JOIV : Int. J. Inform. Visualization, 7(2) - June 2023 430-438

INTERNATIONAL JOURNAL
ON INFORMATICS VISUALIZATION

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

Early Detection of Asymptomatic Covid-19 Infection with Artificial Neural Network Model Through Voice Recording of Forced Cough

Aisyah Khairun Nisa [a], I Gede Pasek Suta Wijaya [a,*], Arik Aranta [a]

[a] *Department Informatics Engineering, Mataram University, Jl Majapahit 62, Mataram, Lombok NTB, 83114, Indonesia*
*Corresponding author: *gpsutawijaya@unram.ac.id*

***Abstract*— SARS-CoV-2 is a virus that spreads the infection known as COVID-19, or Coronavirus 2019. According to data from the World Health Organization as of March 15, 2021, Indonesia has 1,419,455 cumulative cases and 38,426 cumulative deaths, ranking third among countries in terms of fatalities, behind Iran and India. Because COVID-19 was disseminated through direct contact with respiratory droplets from an infected individual, it spread swiftly and widely. According to the American Centers for Disease Control and Prevention, more than 50% of transmission rates are anticipated from asymptomatic individuals. The antigen tests have an accuracy of results ranging from 80–90% and are utilized for early detection of COVID-19. The cost of the antigen test is set to increase as of September 3, 2021, with prices ranging from IDR 99.000 to IDR 109.000; however, researchers are steadfastly searching for the best alternate methods for the early diagnosis of COVID-19. According to MIT News Office, a forced cough recording can identify an asymptomatic COVID-19 infection. Through the vocal recording of a forced cough, this study uses an artificial neural network (ANN) deep learning model to identify asymptomatic COVID-19 patients. The Artificial Neural Network (ANN) can distinguish asymptomatic people from forced cough recordings with an accuracy of up to 98% and a loss value of less than 3% by employing oversampling data. This model can be applied as a free, universal method for the early identification of COVID-19 infection.**

***Keywords*— ANN; asymptomatic; COVID-19; forced cough; oversampling.**

Manuscript received 17 Mar. 2022; revised 19 Sep. 2022; accepted 15 Dec. 2022. Date of publication 30 Jun. 2023.

I. INTRODUCTION

A disease known as coronavirus disease is caused by coronavirus 2 that causes severe acute respiratory syndrome (SARS-CoV-2). Before the outbreak began in Wuhan last December 2019, the aliment was not known [1]. Because COVID-19 can spread directly from one person to another through close contact, it spreads swiftly and rapidly[2]. Indonesia started to see a spike in COVID-19 cases based on a report from the World Health Organization as of March 15, 2021, Indonesia's total number of patients reached 1,419,455, the fourth position in the fourth case in Asia after India, Turkey, and Iran. Indonesia is also ranked as the third-highest country with a fatality rate, with 8,426 cases [3].

Because COVID-19 spreads by direct human contact with an infected person's respiratory tract, it spreads swiftly and rapidly. People can also be contaminated if they touch a contaminated surface [4]. This rapid spreading of COVID-19 requires an immediate testing process to deal with the spread risk [5]. The American Centers for Disease Control and Prevention (CDC) says that people without symptoms (OTG/asymptomatic) and COVID-19 patients who are not yet showing signs are likely to have more than 50% of the transmission rate. According to the CDC, 24% of asymptomatic people transmit the virus to others, and 35% of COVID-19 patients who have not shown symptoms infect others before they develop symptoms [6]. The severe acute respiratory syndrome coronavirus is the source of COVID-19 infection., which begins with an infection of the mucous membranes of the throat, where cough is a common symptom, and travels through the airways to the lungs.[7]. Asymptomatic COVID 19 infection is characterized as the absence of clinically important associated symptoms.

As a result, individuals are less likely to undergo a virus testing and could accidently transmit the virus to others. However, these asymptomatic people might not be entirely immune to the virus's changes. Researchers at MIT have found that asymptomatic people may cough differently than healthy people. These differences cannot be deciphered by the human ear. However, they turned out to be detectable by artificial intelligence because AI can see details and identify patterns that are invisible to the human eye. AI can also automatically generate features. Artificial Neural Networks are a popular type of AI system (ANN). The advantage of ANN is the ability to determine how the input and output data are related. This enables ANN to analyze the data under the volume of information we obtain. The quality of the model created by the ANN approach depends on the data quality. In order to properly classify the existing data, the ANN approach, therefore, requires feature extraction of the data. The feature extraction results will then be used to determine whether or not a forced cough recording is recognized as COVID-19.

The identified speaker is detected through the coefficient of correlation analysis, and the accuracy of fit of the identified speaker's speech feature frames from the ANN and GMM frames is measured in prior research on the prediction system. Experiments were run on spoken utterances from 30 unique speakers to verify the system's performance (20 males and 10 females). For 5-word utterances, system performance displayed an average recognition rate of 77%, and for trained voice utterances, it showed 43% when utterance length was increased to her 20-word utterances[8]. Coswara Respiratory Dataset, Speech and Cough Respiratory Datasets, and a combination of ICBHI and Coswara Respiratory Databases, another study on ensemble models and baseline networks evaluated using ICBHI achieves ICBHI scores between 0.920 and 0.9766.

Most notably, empirical findings demonstrate that when voice recordings are paired with the ICBHI and Coswara respiratory datasets, positive diagnoses of COVID-19 can be separated from other more prevalent respiratory disorders. [9]. Based on this premise, the authors propose researching to develop a machine-learning model to identify COVID-19 from forced cough audio using the ANN model. This model intends to assist medical professionals working in the health field.

II. Materials and Method

A. Materials and Tool

The dataset used in this research is a cough recording voice consisting of two classes, Positive COVID-19 and Negative COVID-19. The number of voice recordings for each class varies, with 19 for the Positive COVID-19 class and 151 for the Negative COVID-19 class. Recordings from each class are in the format of .wav with a maximum duration of 10 seconds. The tools used in this research process are hardware and software that are within the specifications shown in Tables 1.

TABLE I
HARDWARE SPECIFICATIONS

No	Name	Specification
1	Processor	Intel Core i3 6006u 2.6GHz
2	GPU	NVIDIA GeForce 920MX 2GB
3	RAM	8 GB + 4GB
4	Operating System	Windows 10 64bit
5	Programming Language	Python 3.8.5
6	Microsoft Office	Office 2019
7	Text Editor	GoogleColab

B. Research Flowchart

The research flowchart from literature research to writing the final report is shown in Fig. 1. The research begins with literature research, which studies various sources related to the problems raised in this research. Introductory and in-depth studies were conducted to help the researcher to understand the problem. Then, the data that has been collected will be the base to make the model, both as training data and test data. After the data was collected entirely, the ANN model was developed to classify the data. After the model is created, it is continued by testing the model with a confusion matrix, precision, and recall. The last stage in this research is making reports as documentation and learning for future research.

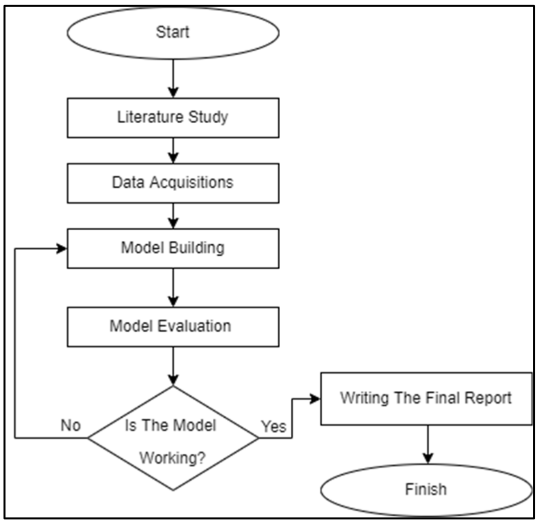

Fig. 1 Research Flowchart

It is fundamental to conduct literature research to strengthen the study by checking up books, journals, and other sources relevant to the research issue. The literary research's material is connected to research about extracting spectrogram from recorded voice, extracting the value of MFCC from the spectrogram, and Back Propagation-ANN.

1) COVID-19: the coronavirus is the infectious disease known as COVID-19, or coronavirus disease. A broad family of viruses called coronavirus can inhabit both humans and animals and make sick. The common cold and more severe conditions like Middle East Respiratory Syndrome (MERS) and the Severe Acute Respiratory Syndrome (SARS) are all caused by certain coronaviruses in humans [10]. Before the outbreak in Wuhan, China in December 2019, COVID19 was undetected. Around March 2, 2020, Indonesia reported its first COVID-19 case. The main symptoms of COVID19 infection are including sneezing and a dry cough.

2) Asymptomatic COVID-19: Asymptomatic infection is a condition where a person is positive for a disease but does not give any clinical symptoms to the person concerned. There are two mechanisms by which asymptomatic transmission can potentially occur [11]:

- Transmission from someone who has never experienced symptoms, if the infected person is asymptomatic throughout their infections, they are is still contagious.

- Transmission from a person during the incubation period if the infected person is contagious before developing symptoms.

Asymptomatic infection is different from a pre-symptomatic infection. The pre-symptomatic people show early disease symptoms, including high fever, dry cough, and loss of smell and taste, while asymptomatic people do not. Still, they can transmit COVID-19 without showing signs of being contaminated by the disease. They spread the disease through breathing, coughing, sneezing, singing, or talking [12].

3) Deep Learning: A subset of machine learning called deep learning excels at analyzing unstructured data. Deep learning methods exceed the state-of-the-art in machine learning. Therefore, features from multilevel data can be gradually learned through computational models. Deep learning is growing in popularity as the amount of data available increases and hardware advancements are offered by powerful computers[13]. Deep learning algorithms perform calculations and predictions repeatedly and gradually increase the accuracy of results over time. There are types of deep learning that are commonly used, namely:

- Artificial Neural Network (ANN)

ANN is a multi-layered algorithm structure that is modeled after the human brain. Artificial Neural Network has many advantages, such as it stores information throughout the network rather than in a database. Losing a little information in one place does not prevent the network from functioning properly, and it also can work with incomplete knowledge. ANN has a fault tolerance, which means damage to one or more cells in the ANN does not interfere with output production. This characteristic makes the network fault tolerant. For ANN to learn, it must determine examples, display them on the network, and teach the network according to the desired output. Network problems do not corrode immediately, and artificial neural networks make decisions by learning events and commenting on similar events. Finally, Artificial Neural Networks have the numerical strength to perform multiple tasks simultaneously.

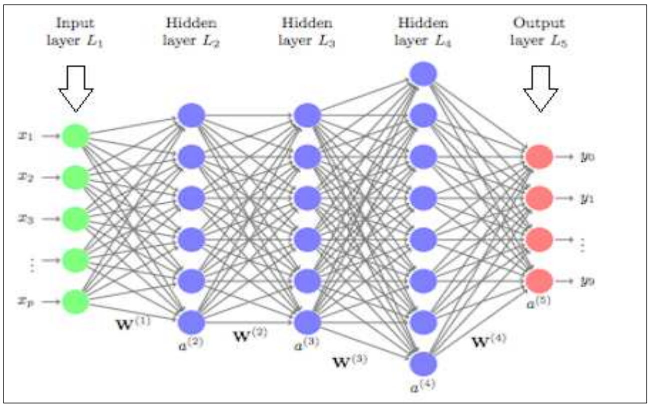

Fig. 2 Artificial Neural Network Architecture

- Convolutional Neural Network (CNN)

The CNN algorithm uses an advanced propagation type neural network with local information abstraction and positional universality. CNN is also one of the most popular models in use today. The computational model of this neural network involves one or more convolution layers that can be fully connected or aggregated and employs a variant of the Multilayer Perceptron. These convolutional layers produce a feature map that identifies the image's region. Finally, rectangles are created from this region and transmitted for non-linear processing.

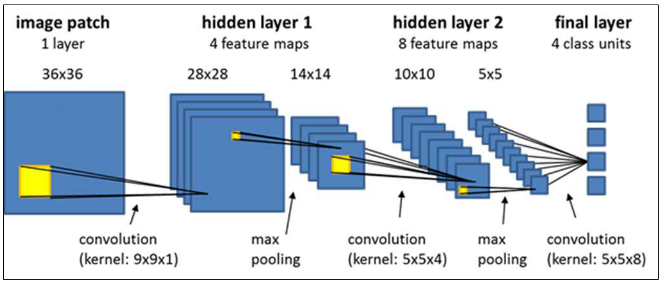

Fig. 3 Convolution Neural Network Architecture

- Recurrent Neural Network (RNN)

RNN is an algorithm using a neural network. A two-way signal is propagated with a recursive structure in the middle layer to handle variable-length data such as voice data and moving images.

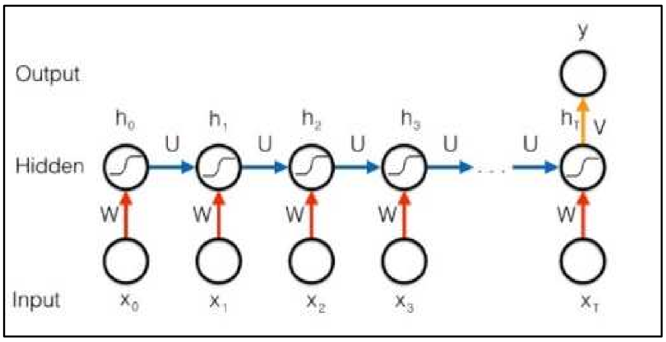

Fig. 4 Recurrent Neural Network Architecture

4) Artificial Neural Network: Artificial Neural Networks (ANNs) are computer programs inspired by biology that imitate how the human brain processes information. An artificial neural network (ANN) learns (or trains) from experience rather than programming by identifying patterns and relationships in data. ANNs are made up of hundreds of discrete units, artificial neurons, or processing elements (PEs), which connect to layers of coefficients to form a neural structure (weights). The network's connections between neurons provide neural computation its strength. A weighted input, a transfer function, and an output are all present in each PE. The transfer functions of a neural network's neurons, the learning rule, and the design itself all affect how the network behaves. In this sense, neural networks are parameterized systems since the weights are movable variables. The weighted total of the inputs determines how neurons are activated. The neuron generates a single output after the transfer function has processed the activation signal. The network becomes non-linear due to the transfer function. The interunit connections are improved during training until the prediction error is reduced and the network gets the required degree of accuracy. The network can be given fresh input data to predict the outcome after being trained and tested.[14].

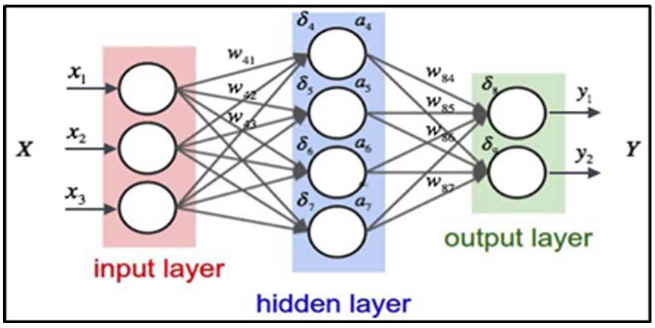

Fig. 5 Artificial Neural Network Architecture

5) WAV Recordings: WAV (Waveform Audio File Format) is a standard digital audio file with the WAV format and storing waveform data. WAV is one of the best audio file formats because of its lossless sound quality. WAV generally stores compressed audio in the 44.1 kHz, 16-bit stereo format, the standard format used for CD audio [15].

6) Biomarker: Biomarkers are molecules that indicate normal or abnormal biological functions to assess a person's health status. Various substances, including proteins, hormones, and DNA (genes), can act as biological markers. Blood, feces, urine, tumor tissue, and other bodily fluids and tissues may include biomarkers. There are numerous additional diseases, such as Alzheimer's, COVID-19, and others, for which biomarkers are used to detect disease alongside cancer. [16]

7) Activation Function: The activation function must be continuous, easily differentiable conditions, and a non-decreasing function[17] In this research, SoftMax is used as the activation function. The equation for the binary sigmoid activation function is as follows.

$$f(x) = \frac{\exp(x_i)}{\sum j \exp(x_j)} \quad (1)$$

The variation of output probabilities is SoftMax's key benefit. The probability varies from 0 to 1, and the total is 1. The probability of each class is returned, and the probability of the target class is raised when the SoftMax function is used in a multi-classification model.

8) Mel Frequency Cepstrum Coefficient (MFCC): One technique that is frequently utilized in the area of speech technology, including speaker and speech recognition, is MFCC (Mel Frequency Cepstral Coefficients) [18]. Utilizing this technique, feature extraction transforms voice signals into a number of metrics. Some of the advantages of this method are [18].

- Able to capture the characteristics of sound that are very important for speech recognition, or in other words, can capture critical information contained in the voice signal.
- Produce minimal data without losing the vital information it contains.
- Replicate the human hearing organ in perceiving signals

The voice data is initially retrieved using the MFCC method before being processed. The matching procedure compares the feature extraction results from the test data with the feature extraction results from the training data stored in the database using the results of feature extraction using the MFCC method. LVQ is used to classify the feature extraction data, and LVQ learns vectors from the output of the MFCC filter. The closest distance can be determined using Euclidean distance in the sound matching process to determine how similar the sounds from the test data to the pattern data in the database are. For accuracy testing, 151 training data and 50 test data were used from 4 resource persons. The results of the accuracy test show that the application of the MFCC and LVQ methods can be applied to speech recognition. The test results get an average percentage level of accuracy of 88.89% using a frame size of 512 and the parameter value in LVQ learning using a learning rate (dec = 0.05), a decrease in learning rate (dec = 0.1), and a maximum epoch = 1000.

9) Oversampling Data: One of the problems with the imbalanced data used for classification is that there are a few examples from minority classes for the model to learn the decision boundaries effectively. Counting the instances of minority classes is one method for resolving this issue. Simple replication of the minority class example in the training dataset before model fitting can do this. For balancing, the distribution of the classes, the model is not given any new information. Synthesizing a new minority class example is preferable to copying an existing minority class example. This is a very useful form of data extension for table data.

The Synthetic Minority Oversampling Technique, or SMOTE for short, is perhaps the most used method for creating new samples [19]. When using SMOTE, nearby patterns in the feature space are chosen, a line is drawn between them, and additional patterns are drawn at various locations along the line. A random sample is initially chosen from the minority class in particular. The k-nearest neighbor is thus located in this case (k is typically equal to 5). A composite example is created between the two examples in the feature space at randomly chosen places between randomly chosen neighbors.

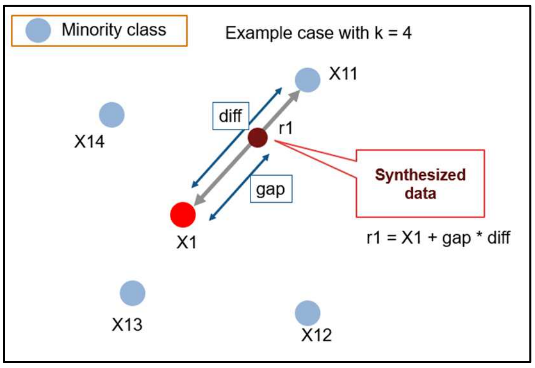

Fig. 6 Oversampling using SMOTE method

C. Related Research

Medical professionals in the health field have used audio recording recognition research extensively as a technique to assist in gender identification [20] to detect autistic behavior in children [21]. In the past five years, previous studies will serve as a reference in this research. Research on the recognition of COVID-19 through forced coughing recordings has been carried out several times, and most of the studies recorded a high degree of accuracy.

Previous research using AI speech processing to framework leverages acoustic biomarker features from patient's cough recording audio provides a personalized patient saliency map to monitor the patient in real-time, non-invasively longitudinally, and at essentially zero cost. The cough recordings are transformed with Mel-Frequency Cepstral Coefficient and inputted into the CNN model, resulting in 98.5% accuracy 22].

Another research using machine learning to distinguish positive and negative covid-19 patients through voice recording of cough via global smartphone recording resulted in 98% accuracy from training data dan 94% accuracy from testing data. Seven machine learning classifiers, including logistic regression, k-nearest neighbor, support vector machine, multilayer perceptron, convolutional neural network, long short-term memory, and residual-based neural network architecture, are used in the research [23].

In another related research of voice recording in the medical field, the recording of cough, speech, and breathing voices has been used to detect various diseases. There are several studies using speech processing to detect diseases in the medical field, such as Parkinson [24] Dementia [25], and even neurodegenerative and psychiatric diseases [26]. There are also several studies on repository system studies such as asthma [27], and neck cancer [28]. Asthma and Covid-19 share common symptoms. Therefore, some studies have been conducted to prove the validity of voice recording in detecting Covid-19. One of the studies focusing on vocal signs and symptoms related to Covid-19 and risk factors for their persistence resulted in positive outcomes of the difference between healthy and infected people. Individuals affected by Covid-19 had a higher frequency of vocal signs during the disease [24].

Another research using voice quality and vocal tract discomfort symptoms in a patient with Covid-19 also resulted in deviations in the voice quality of infected subjects, with mild vocal tract discomfort, then the healthy subjects [30]. All previous research has been conducted to prove that there is a chance of using voice recording of forced cough as early detection of Covid-19.

D. Data Acquisitions

The dataset used to build and evaluate the model must be split into training and testing data. In this research, the data will be split into 70% of data training consisting of 106 data labeled Negative Covid-19, 13 data labeled Positive Covid-19, while the remaining 30% will be data testing with 45 data labeled Negative Covid-19, and 6 data labeled Positive Covid-19.

E. Model Building

The design of the algorithm for this model can be seen in Figure 7.

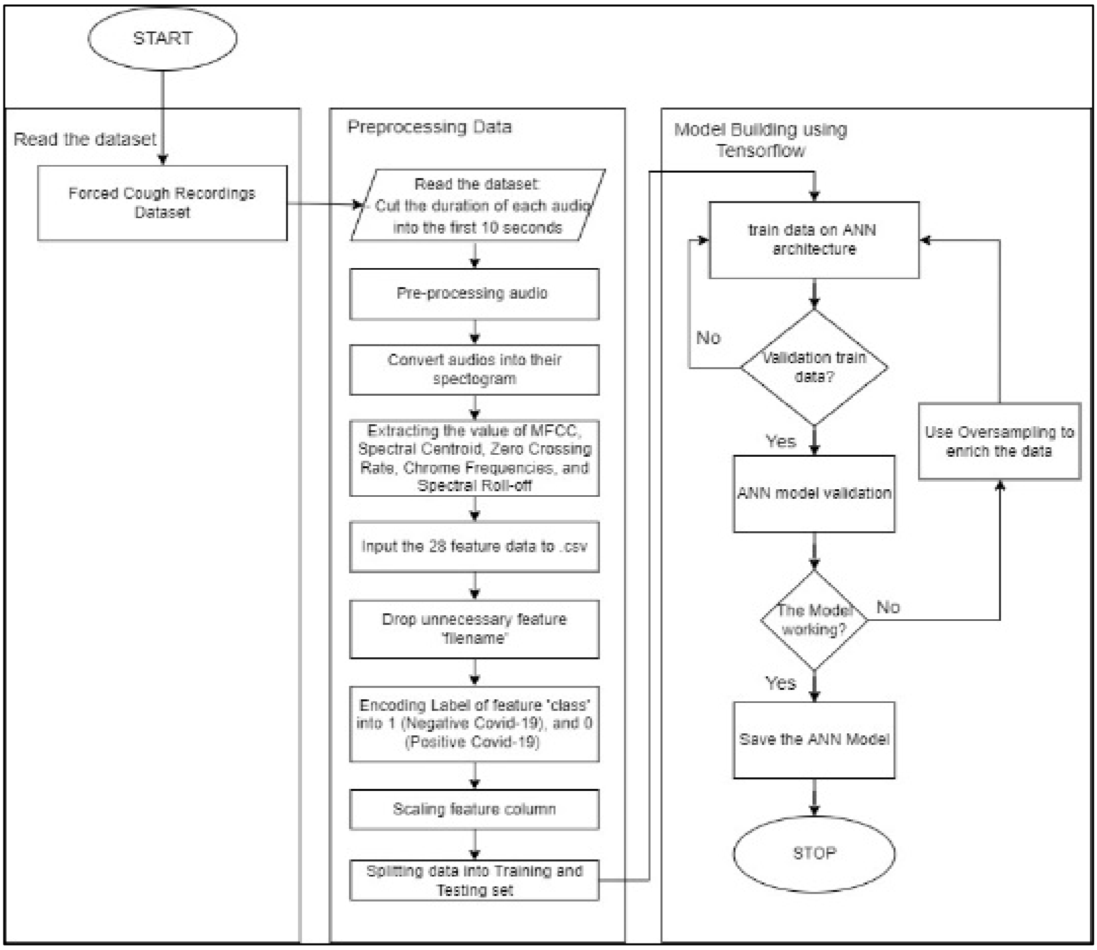

Fig. 7 Analyze model building process

Figure 7 shows three main processes: read the dataset, pre-process the data, and then train and test the data using the model built using the TensorFlow library. The process is described as follows:

- Reading the dataset
- Pre-processing the data
- Model building using TensorFlow

F. Training

The training stage is forming a classification model by training the model with training data. The training process helps the model recognize the characteristics of the data and classify them. The training process begins with converting forced cough recording into a spectrogram, followed by feature extraction. Feature data to be taken are MFCC, Spectral Centroid, Zero Crossing Rate, Chrome Frequencies, and Spectral Roll-off, then saved in a .cvs file. After that, the data will be analyzed using pandas then the classification will be carried out using the ANN method.

The following are the stages in the training process:

1) Convert the spectrogram from the forced cough recording.

The conversion is done by changing the sound recording from .wav to spectrogram image using the 'specgram()' function as a technique used to identify coughing sound data. The sound taken is 10 seconds long. In forming the spectrogram, several things need to be considered, namely:

- The number of data points used in each block for DFT in the FFT transformation (NFFT) is 2048 for more frequency resolution.
- The number of sample points per second (FS) is 2.0.
- The center frequency (FC) offset plot area to reflect the frequency range used when the signal was acquired and filtered.
- The number of overlapping points between blocks (noverlap) is 128.
- Spectrum side a to be returned (set default value)
- Scaling the values in the specification using 'dB', which returns the amplitude dB (20 * log10)

Below is the result of a conversion of a voice recording of positive Covid-19 cough and negative Covid-19 cough into their respective spectrogram. Fig. 8 shows the spectrogram from a negative Covid-19 cough, and Fig. 9 shows the spectrogram from a positive Covid-19 cough; the positive Covid-19 cough tends to have a larger range of cough sound than negative Covid-19 cough, which are visualized by the orange-yellow area in the spectrogram.

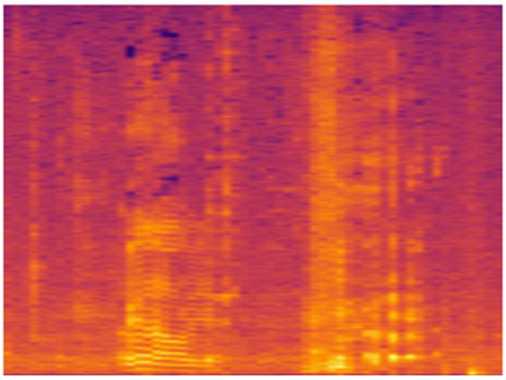

Fig. 8 Spectrogram of Negative Cough

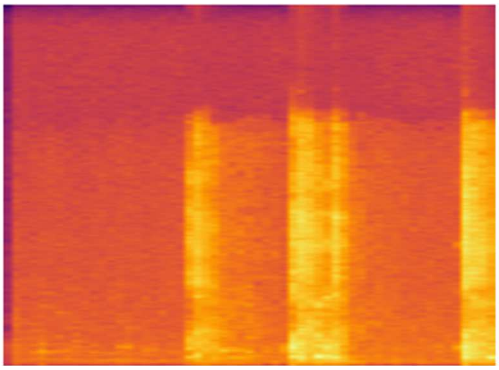

Fig. 9 Spectrogram of Positive Covid-19 Cough

2) Extracted the value of MFCC, Spectral Centroid, Zero Crossing Rate, Chrome Frequencies, and Spectral Roll-off from the spectrogram and save it to csv file.

3) Analyzed dataset using pandas

4) Classification using ANN

The classification carried out in this model is ANN. Input from ANN is the output of the previous feature extraction process. The output classification results in this model will produce 2 labels, namely Positive COVID-19 or Negative COVID-19. In this research neural network will be build using 'Sequential () class.' Input layer will be consisting of 64 nodes, with 26 input features. To keep things simple, there will be two hidden layers. The initial hidden layer that acts as input layer has 64 nodes, then the next will have 32 nodes, and then 8 nodes while the last hidden layer has 2 nodes.

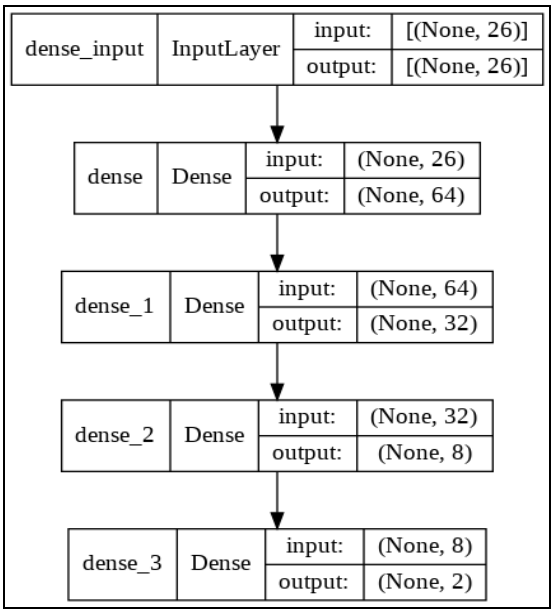

Fig. 10 Model Artificial Neural Network

G. Testing

The testing phase aims to determine the performance of the model that has been made. At the testing stage, the accuracy consists of the test data and training data, the computational time for test data and training data, and the accuracy gap between test data and training data. The method used in the testing process is the confusion matrix, precision, recall, dan f1-score.

1) Confusion Matrix: The confusion matrix is the method used to test the model generated at the training stage. Table 2 is the confusion matrix of this research.

TABLE II
THE CONFUSION MATRIX

		Predicted Class	
		Positive Covid-19	**Negative Covid-19**
Actual Class	**Positive Covid-19**	True Positive Covid-19	False Negative Covid-19
	Negative Covid-19	False Positive Covid-19	True Negative Covid-19

Calculation of confusion matrix using the following equation:

$$accuracy = \frac{(True\_Pos\_Covid19)+(True\_Neg\_Covid19)}{Total}$$

2) Precision: Precision is the measure to which the model can accurately forecast the class that will exist. When the model incorrectly recognizes another class as the COVID-19 class, which indicates that the accuracy is low, precision is important for assessing the impact of erroneous COVID-19 (in the COVID-19 class). The following equation is used to calculate the precision model:

$$precission = \frac{(True\_Pos\_Covid19)}{(True\_Pos\_Covid19)+(Flase\_Pos\_Covid19)} \quad (2)$$

3) Recall: The model's success rate in recovering data is measured by a recall. When the model should recognize the input as COVID-19, a recall helps to assess the effects of false-positive COVID-19 and false-negative COVID-19 on the true class affirmative COVID-19. The model still makes a mistake, so the recall is of poor value. This is risky because the created model needs to be able to accurately categorize the COVID-19 class to reduce errors that make the COVID-19 class unpredictable. The equation can be used to calculate the recall model for the COVID-19 class:

$$recall = \frac{(True\_Pos\_Covid19)}{(True\_Pos)+(Flase\_Pos)-(False\_Neg)} \quad (3)$$

4) F1 Score: F1 Score is a weighted comparison of the average precision and recall value. To evaluate this study, the recall metric is calculated using the equation:

$$F1\,Score = \frac{Recall*Precission)}{Recall+Precission)} \quad (4)$$

III. Result and Discussion

A. Dataset Sample

As mentioned earlier, the voice recordings dataset consists of 19 for the Positive COVID-19 class and 151 for the Negative COVID-19 class. Building and evaluating the model data needs to be split into training and testing data under the ratio of 70% versus 30%. It means a training set consisting of 106 data labeled Negative Covid-19, and 13 data labeled Positive Covid-19, while the testing set consists of 45 data labeled Negative Covid-19 and 6 data labeled Positive Covid-19. Fig. 10 and Fig. 11 show the wave shape of .wav recording of forced cough in labels positive COVID-19 and negative COVID-19, respectively.

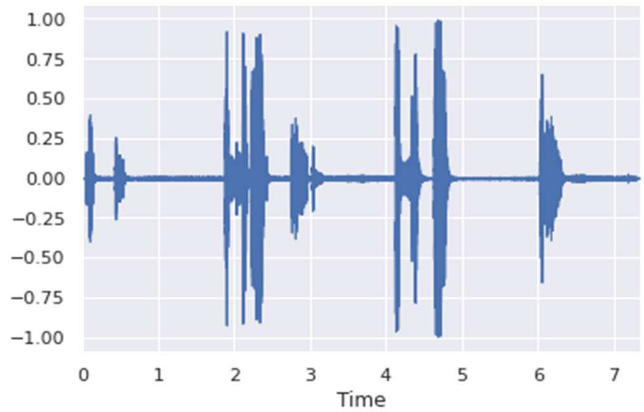

Fig. 11 Wave plot of Negative COVID-19 forced cough recording

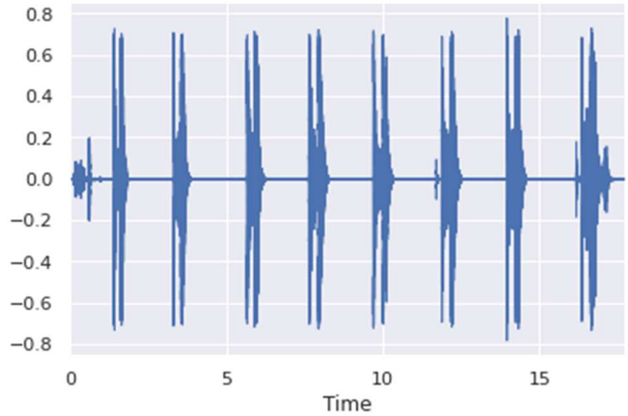

Fig. 12 Waveplot of Positive COVID-19 forced cough recording

The following process is spectrogram extraction which aims to visualize audio and make it easier to extract features in audio.

B. Data- Pre-processing

Data pre-processing is crucial because it directly impacts the project's success rate. Data preparation decreases the complexity of the data being analyzed because real-world data is often filthy. The data is unclean if there are duplicates or incorrect data, missing attributes, missing attribute values, noise, or outliers. Any of these will lower the effectiveness of the outcomes. In this case, the pre-processing data starts with converting audios into their spectrogram using the 'specgram()' function as a technique used to identify coughing sound data. The sound taken is 10 seconds long. From the spectrogram, MFCC (20 ordos), Spectral Centroid, Zero Crossing Rate, Chrome Frequencies, and Spectral Roll-off will be extracted. After that, the features in the dataset are to be scaled using the library "StandardScaler()." StandardScaler is a class from sklearn to normalize data so that the data used does not have large deviations. After the data is clean and ready to be processed, the data will be split into 70% of data training, while the remaining 30% will be data testing.

C. The Performance ANN on Original Data

The training process showed convergence when the model was deployed using the original dataset without oversampling and augmentation, as presented in Fig. 13 and 14. These results seem that the ANN model has worked well in classifying COVID-19 and non-covid coughs. The accuracy gets better with each iteration (Fig. 13), and the loss value gets smaller with each iteration (Fig. 14). This is not enough to prove whether the model is working correctly. So, validation of the model is a must, and the model can be validated using a confusion matrix and the value of precision, recall, and F1-score.

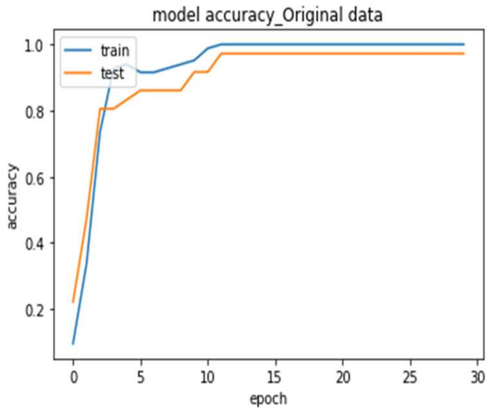

Fig. 13 Model accuracy from Original Data

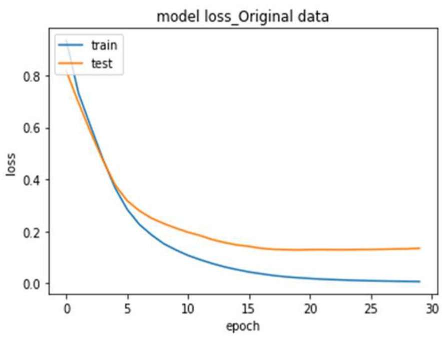

Fig. 14 Model loss from Original Data

Furthermore, the model was validated with data testing, so a confusion matrix was obtained as in Table 3, which means that the accuracy, precision, recall, and F1 score were 92%, 96%, 67%, and 73%, respectively. These results indicate that the ANN model cannot work correctly and indicate overfitting where the difference in accuracy between training and test data is more than 10%, as shown in Fig. 13. The leading cause of overfitting is the imbalance ratio of a positive and negative class of original dataset samples, namely 151 versus 19. Therefore, an oversampling was carried out in the next test to enrich the data.

TABLE III
THE CONFUSION MATRIX OF THE ORIGINAL DATA TESTING

True Labels	**Pos**	2	4
	Neg	0	45
		Pos	**Neg**
		Predicted Labels	

From the ROC point of view, as shown in Fig. 15, the True Positive Rate value is 98% against the False Positive Rate. It means that the ANN model that has been created has successfully diagnosed the Covid-19 positive forced cough recording correctly. However, based on the values of precision, recall, F1-score, and the accuracy proven to be overfitting, there is a need to augment the data. In this case, the researcher used the oversampling method, using the SMOTHE technique to enrich the data in the following experiment.

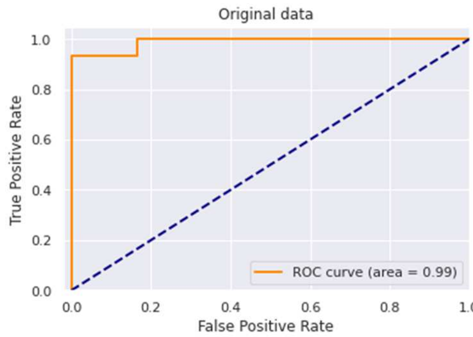

Fig. 15 ROC of original data

D. The Performance ANN on Oversampling Data

Oversampling data from the original dataset using SMOTE method. SMOTE is a method of oversampling in which artificial specimens are created for the minority class, in this case, the positive class. The overfitting issue is countered by this algorithm. With the aid of interpolation between the positive instances that lay together, it concentrates on the feature space to create new instances.

```
original positive cases: 19  and  total cases: 170

After oversampling
negative cases: 106
positive cases: 74 (41.11% of total)
```

Fig. 16 Output oversampling data

After oversampling the original data, the positive class value increased to 74 samples from 19 samples (see Fig. 16), which means it increased as much as 41.11% from the total sample. The oversampling data were tested into the ANN model to distinguish between the original and oversampling data. The training process using oversampling data provided a loss of 3.63% and an accuracy of 98.04%. These results show satisfactory results, and the overfitting condition can be overcome. Next, similar to the previous evaluation, the model was validated with data testing, so a confusion matrix was obtained as in Table 5, which means that the accuracy, precision, recall, and F1 score were 98%, 99%, 92%, and 95%, respectively.

Furthermore, the oversampling can improve the overfitting problem ANN model for the original data. Using the ROC in Fig. 17, the ANN model provides accuracy almost perfectly indicated by an orange line with a 1.0 True positive rate threshold. According to the achievements, the ANN model can potentially be implemented as a tool for Covid-19 pre-screen from forced cough recordings.

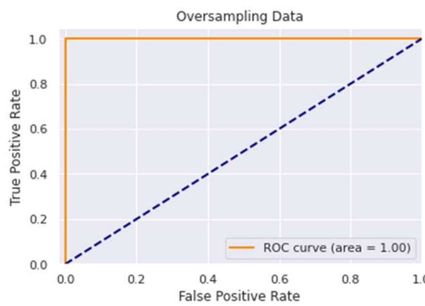

Fig. 17 ROC of oversampling data

IV. Conclusion

Based on the experimental results and discussions, it can be concluded that the Artificial Neural Network (ANN) model with four layers, 26 features as input, and activation function of ReLu and SoftMax succeeded in detecting COVID-19 patients through the voice recording of forced cough. The best ANN model was obtained with an accuracy of 98%, with a loss value of less than 3%. It shows that this model can be a solution for early detection of COVID-19 infection at no cost.

The research that has been carried out is still far from perfect. Several efforts will be made to develop this research, including adding the number of datasets and augmenting audio files to increase data variations to reduce the risk of overfitting or underfitting. Additionally, the model will be deployed for Android or web applications so that people can easily access it anytime and anywhere directly.

Acknowledgment

This work is dedicated to those affected by the COVID-19 pandemic and those helping to fight this battle. Also, we are grateful to the research team from the Indian Institute of Technology Kharagpur, who have provided us with the dataset of forced cough recording that is available for the public through Kaggel.com.

References

[1] Y. C. Wu, C. S. Chen, and Y. J. Chan, "The outbreak of COVID-19: An overview," *Journal of the Chinese Medical Association*, vol. 83, no. 3. Wolters Kluwer Health, pp. 217–220, 2020. doi: 10.1097/JCMA.0000000000000270.

[2] L. Morawska *et al.*, "How can airborne transmission of COVID-19 indoors be minimised?," *Environment International*, vol. 142. Elsevier Ltd, Sep. 01, 2020. doi: 10.1016/j.envint.2020.105832.

[3] World Health Organization, "Indonesia: WHO Coronavirus Disease (COVID-19) Dashboard | WHO Coronavirus Disease (COVID-19) Dashboard.," *https://covid19.who.int/region/searo/country/id*, Indonesia, March 15, 2021.

[4] N. H. L. Leung, "Transmissibility and transmission of respiratory viruses," *Nature Reviews Microbiology*, vol. 19, no. 8. Nature Research, pp. 528–545, August 1, 2021. doi: 10.1038/s41579-021-00535-6.

[5] A. M. Pollock and J. Lancaster, "Asymptomatic transmission of covid-19," *The BMJ*, vol. 371. BMJ Publishing Group, Dec. 21, 2020. doi: 10.1136/bmj.m4851.

[6] D. P. Oran and E. J. Topol, "Prevalence of asymptomatic SARS-CoV-2 infection. A narrative review," *Ann Intern Med*, vol. 173, no. 5, pp. 362–368, Sep. 2020, doi: 10.7326/M20-3012.

[7] T. Singhal, "A Review of Coronavirus Disease-2019 (COVID-19)," *Indian Journal of Pediatrics*, vol. 87, no. 4. Springer, pp. 281–286, April 1, 2020. doi: 10.1007/s12098-020-03263-6.

[8] A. Nichie and G. A. Mills, "Voice Recognition Using Artificial Neural Networks And Gaussian Mixture Models."

[9] C. Wall, L. Zhang, Y. Yu, A. Kumar, and R. Gao, "A Deep Ensemble Neural Network with Attention Mechanisms for Lung Abnormality Classification Using Audio Inputs," *Sensors*, vol. 22, no. 15, Aug. 2022, doi: 10.3390/s22155566.

[10] World Health Organization, "Coronavirus disease (COVID-19)," *https://www.who.int/westernpacific/health-topics/detail/coronavirus*, Oct. 18, 2021.

[11] Y. Li *et al.*, "Asymptomatic and Symptomatic Patients With Non-severe Coronavirus Disease (COVID-19) Have Similar Clinical Features and Virological Courses: A Retrospective Single Center Study," *Front Microbiol*, vol. 11, Jun. 2020, doi: 10.3389/fmicb.2020.01570.

[12] M. Zuin, V. Gentili, C. Cervellati, R. Rizzo, and G. Zuliani, "Viral load difference between symptomatic and asymptomatic COVID-19 patients: Systematic review and meta-analysis," *Infectious Disease Reports*, vol. 13, no. 3. MDPI, pp. 645–653, 2021. doi: 10.3390/IDR13030061.

[13] A. Mathew, P. Amudha, and S. Sivakumari, "Deep learning techniques: an overview," in *Advances in Intelligent Systems and Computing*, 2021, vol. 1141, pp. 599–608. doi: 10.1007/978-981-15-3383-9_54.

[14] R. Dastres and M. Soori, "Artificial Neural Network Systems Virtual Optimization of Manufacturing Processes View project Computer Integrted Manufacturing View project Artificial Neural Network Systems," 2021. [Online]. Available: https://www.researchgate.net/publication/350486076.

[15] F. Djebbar and B. Ayad, "Journal of Information Hiding and Multimedia Signal Processing c," 2017.

[16] K. Waury, E. A. J. Willemse, E. Vanmechelen, H. Zetterberg, C. E. Teunissen, and S. Abeln, "Bioinformatics tools and data resources for assay development of fluid protein biomarkers," *Biomark Res*, vol. 10, no. 1, p. 83, Nov. 2022, doi: 10.1186/s40364-022-00425-w.

[17] V. Bansal, "Activation Functions: Dive into an optimal activation function," Feb. 2022, [Online]. Available: http://arxiv.org/abs/2202.12065

[18] T. Hori *et al.*, "Multi-microphone speech recognition integrating beamforming, robust feature extraction, and advanced DNN/RNN backend," *Comput Speech Lang*, vol. 46, pp. 401–418, Nov. 2017, doi: 10.1016/j.csl.2017.01.013.

[19] D. D. Saputra *et al.*, "Optimization Sentiments of Analysis from Tweets in myXLCare using Naïve Bayes Algorithm and Synthetic Minority over Sampling Technique Method," in *Journal of Physics: Conference Series*, Mar. 2020, vol. 1471, no. 1. doi: 10.1088/1742-6596/1471/1/012014.

[20] V. G. Enriquez and M. Singh, "Gender Detection Using Voice Through Deep Learning," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2022, vol. 13184 LNCS. doi: 10.1007/978-3-030-98404-5_50.

[21] M. Milling *et al.*, "Evaluating the Impact of Voice Activity Detection on Speech Emotion Recognition for Autistic Children," *Front Comput Sci*, vol. 4, 2022, doi: 10.3389/fcomp.2022.837269.

[22] J. Laguarta, F. Hueto, and B. Subirana, "COVID-19 Artificial Intelligence Diagnosis Using only Cough Recordings," *IEEE Open J Eng Med Biol*, vol. 1, pp. 275–281, 2020, doi: 10.1109/OJEMB.2020.3026928.

[23] M. Pahar, M. Klopper, R. Warren, and T. Niesler, "COVID-19 cough classification using machine learning and global smartphone recordings," *Comput Biol Med*, vol. 135, Aug. 2021, doi: 10.1016/j.compbiomed.2021.104572.

[24] S. Arora, L. Baghai-Ravary, and A. Tsanas, "Developing a large scale population screening tool for the assessment of Parkinson's disease using telephone-quality voice," *J Acoust Soc Am*, vol. 145, no. 5, pp. 2871–2884, May 2019, doi: 10.1121/1.5100272.

[25] S. Merugu, A. Kumar, and G. Ghinea, "Advanced Technologies and Societal Change Track and Trace Management System for Dementia and Intellectual Disabilities." doi: 10.1007/978-981-19-1264-1_10.

[26] V. Rentoumi *et al.*, "LANGaware: Leveraging machine learning on natural language for the early detection of neurodegenerative and psychiatric diseases," *Alzheimer's & Dementia*, vol. 17, no. S11, 2021, doi: 10.1002/alz.052520.

[27] S. Yadav, M. Keerthana, D. Gope, U. Maheswari K., and P. Kumar Ghosh, "Analysis of Acoustic Features for Speech Sound Based Classification of Asthmatic and Healthy Subjects," in *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2020, pp. 6789–6793. doi: 10.1109/ICASSP40776.2020.9054062.

[28] T. Bressmann, "Speech Disorders Related to Head and Neck Cancer," in *The Handbook of Language and Speech Disorders*, 2021. doi: 10.1002/9781119606987.ch22.

[29] A. P. Dassie-Leite, T. P. Gueths, V. V. Ribeiro, E. C. Pereira, P. do N. Martins, and C. R. Daniel, "Vocal Signs and Symptoms Related to COVID-19 and Risk Factors for their Persistence," *Journal of Voice*, Aug. 2021, doi: 10.1016/j.jvoice.2021.07.013.

[30] S. A. Tohidast, B. Mansuri, M. Memarian, A. H. Ghobakhloo, and R. C. Scherer, "Voice Quality and Vocal Tract Discomfort Symptoms in Patients With COVID-19," *Journal of Voice*, Oct. 2021, doi: 10.1016/j.jvoice.2021.09.039.