] K. Shaheed, Q. Abbas, A. Hussain, and I. Qureshi, "Optimized Xception Learning Model and XgBoost Classifier for Detection of Multiclass Chest Disease from X-ray Images," *Diagnostics*, vol. 13, no. 15, pp. 2583–2583, Aug. 2023.

[60] D. Li, S. Yuan, and G. Yao, "Classification of lung nodules based on the DCA-Xception network," *Journal of X-ray Science and Technology*, vol. 30, no. 5, pp. 993–1008, Sep. 2022.

[61] B. Gülmez, "A novel deep neural network model based Xception and genetic algorithm for detection of COVID-19 from X-ray images," *Annals of Operations Research*, vol. 328, no. 1, pp. 617-641, 2022.

[62] K. Shaheed, Q. Abbas, A. Hussain, and I. Qureshi, "Optimized Xception Learning Model and XgBoost Classifier for Detection of Multiclass Chest Disease from X-ray Images," *Diagnostics*, vol. 13, no. 15, pp. 2583–2583, Aug. 2023.

[63] M. Rahimzadeh and A. Attar, "A modified deep convolutional neural network for detecting COVID-19 and pneumonia from chest X-ray images based on the concatenation of Xception and ResNet50V2," *Informatics in Medicine Unlocked*, vol. 19, p. 100360, 2020.

[64] I. Lahsaini, M. E. H. Daho, and M. A. Chikh, "Deep transfer learning-based classification model for covid-19 using chest CT-scans," *Pattern Recognition Letters*, vol. 152, pp. 122–128, Dec. 2021.

[65] M. S. A. Reshan *et al*., "Detection of Pneumonia from Chest X-ray Images Utilizing MobileNet Model," *Healthcare*, vol. 11, no. 11, p. 1561, May 2023.

[66] P. Podder *et al*., "LDDNet: A Deep Learning Framework for the Diagnosis of Infectious Lung Diseases," *Sensors*, vol. 23, no. 1, p. 480, Jan. 2023.

[67] C. Song *et al*., "Maize seed appearance quality assessment based on improved Inception-ResNet," *Frontiers in Plant Science*, vol. 14, Aug. 2023.

[68] Y. Chen *et al*., "Classification of Lungs Infected COVID-19 Images based on Inception-ResNet," *Computer Methods and Programs in Biomedicine*, p. 107053, Jul. 2022.

[69] C. Lin *et al*., "CIR-Net: Automatic Classification of Human Chromosome based on Inception-ResNet Architecture," *IEEE/ACM Trans. Comput. Biol. Bioinf*., vol. 19, no. 3, pp. 1285-1293, 2020.

[70] H. M. Emara *et al*., "Simultaneous Super-Resolution and Classification of Lung Disease Scans," *Diagnostics*, vol. 13, no. 7, pp. 1319–1319, Apr. 2023.

[71] S. Dash, P. K. Sethy, and S. K. Behera, "Cervical Transformation Zone Segmentation and Classification based on Improved Inception-ResNet-V2 Using Colposcopy Images," *Cancer Informatics*, vol. 22, p. 11769351231161477, Mar. 2023.

[72] H. Zhang *et al*., "Recurrence Plot-Based Approach for Cardiac Arrhythmia Classification Using Inception-ResNet-v2," *Frontiers in Physiology*, vol. 12, May 2021.

[73] K. Pawar, Z. Chen, N. J. Shah, and G. F. Egan, "Suppressing motion artefacts in MRI using an Inception-ResNet network with motion simulation augmentation," *NMR in Biomedicine*, vol. 35, no. 4, 2019.

[74] X. Wan, J. Luo, Y. Liu, Y. Chen, X. Peng, and X. Wang, "An image classification method for arrhythmias based on Gramian angular summation field and improved Inception-ResNet-v2," *Journal of Biomedical Engineering*, vol. 40, no. 3, pp. 465–473, Jun. 2023.

[75] Nillmani *et al*., "Segmentation-Based Classification Deep Learning Model Embedded with Explainable AI for COVID-19 Detection in Chest X-ray Scans," *Diagnostics*, vol. 12, no. 9, p. 2132, Sep. 2022.

[76] H. A. Sanghvi, R. H. Patel, A. Agarwal, S. Gupta, V. Sawhney, and A. S. Pandya, "A deep learning approach for classification of COVID and

pneumonia using DenseNet -201," *International Journal of Imaging Systems and Technology*, Sep. 2022.

[77] P. Podder, F. B. Alam, M. R. H. Mondal, M. J. Hasan, A. Rohan, and S. Bharati, "Rethinking Densely Connected Convolutional Networks for Diagnosing Infectious Diseases," *Computers*, vol. 12, no. 5, p. 95, May 2023.

[78] Z. Tao, H. Bingqiang, L. Huiling, Y. Zaoli, and S. Hongbin, "NSCR-Based DenseNet for Lung Tumor Recognition Using Chest CT Image," *BioMed Research International*, vol. 2020, pp. 1–9, Dec. 2020.

[79] A. M. Abdelmula, O. Mirzaei, E. Güler, and K. Süer, "Assessment of Deep Learning Models for Cutaneous Leishmania Parasite Diagnosis Using Microscopic Images," *Diagnostics*, vol. 14, no. 1, pp. 12–12, Dec. 2023.

[80] Y.-D. Zhang, S. C. Satapathy, X. Zhang, and S.-H. Wang, "COVID-19 Diagnosis via DenseNet and Optimization of Transfer Learning Setting," *Cognitive Computation*, pp. 1-17, 2021.

[81] X. Xue *et al*., "Design and Analysis of a Deep Learning Ensemble Framework Model for the Detection of COVID-19 and Pneumonia Using Large-Scale CT Scan and X-ray Image Datasets," *Bioengineering*, vol. 10, no. 3, pp. 363–363, Mar. 2023.

[82] I. Kanjanasurat, K. Tenghongsakul, B. Purahong, and A. Lasakul, "CNN–RNN Network Integration for the Diagnosis of COVID-19 Using Chest X-ray and CT Images," *Sensors*, vol. 23, no. 3, p. 1356, Jan. 2023.

[83] H. Malik, A. Naeem, R. A. Naqvi, and W.-K. Loh, "DMFL_Net: A Federated Learning-Based Framework for the Classification of COVID-19 from Multiple Chest Diseases Using X-rays," *Sensors*, vol. 23, no. 2, p. 743, Jan. 2023.

[84] C. Bada *et al*., "Inter-observer Agreement in Interpreting Chest X-rays on Children with Acute Lower Respiratory Tract Infections and Concurrent Wheezing," *Sao Paulo Medical Journal*, vol. 125, pp. 150-154, May 2007.

[85] V. P. Shah, M. G. Tunik, and J. W. Tsung, "Prospective Evaluation of Point-Of-Care Ultrasonography for The Diagnosis of Pneumonia in Children and Young Adults," *JAMA pediatrics*, vol. 167, pp. 119-125, 2013.

[86] V. A Caiulo *et al*., "Lung Ultrasound Characteristics of Community-acquired Pneumonia in Hospitalized Children," *Pediatric Pulmonology*, vol. 48, pp. 280-287, 2012.

[87] M. F. Hashmi *et al*., "Efficient Pneumonia Detection In Chest Xray Images Using Deep Transfer Learning," *Diagnostics*, vol. 10, p. 417, 2020.

[88] A. Hosny *et al*., "Artificial Intelligence In Radiology," *Nature Reviews Cancer*, vol. 18, pp. 500-510, 2018.

[89] S. Sharma and K. Guleria, "A Deep Learning based model for the Detection of Pneumonia from Chest X-Ray Images using VGG-16 and Neural Networks," *Procedia Computer Science*, vol. 218, pp. 357–366, 2023.

[90] M. Rahimzadeh and A. Attar, "A modified deep convolutional neural network for detecting COVID-19 and pneumonia from chest X-ray images based on the concatenation of Xception and ResNet50V2," *Informatics in Medicine Unlocked*, vol. 19, p. 100360, 2020.

[91] C. Ieracitano *et al*., "A fuzzy-enhanced deep learning approach for early detection of Covid-19 pneumonia from portable chest X-ray images," *Neurocomputing*, vol. 481, pp. 202–215, Apr. 2022.

[92] R. Kundu, R. Das, Z. W. Geem, G.-T. Han, and R. Sarkar, "Pneumonia detection in chest X-ray images using an ensemble of deep learning models," *PLoS ONE*, vol. 16, no. 9, p. e0256630, Sep. 2021.