



Cough Detection System Using Machine Learning Technique

^[1] Noorfadzli Abdul Razak, ^[2] Muhammad Abdur Razaq Abu Bakar, ^[3] Muhammad Saifullah Zahari, ^[4] Juliana Johari, ^[5] Noor Azlina Mohd Salleh, ^[6] Yupiter Harangan Prasada Manurung
^[1] noorfadzli@uitm.edu.my, ^[2] muhammadabdurazaq98@gmail.com, ^[3] 2020996725@student.uitm.edu.my
^[4] julia893@uitm.edu.my, ^[5] noorazlinamohdsalleh@uitm.edu.my, ^[6] yupiter.manurung@uitm.edu.my

Abstract—Coughing is the sudden expulsion of air from the lungs to clear the breathing passageways of unwanted irritants. However, if it sounds too loud and occurs for long periods, then it will cause a person chest pain, difficulty breathing and a high fever which requires treatment from a doctor. Knowing the seriousness of the matter, this research develops a cough detection system using machine learning technique. The system employed five distinct modules namely audio sampling, sound feature extraction, model training, cough detection, and audio data testing. Via the modules, when a person coughs then an important audio signal characteristic will be retrieved. It became an input for an Artificial Neural Network (ANN) model that has been trained with a cough sound dataset. The model will analyze the sound for a coughing fit and identify whether the sound is a cough or vice versa. A real-time test was arranged to test the performance of the model. It was programmed in an embedded controller and then installed to a device namely a Modular and Open System (MOST). The results demonstrate the model has a precision of 97.5% and successfully detects cough from a user.

Index Terms—cough, machine learning, online detection, training model

I. INTRODUCTION

The coughing reflex is an important part of the human defence mechanism. It assists in the removal of particles and secretions from the airways, while also providing protection for the lower airways. It is one of the signs of respiratory disorders that is seen the most frequently. A persistent cough can occasionally have a negative influence on the quality of life, and it also has the potential to cause organ damage in addition to high levels of mental and social stress. The very first thing that needs to be done to treat a cough is to determine how severe it is. The doctors can have a better knowledge of the patient's state and even track the progression of the disease thanks to accurate evaluations of the cough's intensity. Subjective measures such as cough scores, Visual Analogue Scales (VAS), and quality-of-life questionnaires are typically used in clinical practice to assess the intensity of coughing [1].

Coughing is a symptom that is shared by a few different illnesses, including asthma, bronchitis, pertussis, and Coronavirus disease (COVID-19). The sound of the cough is often distinct for each respiratory condition, which enables medical professionals to identify the sickness based only on the sound of the cough. For this reason, a great number of digital technology solutions that made use of big data analysis, the Internet of Things (IoT), Blockchain, Artificial Intelligence (AI) in Machine Learning (ML) and Deep Learning (DL) are offered to detect the disease based on the cough sound [2]. An example of this work was done by Woochang Shin that combined the IoT and AI to produce a

cough identification system [3]. For the system, a sensor was connected to a respirator to record coughing and other types of sounds. Then AI model was trained using these recordings and the Mel-spectrogram conversion method was used to identify sound data. The system performance was declared to have a sensitivity of 95.12%, a specificity of 100%, and an overall accuracy of 97.94%.

J. Amoh, S. Member, K. Odame, and S. Member introduced two distinct methods by employing Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN) as a cough detector [4]. These methods were used to solve the visual recognition issue and a sequence-to-sequence labelling problem in the detector system. Outcomes of the system show the CNN achieves a greater sensitivity of 87.7% than the RNN, which achieves a higher specificity of 92.7%. As in reference [5], a prototype consists of electrocardiogram (ECG), a thermistor, a chest belt, an accelerometer, an oximeter, contact microphones, and audio microphones was developed for cough detection and classification. The significance of each sensor was evaluated based on its performance in relation to three aspects, which were the capacity to identify coughing from other event categories, the ability to detect coughing, and the mutual information acquired between the features. The results from the sensors were interpreted and shown using a graphical view. In this view, the cough extracts that were identified are visualised and categorised according to their similarity in terms of the audio attributes timbre, cough length, and signal energy.

Besides that, S. H. Shin, T. Hashimoto, and S. Hatano

proposed an ANN model and a Hidden Markov Model (HMM) to form a hybrid model for a system that can differentiate a cough sound from other environmental sounds [6]. The ANN model made use of energy cepstral coefficients that were created by filter banks that were based on human auditory characteristics. Then, an output of this ANN model and a filtered envelope of the signal were utilised to make an input sequence for the HMM that deals with the temporal fluctuation of the sound signal. This sequence was then fed into the HMM. Additionally, A. Rana, Y. Dhiman, and R. Anand present a method for the detection of cough that is constructed using Edge Impulse Software and Arduino 33 BLE Sense board [7]. It was able to differentiate between a genuine cough and a generic unwanted signal that was present in the background. The Edge Impulse Software was used to train a huge dataset consisting of a variety of samples for cough as well as unwanted noise. For identifying the resonance of a cough on an immediate basis, a TinyML model based on machine learning has been designed. It has been determined that the proposed method can achieve an accuracy of recognition that was nearly 97%. Next, S. Khomsay, R. Vanijirattikhon and J. Suwatthikul purposed a method for cough detection that makes use of Principal Component Analysis (PCA) and Deep Learning Networks (DLN) that are powered by TensorFlow [8]. PCA was utilised in the process of doing feature extraction prior to transmitting the data to be used by DLN to train a model. TensorFlow's graph model was used to carry out computations, and these computations were quite effective in practice. According to the findings, DLN in conjunction with PCA was an effective method for detecting cough signals. PCA+DLN had a performance that is approximately 99.91%, but DLN only has a performance of about 98.45%. Thus, according to the findings, the PCA+DLN model was superior to the DLN model in terms of both the amount of time it takes to run and the accuracy it provides. Finally, F. Barata, K. Kipfer, M. Weber, P. Tinschert, E. Fleisch and T. Kowatsch introduced agnostic mobile cough detection using Convolutional Neural Networks (CNN) [9]. A total of 43 participants, using five distinct recording devices to capture a total of 6737 cough samples and 8854 control noises were carried out. After that, two different strategies from the previous work were reimplemented and analysed how well they performed in two distinct circumstances on a variety of devices. To lessen the differences between the various devices, an effective CNN design and an ensemble-based classifier were created. The method achieved mean accuracy levels ranging from 84.9% to 90.9%, demonstrating consistency across devices and exceeding earlier learning algorithms.

Seeing this opportunity to contribute, this project is focused on developing a system that can identify a person with a cough. Having this system allows precautionary measures can be taken by the responsible party to control or not allow the person to enter the premises. At present, the system is meant to detect a normal cough only. It will be an initial setup and work to enable the system later to differentiate between coughs induced by COVID-19 and ordinary coughs. To do this, an embedded controller was

programmed using ML software to learn from several samples of cough sound and later, it can make predictions and judgments about whether the sound is a cough or not. For this research, the ML software that had been employed was Edge Impulse. The software was selected since it has the advantages of ease of setup to interface with collected data, perform data sampling, data feature extraction, train a detection model, and capture sound and test model in real-time. These are the primary components that make up this system.

II. METHODOLOGY

In this section, the methods used to develop and test the cough detection model will be discussed. Each method employed and its flow is illustrated in Fig. 1.

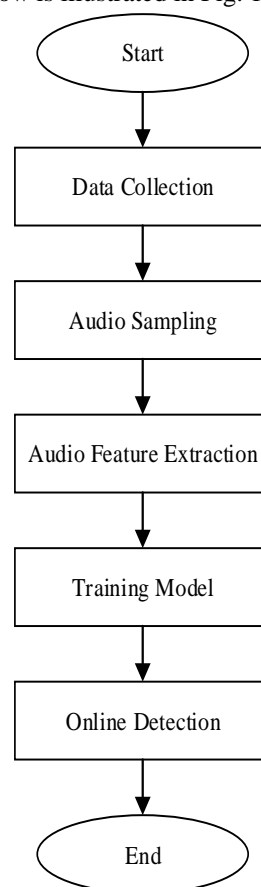


Fig. 1: Methods applied and flow process of a cough detection system

According to Fig. 1, the first method that was taken is to collect and acquire the data specifically the cough sound from several persons. Then, audio sampling was performed on each of the sounds and later, the desired feature data from the sounds were extracted. By having the data, the training process was performed on the model. Once the training process was completed, the model was tested to perform online cough detection from a person. A detailed discussion of each of these methods is discussed next.

A. Dataset collection and sampling

In this phase, cough sound data gathering is emphasized. This sound was captured with a handheld digital voice recorder by several persons. Each of the noises was captured for a length of five seconds. Currently, 500 sounds have been captured, and their details are shown in Table 1.

Table 1: Number of sample cough sounds recorded from persons by age categories

Gender	Kids (3-12) years old	Teens (13-19) years old	Adults (20-49) years old	Old Adults (50-80) years old
Male	50	65	75	56
Female	46	62	82	64
Total	96	127	157	120

The cough data were then kept in the database of the Mel-Frequency Cepstral Coefficients (MFCC) section in Edge Impulse software. Via the section, audio sampling is possible, and an example of the waveform yielding to one of the cough noises is depicted in Fig. 2. This procedure's sampling rate was set at 16 kHz to reduce aliasing during the Analog to Digital Conversion (ADC) process.



Fig. 2: A sample of the waveform from cough sound for the sampling process

B. Sound feature extraction

Once the sampling procedure was complete and the dataset was obtained, this step, namely the extraction of sound features, was carried out. MFCC were utilized for this purpose. In general, this method consists of the following steps: performing pre-emphasis, windowing the signal, executing Discrete Fourier Transform (DFT), applying Mel filter bank, calculating the logarithm of the magnitude, employing Inverse Discrete Fourier Transform (IDFT), executing dynamic feature, and finally obtaining transformed feature. Fig. 3 depicts the technique's flow, and the role of each phase is detailed as follows. According to the diagram, the first process is pre-emphasis, whose function is to amplify the sound energy at higher frequencies. This transforms the sound into higher formants, which improves detection accuracy. In the same procedure, filtering is used to equalize the frequency range of voice sounds with a quick roll-off of high-frequency sounds. The sound is hence less vulnerable to noise supplied later in the process. The windowing procedure follows. It tries to generate sound signal characteristics that

can be utilized to identify coughing. As expected, the sound signal is time-varying or quasi-stationary, with gradual fluctuations over time.

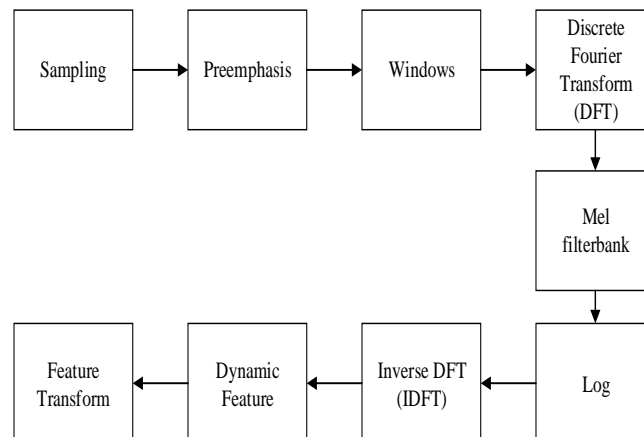


Fig. 3: Flow process of MFCC technique

To have stable acoustic characteristics, this sound must be analysed over an insufficiently long length of time. In this method, the sound analysis was conducted on a brief segment or frame when the signal is assumed to be stable. Later, the overlapping analysis was performed to verify that each speech sound in the input sequence is roughly centred at a frame. A Hanning or Hamming window was deployed for this purpose. A DFT was used to transform the signal from the time domain to the frequency domain in the subsequent procedure. Each window frame was transformed into a magnitude spectrum in this manner. As a result, the harmonics are enhanced, the edges become smoother, and the edge effect is diminished. Later, the Mel filter bank was implemented. It was used to convert the actual frequency into one that humans can sense. This procedure was required to enhance the efficiency of feature extraction. The subsequent stage in the procedure extracts the log from the output of the Mel filter bank, which was the power spectrum. This will aid in the reduction of insignificant acoustic signal variations and make the retrieved characteristics independent of one another.

The IDFT was then applied. The power spectrum was transferred to the time domain throughout this procedure. The output was known as cepstrum and comprises each frame's cepstral coefficients. In the final step of the dynamic feature process, the frame's energy was added to the cepstrum. In addition, characteristics pertaining to the evolution of cepstral features across time were included using both deltas and delta-deltas coefficients. First-order derivatives were represented by delta coefficients, whereas second-order derivatives were represented by delta-delta coefficients. Delta coefficients represent the pace of the signal, whereas delta-delta coefficients represent its acceleration. Finally, arrays containing the modified features were correctly collected from the audio stream produced.

C. Model Training

To train the data, Edge Impulse software was employed. Via the software, the acquired data were loaded into the program and the training process was executed. During this

process, the MFCC signal processing block was applied. It is a way of simplifying raw audio that typically contains a great deal of information that is not essential. Other audio processing blocks in Edge Impulse include the Mel Frequency Energy (MFE) and spectrogram blocks for handling non-voice audio. However, the MFCC block is very helpful when it comes to dealing with noises, including human speech, and coughing. The training model's accuracy and confusion matrix is depicted in the panel as in Fig. 4.

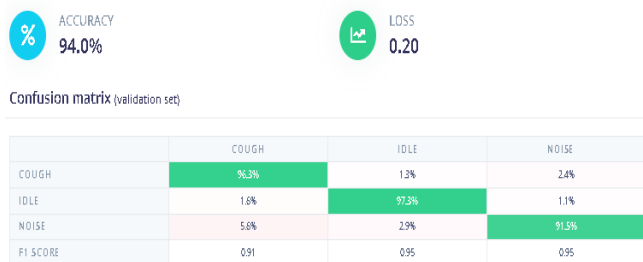


Fig. 4: Panel shows training performance

The trained model's accuracy is 94.0%. The greater the precision, the better. Although near-perfect accuracy is uncommon, it is often an indication that the model has to overfit the training data. For many applications, an accuracy of greater than 85% is considered suitable. As a result, the trained model's accuracy is regarded as acceptable. The confusion matrix, on the other hand, is a table that illustrates the proportion of windows that are properly classified versus those that are incorrectly labelled. According to the trained model's confusion matrix, 96.3% of the cough sounds were properly identified as cough, whereas 1.3% and 2.4% were misclassified. Following that, the results reveal that 97.3% of the idle sounds were accurately labelled as idle, whereas 1.6% and 1.1% were misclassified. Moreover, 91.5% were correctly identified as noise sounds, while 5.6% and 2.9% were misclassified. This looks to be an excellent outcome. In conclusion, all the data has been successfully trained.

Meanwhile, the output of the feature explorer resulting from the training of all data is depicted in Fig. 5.

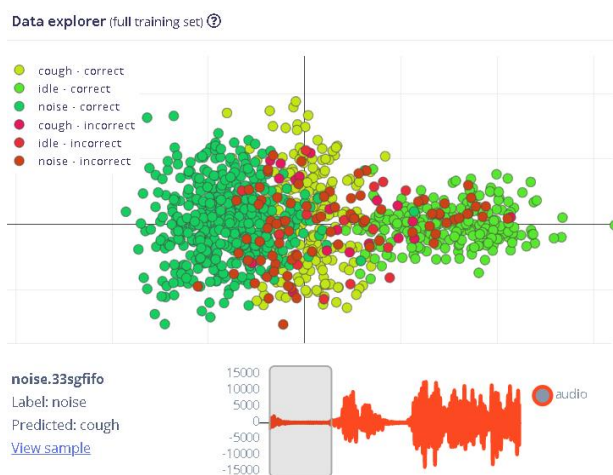


Fig. 5: Feature explorer resulted from model training

This feature explorer is beneficial since it displays the appearance of the two classes in a very clear and visible manner. It becomes a method for evaluating the quality of a dataset based on how well the data separates. Yellow, light green and dark green dots indicate the cough, idle and noise sounds are accurately classified and anticipated to respective categories. While the red, light orange and dark orange dots depict cough, idle and noise sounds that are improperly categorized and forecasted. Temporarily, at the bottom of the feature explorer, there is a cluster of audio files labelled as noise that was projected to be a cough category. Technically, this error comes from an audio problem, and it will not give a huge effect on the model. This error can be eliminated by retraining the model later so that precision can be increased.

D. Experimental Setup

To test the performance of the cough detection system to detect a person with a cough, an experiment platform was arranged as shown in Fig. 6. The platform was named a Modular and Open System (MOST) and had a dimension of 150 cm in length, 60 cm in width and 150 cm, in height. It is built using an aluminium pipe that has a diameter of 28 mm. To join the pipes to form the chassis of MOST as shown in the figure, specialized connectors were utilized. By using this connector, the chassis becomes modular and ease the process of modification when necessary.

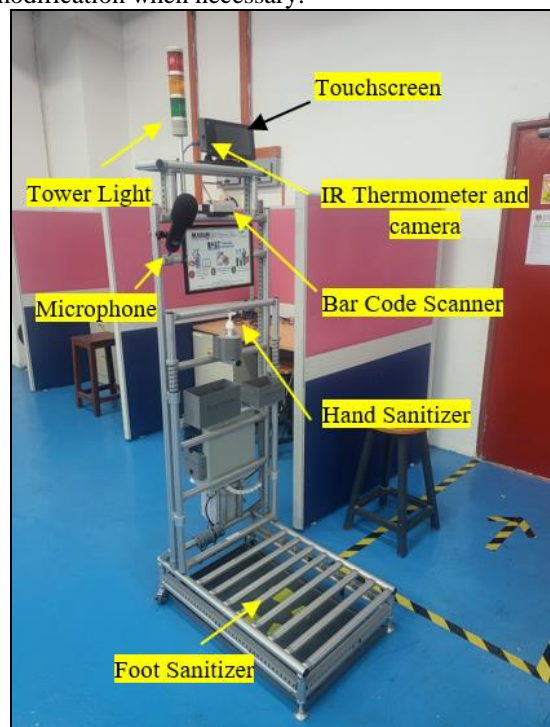


Fig. 6: Modular and Open System (MOST)

Meanwhile, several electronic devices were installed on this MOST which are a camera, infrared temperature sensor, 7 inches touchscreen, bar code scanner, tower light and Raspberry Pi 4B controller board. Having these devices on the MOST allows it to perform contactless temperature measurements, record information, and displays a user's temperature and health condition. More information and

details about the MOST device can be found in [10]. The detection system was then programmed into the controller board and integrated with the MOST system.

Once the MOST device is activated, a simple Human Machine Interface (HMI) of the system as depicted in Fig. 7 will be displayed.

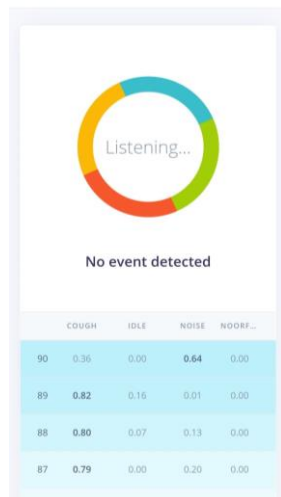
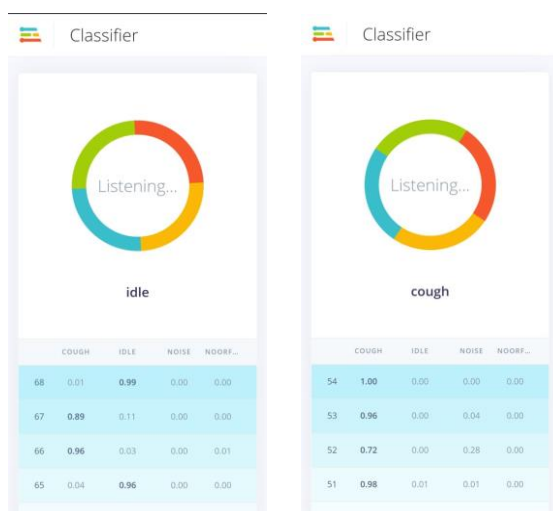


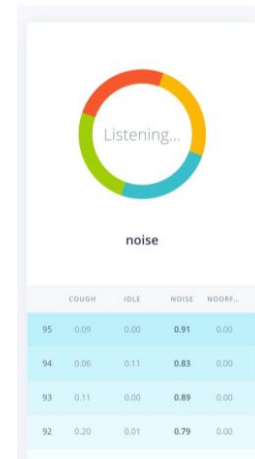
Fig. 7: HMI of live classification

In this HMI several features were arranged to show the real-time process of the system to perform the classification process. The first part is a circle progress bar with a listening word in it used to show the detection system ready to capture the sound. At the bottom, there is a phrase used to show the outcome of the detection and classification. It was set there only three categories of classification will be displayed. The first category is idle which means no cough sound is detected. The second category is a cough which indicates a cough sound detected. The last category is noise which means the sound cannot be recognized as either idle or sound. To identify the category of the captured sound, the system will analyse first using the training model. Then it will classify it to score a number for each category where the sum of the numbers is 1. The category with the highest score number will be shown. Fig. 8 shows examples of HMI display when it identified the sound category as idle, cough or noise.



(a) Idle

(b) Cough



(c) Noise

Fig 8: HMI classification for each sound category

To test the performance of the detection system, the MOST was placed at the entrance of a Dream Factory Lab, School of Electrical Engineering, College of Engineering, Universiti Teknologi MARA, Shah Alam, Selangor, Malaysia. Every person who come into the lab was required to scan staff or student ID to log attendance, go through temperature screening, and hand and foot sanitizing process using the MOST. Concurrently, the detection system also runs to detect a person with a cough. Figure 9 depicts such a setup that had been carried out as well as demonstrates a person using the MOST. At that moment, this system was arranged to function during office hours and log specific data, especially the score number for cough, idle, noise and uncertain sounds limited to 30 persons. This arrangement is intended to evaluate the performance of the model to perform the task of detecting cough



Fig. 9: Experimental setup

III. RESULTS AND DISCUSSION

The experiment was completed on that day when 30 persons in the category of adults had used the MOST device. Out of the 30 persons, 12 persons are men, and 18 persons are women. The score number of each sound category for every people are shown in Fig. 10.

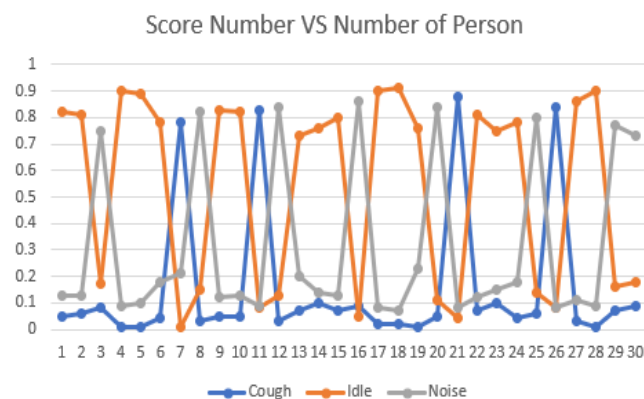


Fig. 10: The score number for each person obtained from the experiment

It can be spotted from Fig.10, the model detects cough from 4 persons and meanwhile, there were 18 persons who are not detected with cough. However, the model did show the detection results for 8 persons were noisy. In the meantime, the average score number that the model attained when detecting the cough category for 4 persons is 0.83. For the idle category that signifies no cough detected, the average score number for 18 persons is 0.78. As for the noise category, the average score number calculated for 8 persons is 0.8. With this outcome from the experiment, the detection model can be said only satisfactorily performed. Indeed, the model can detect a person with a cough or not, but it still has high detection error. This is due to the number of persons detected by the model as noise is quite high.

The cause of this problem was investigated, and it was identified that the noise of the surrounding experiment area had interfered with the accuracy of model detection. Hence, to overcome this problem, the model was arranged to be trained with additional sound data. These data consist of cough, idle and noise sounds captured at the experiment area. At that moment, the model had also verified whether it works or not to perform detection in that surrounding sound environment. Fig. 11 and Fig. 12 demonstrate the outcomes of the model training.

Last training performance (validation set)



Confusion matrix (validation set)

	COUGH	IDLE	NOISE
COUGH	98.9%	0.7%	0.5%
IDLE	1.9%	97.9%	0.3%
NOISE	3.5%	0.4%	96.1%
F1 SCORE	0.97	0.98	0.98

Fig. 11: Panel shows training performance after model retraining

Data explorer (full training set)

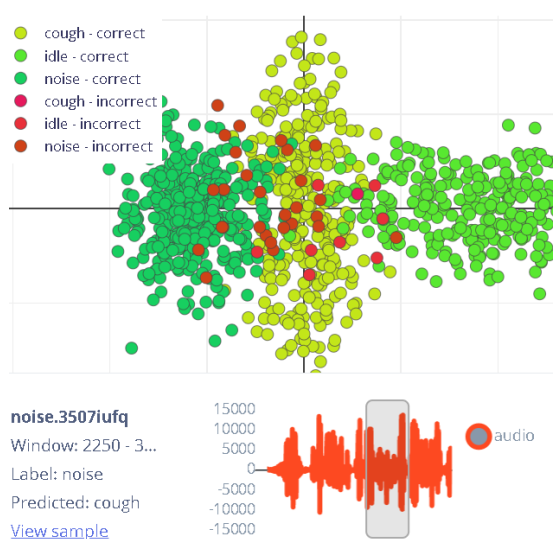


Fig. 12: Feature explorer resulted from model retraining

According to Fig. 10, the detection performance accuracy of the model was increased by 97.5%. This signifies that the model at that time was very capable to differentiate the sound of cough, idle, or noise in the experiment area. Seemingly, it is difficult for the model to achieve perfection since there is still slight error. This is shown by the confusion matrix in the figure. About 98.9% of the cough audio sounds were correctly detected as cough, whereas 0.7% and 0.5% were misclassified respectively as idle and noise. For idle, 97.9% of the idle audio windows were correctly labelled whereas only 1.9% and 0.3% were incorrectly labelled. At last, 96.1% of the samples were accurately categorized as noise, with just 3.5% and 0.4% mislabelled. With this data, the model was justified to be successfully trained in the experiment area.

Fig.12 presents the feature explorer results after the model has been retrained using the new data. It can be recognized, this time, the sounds of cough, idle, and noise are appropriately categorized as represented respectively by the yellow, light green, and dark green dots. As for the red, light orange, and dark orange dots that signifies as the sound of the cough, idle, and noise, they are poorly classified and predicted. Besides the number of dots turns out to be less as compared to the feature explorer output in Fig. 5. Temporarily, there was still a collection of audio files labelled

as noise that was expected to be cough and can be found below the feature explorer. It is expected that this matter is still caused by the audio problem. Even though the model was retrained, and adjustments were done to the microphone and controller interfacing settings, this problem still occurred. Yet, the error is still acceptable since it does not have a huge effect on the model when performing the detection.

The experiment was then repeated to evaluate again the detection model after going through the retraining process. The same number of persons and operating hours of MOST used in the previous experiment was employed again. For this experiment, 14 males and 16 females were tested. Fig. 13 displays the score data of each sound category for every person. For this experiment, the model could detect 4 persons with cough and 24 persons were detected as idle which represents no cough detected. Only 2 persons were detected and labelled as noise. Meanwhile, the average score number for the model when detecting cough for 4 persons is 0.91. The idle category has an average score number of 0.91 calculated for 24 persons. At last, average score number for noise category is 0.86. This outcome is better than the previous experiment. Now, the model can perform better detection and the error is still acceptable.

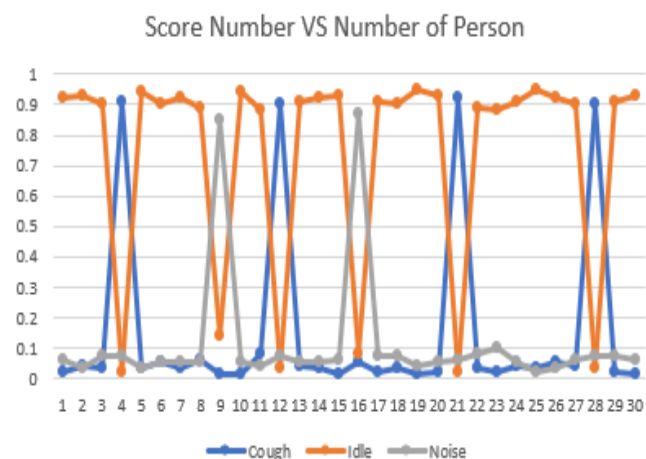


Fig. 13: The Score number for each person obtained from the experiment after the model retrained

IV. CONCLUSION

In this research, a model was successfully developed to have a function to detect a person with a cough. Such a function was done by using a sound captured from a microphone and later fed to the model to analyse and determine whether the sound is cough or vice versa. This model was embedded in the MOST system which enhances the device functionality which is now not limited to log attendance, performing temperature screening and sanitizing processes. The development of this model has gone through the collection of cough sound samples, sound feature extraction, model training, testing and evaluation to ensure high accuracy performance before it can perform for real-time application. However, more data are needed from a range of nations, localities, ages, genders, and ethnicities for the model's generalizability and its efficiency to perform the detection. In the meantime, the completion of this research provides a groundwork for future research. Ever since cough

is one symptom for those infected by COVID-19, hence it is our intention to upgrade this model to identify a person infected by COVID-19 or not using cough sound.

V. ACKNOWLEDGEMENT

This research was supported by a fund from Universiti Teknologi MARA (UiTM) and Malaysia Automotive, Robotics and IoT (MARii), 100-RMC 5/3/SRP (015/2020). The authors would like to thank UiTM, MARii, Innovative Electromobility (ITEM) Research Lab and College of Engineering, UiTM for providing the financial supports, equipment, and place to conduct this research.

VI. REFERENCES

- [1] J.M. Liu, M. You, Z. Wang, G.Z. Li, X. Xu, and Z. Qiu, "Cough detection using deep neural networks," IEEE International Conference on Bioinformatics and Biomedicine (BIBM), pp. 560-563, November 2014.
- [2] K. S. Alqudaihi, N. Aslam, I. U. Khan, A. M. Almuhaideb, S. J. Alsunaidi, N. M. A. R. Ibrahim, F. A. Alhaidari, F. S. Shaikh, Y. M. Alsenbel, D. M. Alalharith, H. M. Alharthi, W. M. Alghamdi and M. S. Alshahrani, "Cough sound detection and diagnosis using artificial intelligence techniques: challenges and opportunities," IEEE Access, vol. 9, pp. 102327-102344, 2021.
- [3] W. Shin, "Implementation of cough detection system using IoT sensor in respirator," International Journal of Advanced Smart Convergence, vol. 9, pp. 132-138, 2020.
- [4] J. Amoh and K. Odame, "Deep Neural Networks for Identifying Cough Sounds," IEEE Transactions on Biomedical Circuits and Systems, vol. 10, no. 5, pp. 1003-1011, Oct. 2016.
- [5] S.A. Mahmoudi, P.D. C. Possa, T. Ravet, T. Drugman, R. Chessini, T. Dutoit, and C. Valderrama, "Sensor-based system for automatic cough detection and classification," ICT Innovations Conference, September 2016.
- [6] S. H. Shin, T. Hashimoto, and S. Hatano, "Automatic detection system for cough sounds as a symptom of abnormal health condition," IEEE Transactions on Information Technology in Biomedicine, vol. 13, no. 4, pp. 486-493, July 2009.
- [7] A. Rana, Y. Dhiman and R. Anand, "Cough Detection System using TinyML," IEEE International Conference on Computing, Communication and Power Technology (IC3P), pp. 119-122, 2022.
- [8] S. Khomsay, R. Vanijjirattikhan, and J. Suwatthikul, "Cough detection using PCA and Deep Learning," IEEE International Conference on Information and Communication Technology Convergence (ICTC), pp. 101-106, October 2019.
- [9] F. Barata, K. Kipfer, M. Weber, P. Tinschert, E. Fleisch and T. Kowatsch, "Towards Device-Agnostic Mobile Cough Detection with Convolutional Neural Networks," IEEE International Conference on Healthcare Informatics (ICHI), pp. 1-11, 2019.
- [10] N.A. Razak, M. R. B. Sazali, J. Johari, N.A. M. Salleh, Y. H. P. Manurung and N. K. Mun, "Design and Development of Modular and Open System (MOST) for Pandemic Precaution," In Proceedings of the 8th International Conference on Computational Science and Technology (ICCST) pp. 751-763, August 2021.