Online-Journals.org

iJOE International Journal of Online and Biomedical Engineering

iJOE | eISSN: 2626-8493 | Vol. 19 No. 12 (2023) | OPEN ACCESS

https://doi.org/10.3991/ijoe.v19i12.40471

PAPER

Speech Recognition Algorithms-Based Cough Recognition System

Fatima Barkani[1], Mohamed Hamidi[2](✉), Ouissam Zealouk[1], Hassan Satori[1]

[1]Faculty of Sciences Dhar Mahraz, Laboratory of Computer Science, Signals, Automation and Cognitivism, USMBA, Fez, Morocco

[2]Pluridisciplinary Faculty of Nador, Team of Modeling and Scientific Computing, UMP, Oujda, Morocco

mohamed.hamidi.5@gmail.com

ABSTRACT

This paper introduces an innovative technique for creating a cough detection system that relies on speech recognition algorithms. The strategy utilizes the Kaldi platform, which is open source and incorporates a hybrid system of Gaussian Mixture Model-based Hidden Markov Models (GMM-HMM) through a straightforward monophone training model. Additionally, the study examines the effectiveness of two different feature extraction approaches, Mel Frequency Cepstral Coefficient (MFCC) and Perceptual Linear Prediction (PLP). The proposed system can function as a collection tool for gathering natural and spontaneous cough data from conversations or continuous speech. The paper also compares the Kaldi and CMU Sphinx4 toolkits, concluding that Kaldi's use of GMM-HMM outperforms CMU Sphinx4.

KEYWORDS

Cough recognition, HMM-GMM, Speech recognition, MFCC, PLP

1 INTRODUCTION

Cough is an important component of daily health monitoring, serving as a means to expel foreign bodies like irritants, fluids, and microbes from the respiratory system by forcibly expelling air through the throat. It is a symptom of numerous diseases, such as respiratory illnesses, asthma, and COVID-19 [1–3]. According to a study [4], there are three principal phases of coughing: (1) inspiration, (2) closure of the glottis with the pressure of air in the lower respiratory tract, and (3) ejection of air through the glottis. In recent years, scientists have been working on creating technology capable of detecting coughing sounds through a range of techniques, including audio sensors, machine learning algorithms, and artificial intelligence.

In a research study conducted by Miranda et al. [5], the authors compared various acoustic features for automatic cough detection using deep learning architectures. They utilized DNN, CNN, and LSTM classifiers to evaluate the performance of the MFCC, littered-MFCC, STFT, and MFB features. In another study [6], the researchers utilized two distinct cough acoustic models, which were created from both band-pass

Barkani, F., Hamidi, M., Zealouk, O., Satori, H. (2023). Speech Recognition Algorithms-based Cough Recognition System. *International Journal of Online and Biomedical Engineering (iJOE)*, 19(12), pp. 49–61. https://doi.org/10.3991/ijoe.v19i12.40471

Article submitted 2023-04-15. Resubmitted 2023-06-03. Final acceptance 2023-06-06. Final version published as submitted by the authors.

filtered and unfiltered signals, to develop a cough detection system. Drugman, et al. [7] conducted the classification of acoustic sounds as either event of cough or non-cough using a logistic regression model with three spectral features. The results showed an F1-score of 88.70%, a sensitivity of 90.31%, and a specificity of 98.14%. According to a research paper [8], a deep neural network approach is used for detecting coughs. The method has a specificity of 92.7% and a sensitivity of 87.7%. Some researchers have created a cough detection system that utilizes audio signals to automatically count coughs [9]. The main goal of their work was to attain a high level of precision, as evidenced by reported sensitivity and specificity values of 94.7% and 95%, respectively. Botha et al [10] demonstrated that the disaggregation of cough data can be an efficient and cost-effective approach for tuberculosis detection. Their methodology achieved a sensitivity of 95% and a specificity of around 72%, indicating its potential for successful implementation. Yin et al. [11] explored two methods of combining data to enhance classification accuracy. These methods were feature fusion and classifier fusion. To improve the feature fusion algorithm, they selected better acoustic and image features. Additionally, they proposed a novel classifier fusion algorithm, which combined support vector machine (SVM) classifiers trained using acoustic and deep features using soft voting. Their research focused on predicting pig coughs, and they validated their methods using sound data collected from a pig barn. Their proposed methods resulted in a high classification rate of 97.47% and 99.20% for feature fusion and classifier fusion, respectively. In another research [12], the authors introduced an efficient and adaptable method for detecting symptomatic patterns in biological audio signals, specifically cough audio files. The proposed method involves utilizing a stationary wavelet transform (SWT) for spectral analysis of the audio files while addressing the class imbalance problem through the use of the ADASYN technique. The model also extracts various features such as Mel-frequency cepstral coefficients (MFCCs), log frame energies, zero crossing rate (ZCR), and kurtosis. The deep belief network (DBN) model is employed for classification purposes, and optimal hyper-parameter tuning is achieved through the use of the mayfly optimization (MFO) algorithm. In another study [13], the authors described a machine hearing system that is designed for robust cough segmentation using audio signals and can be easily deployed in mobile environments. The system involves two stages of cough detection. In the first stage, a short-term spectral feature set is computed in five predefined frequency bands to ensure the feature set's robustness in various noisy scenarios. Then, feature selection and combination techniques are applied to obtain a high-level representation of the data by calculating the mean and standard deviation of the short-term descriptors in long-term frames of 300ms. In the second stage, a support vector machine is used for cough detection, which is trained with data from various noisy scenarios. The system's performance is evaluated using a patient signal database, which emulates three real-life scenarios in terms of noise content. The results show that the proposed system outperforms state-of-the-art methods, achieving 92.71% sensitivity, 88.58% specificity, and 90.69% Area Under Receiver Operating Characteristic (ROC) curve (AUC). Rana et al. [14] describe the development of a system for cough detection using Edge Impulse Studio and Arduino 33 BLE Sense. The system is capable of differentiating between genuine cough sounds and other unwanted background noises. The results show that the system achieved a recognition accuracy of nearly 97%. In another study [15], researchers created a cough detection system using cough sound data gathered through the Arduino 33 BLE Sense and Edge Impulse for development. The microphone of the Arduino Nano 33 BLE Sense captured a total of 57 cough sounds which were then processed for classification using machine-learning tools provided by Edge Impulse. The cough sounds were labeled into four categories, namely cough, cough with noise, silence, and noisy, and were classified using a neural network in Edge Impulse. The results showed an accuracy of 73.6% and a loss of 0.86 for the classification. After analyzing related works, the present study introduces a novel approach for creating a cough detection system, utilizing the open-source software Kaldi. In addition, this study presents an open-source framework for assessing the system's efficiency by conducting experiments with various HMMs, GMMs, and feature extraction methods. The objective is to identify the most effective values that lead to optimal performance.

This paper is structured in the following manner: Section 1 is an introduction, Section 2 provides an overview of coughing, and Section 3 gives a brief explanation of the Kaldi toolkit, and the Hidden Markov Model is explained in Section 4. Section 5 presents the proposed method, while Section 6 discusses the experimental results. Section 7 is dedicated to the discussion, while Section 8 provides a comparison. Finally, Section 9 contains the conclusion.

2 COUGH OVERVIEW

The natural process of coughing aids to expel foreign objects from the lungs. Sensory shots in the airways are triggered by the presence of foreign objects or inflammation, sending a signal to the brainstem which activates the respiratory muscles and diaphragm in a coordinated manner to move air through the airways (see Figure 1 [16]).

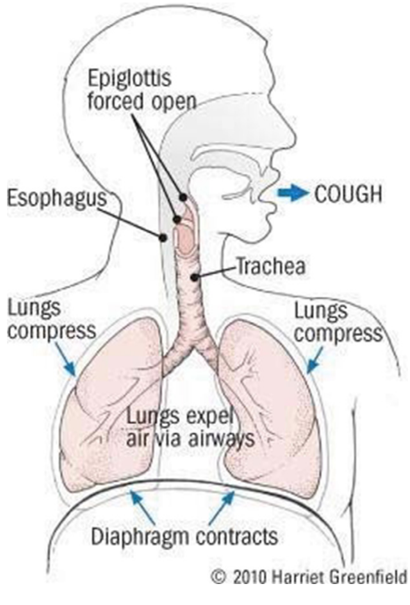

Fig. 1. The process of producing cough [16]

3 KALDI

Kaldi is a speech recognition software that is published under the Apache License version 2.0. It is developed in C++ and comes with a comprehensive set of tools and programs including HMMs, decision trees, neural networks, data preprocessing and feature extraction capabilities [17]. A graphical representation of their internal structure is provided in Figure 2. Our study involves the utilization of the Kaldi toolkit to create a cough detection system. The system is designed to analyze audio data, extract acoustic features, model training, and ultimately classify cough sounds. Through this process, we are able to accurately identify and distinguish coughs from other sounds.

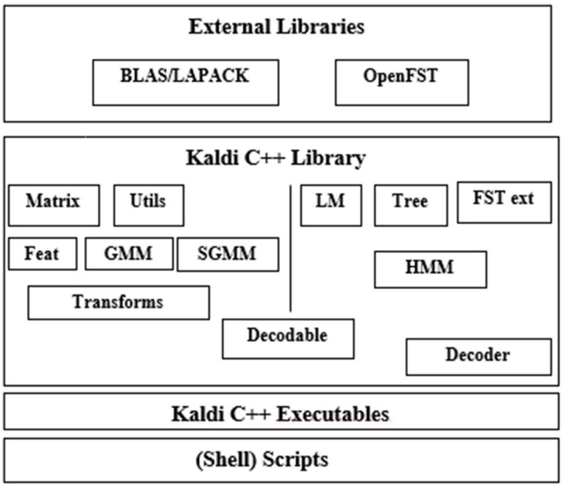

Fig. 2. Kaldi architecture [17]

4 HIDDEN MARKOV MODELS (HMMS)

The Hidden Markov Model (HMM) was originally introduced in the 1960s [18] and has since become widely utilized as a speech recognition modeling technique [19]. In addition, it has also found application in computational molecular biology [20]. A visual representation of a three-state HMM topology is shown in Figure 3.

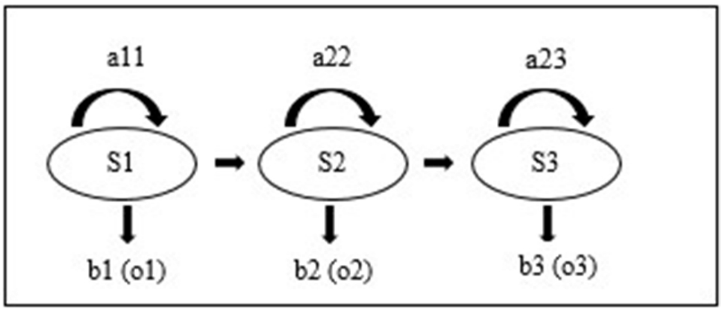

Fig. 3. The 3 states of HMM architecture [21]

5 COUGH DETECTION SYSTEM ARCHITECTURE

Our study introduces an innovative approach to developing a cough detection system that exploits speech-recognition algorithms. The system is designed using the Kaldi open-source platform, which utilizes HMMs methodology and a range of GMMs. Furthermore, we employ MFCCs and PLP feature extraction techniques to enhance the system's accuracy. Figure 4 illustrates the proposed System Architecture. The development of a system to detect cough sounds would involve several steps. Firstly, the system would need a microphone to receive an audio input of the cough sound. Once the audio input has been obtained, the system would then need to divide the input into segments. Next, the system would create feature vectors from each segment of the input. After the creation of feature vectors, the system would subsequently analyze and compare them to a pre-existing database of known coughs. This process would allow the system to identify and detect the occurrence of a cough signal accurately.

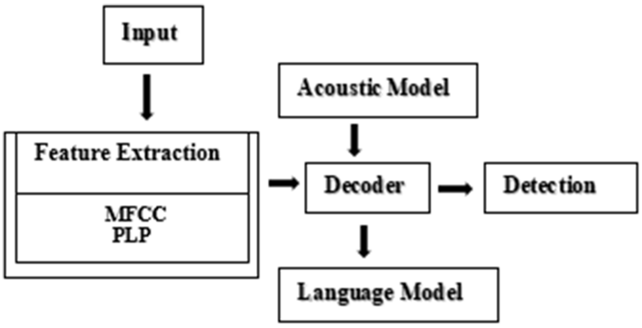

Fig. 4. Proposed system architecture

5.1 Corpus

The corpus utilized in this study is unique to the laboratory where it was generated. It comprises three types of audio samples: coughs, speech, and laughter. The primary focus of this research is the development of a cough detection system. For the speech data, seven individuals who are native Tarifit Berber speakers were enlisted to provide the audio samples used in this paper. Furthermore, the laughter data was utilized to evaluate the system's performance. The audio data was captured using a microphone and recorded using the WaveSurfer recording tool in the wave format. The recorded data was then saved into a single file with the extension ".wav". A sampling rate of 16 kHz and a resolution of 16 bits were employed during the recording process. Additionally, two separate sets of audio files were created for training and testing purposes in this work (show Table 1).

Table 1. Corpus characteristics

Recorder Type	Number of Recorders Used for Training	Number of Recorders Used for Testing
Cough	28	12
Speech	14	6
Laughter	0	6

5.2 Feature extraction methods

The feature extraction phase allows the extraction of characteristics that make it possible to discern the components of the audio signal that are relevant for the identification of linguistic content, by rejecting the other information contained in that signal. In this study, the used feature extraction methods are the Perceptual Linear Prediction (PLP) model [22] and Mel Frequency Cepstral Coefficients (MFCCs) [23].

5.3 Acoustic model

The acoustic model represents the statistical features of the audio signal, with the most significant being used for analysis. The acoustic model is typically trained on large data and uses statistical techniques, such as HMMs or neural networks. In this investigation, the acoustic model utilized for detecting coughing actions employs a simple monophone model trained using a GMM-HMM combining model.

6 EXPERIMENTAL RESULTS

The designed system was trained and generated using an Ubuntu 16.04 LTS (64-bit operating system). We have chosen the open-source Kaldi for the implementation of our system to detect the cough. Also, our system realized 4 experiments, namely Experiment 1, Experiment 2, Experiment 3, and Experiment 4. In Experiment 1, the system was trained and tested only on cough sounds. In Experiment 2, the system was trained and tested on both cough and speech sounds. In Experiment 3, the system was trained on both cough and speech sounds and tested only on cough sounds. In Experiment 4, the system was trained only on cough sounds and tested on a combination of cough and laughter sounds. Table 2 provides information about the training and testing data utilized in our experiments. These tables present specific details about the datasets used in our study.

To create an effective detection system, we conducted training and testing using various numbers of GMMs (200, 400, and 600) and two feature extraction techniques: MFCC and PLP. Our system's ability to recognize coughs was evaluated based on the recognition rate. The cough recognition rate refers to the accuracy of a system or model in detecting and classifying cough sounds. It is determined by calculating the ratio of correctly identified cough instances to the total number of cough instances in a given dataset or evaluation scenario.

To calculate the cough recognition rate, we define the following terms:

- True Positives (TP): The number of cough instances correctly identified by the system.
- False Negatives (FN): The number of cough instances missed or incorrectly classified as non-cough sounds by the system.

The cough recognition rate, usually expressed as a percentage, can be computed using the following formula:

$$\text{Cough Recognition Rate} = (TP / (TP + FN)) * 100$$

Table 2. The data set used in training and testing

Exp ID	Cough		Speech		Laughter	
	Train	Test	Train	Test	Train	Test
1	28	12	0	0	0	**0**
2	14	6	14	6	0	**0**
3	14	12	14	0	0	**0**
4	**28**	**6**	**0**	**0**	**0**	**6**

Figure 5 presents the results of the first experiment where the system underwent training and testing using the cough data with MFCC and PLP Coefficients, along with Gaussian mixture models ranging from 200 to 600. For the MFCC Coefficients, the cough recognition rate was 98.33% for both 200 and 400 GMMs, and corresponded to 99.16% when using 600 GMMs. On the other hand, for PLP coefficients, the recognition rate was 94.16% for both 200 and 400 GMMs, and rose to 97.5% when utilizing 600 Gaussian mixture distributions. The most exceptional cough recognition rate was achieved with the MFCC feature in combination with 600 Gaussian mixture distributions. The obtained results of the second experiment are presented in Figures 6 and 7. The experiment involved using MFCC and PLP features with different numbers of Gaussian mixture distributions to identify cough and speech sounds. For MFCC features, the system's correct rates for 200, 400, and 600 Gaussian mixture distributions were reported. The correct rates for cough and speech were 91.66% and 78.99%, 93.33% and 80.66%, and 96.66% and 80.83%, respectively. For PLP features, the correct rates for 200, 400, and 600 Gaussian mixture distributions were also reported, and the rates for cough and speech were 91.66% and 78.03%, 96.66% and 79.69%, and 98.33% and 82.16%, respectively. Figures 6 and 7 show that the recognition rate for cough was better than that for speech. Figure 8 presents the results obtained from the third experiment. When 200 Gaussian mixture distributions were used, the correct rate of the system was 87.5% and 86.66% for MFCC and PLP features, respectively. When 400 Gaussian mixture distributions were used, the recognition rates increased to 89.16% for both MFCC and PLP features. For 600 Gaussian mixture distributions, the system performances were 90.83% and 95% for MFCC and PLP features, respectively. The highest recognition rate for cough was obtained with PLP features in combination with 600 Gaussian mixture distributions. In the fourth experiment, the system's performance was evaluated for cough and laughter using MFCC and PLP features with 200, 400, and 600 Gaussian mixture distributions. The results showed that for MFCC features, the system's correct rate for 200 Gaussian mixture distributions was 96.66% for cough and 78.33% for laughter. For 400 Gaussian mixture distributions, the correct rate was 96.66% for cough and 95% for laughter. When 600 Gaussian mixture distributions were used, the correct rate increased to 98.33% for cough and 95% for laughter. With PLP features, the system's performance was slightly lower than with MFCC features. For 200 Gaussian mixture distributions, the correct rate was 90% for cough and 85% for laughter. For 400 Gaussian mixture distributions, the correct rate was 91.66% for cough and 88.33% for laughter. When 600 Gaussian mixture distributions were used, the correct rate remained at 91.66% for cough, and 88.33% for laughter. The recognition rate for cough was consistently better than for laughter, as shown in Figures 9 and 10.

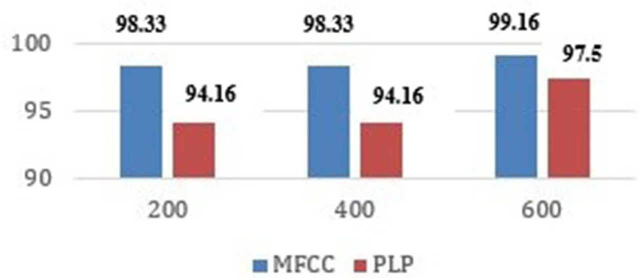

Fig. 5. Experiment 1: Cough recognition rates based on the number of GMMs used for MFCC and PLP features

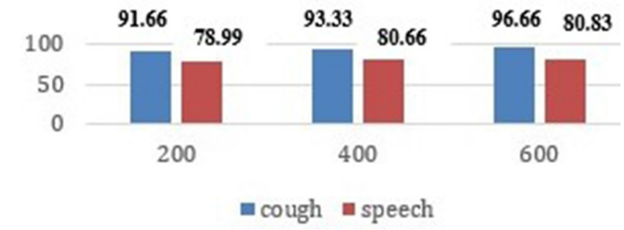

Fig. 6. Experiment 2: Comparing the accuracy of cough and speech signal recognition using MFCC

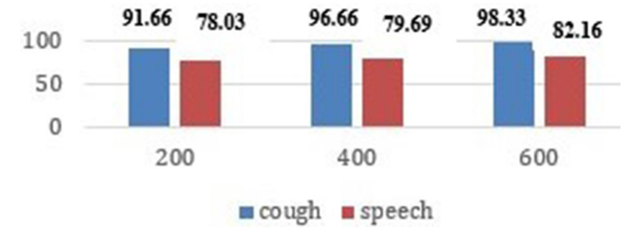

Fig. 7. Experiment 2: Comparing the accuracy of cough and speech signal recognition using PLP

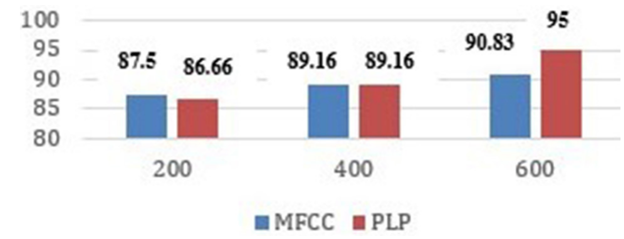

Fig. 8. Experiment 3: Cough recognition rates based on the number of GMMs used for MFCC and PLP features

Fig. 9. Experiment 4: Comparing the accuracy of cough and laughter signal recognition using MFCC and GMMs

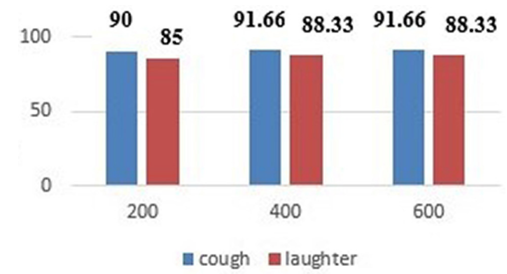

Fig. 10. Experiment 4: Comparing the accuracy of cough and laughter signal recognition using PLP and GMMs

7 DISCUSSION

Based on our analysis and experience, we have found that using 600 GMMs produced the best results. Additionally, we found that the performance of MFCC coefficients was better for detecting cough sounds.

Figures 5, 6, 7, 8, and 10 present the recognition rate differences of cough, speech, and laughter using the MFCC and PLP features with varying total Gaussian distribution numbers. In Figure 5, the difference in recognition rates between MFCC and PLP for cough was found to be between 1.66 and 4.17%. Figure 6 showed that the recognition rate difference between cough and speech using the MFCC feature was between 12.67 and 15.83%, with the highest difference occurring at 600 GMMs. Similarly, Figure 7 demonstrated that the recognition rate difference between cough and speech using the PLP feature was between 13.63 and 16.97%, and the highest difference occurred with 400 GMMs. Figures 9 and 10 illustrated the recognition rate differences between cough and laughter using the MFCC and PLP features, respectively. The difference in recognition rates between cough and laughter using the MFCC feature was between 1.66 and 18.33%, while using the PLP feature, the difference ranged between 3.33 and 5%. Overall, these results indicate that the system is capable of detecting coughs based on the differences in recognition rates between a cough and other vocalizations such as speech and laughter.

8 COMPARISON

The primary objective of this research is to create an automated cough recognition system using the open-source Kaldi platform. To ensure that the system performs well, the research employs HMM-GMM acoustic models with varying numbers of Gaussians (8, 16, 32, and 64 GMMs) and MFCC coefficients trained with both Kaldi and CMU Sphinx4 tools. The study aims to compare the recognition rate of the different models. For training the system, both cough and speech samples were used, and the system was tested using cough sounds. In the cough dataset, a total of 7 recorders were used for training and 7 were used for testing. On the other hand, in the speech dataset, 7 recorders were used for training. Figure 11 shows the recognition rate (%) of cough in the function of GMMs 8, 16, 32, and 64. In the case of the Kaldi toolkit the result achieved using 8 GMMs is 91.4 %, the result obtained using 16 GMMs and 32 GMMs is 92.85 % and the result of 64 GMMs is 94.28%. On the other hand, in the case of CMU Sphinx4 tools, the result given are 90%, 92.85%, 88.57%, and 82.85% for 8, 16, 32, and 64 GMMs, respectively. The difference between the achieved recognition rates of cough based on four experiments is as follows: For the case of 8 GMMs, the difference is 1.4%, for 16 GMMs, the difference is 0%, while, the difference for 32 GMMs is 4.28%, finally in the case of 64 GMMs the difference is 11.43%. The results obtained clearly indicate Kaldi definitely outperformed CMU Sphinx4 with the use of the GMM-HMM. The accuracy of Kaldi is significantly higher than that of CMU Sphinx4, which suggests that Kaldi has a better ability to detect coughs.

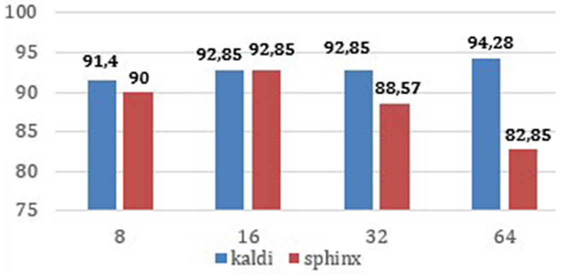

Fig. 11. The difference in recognition rate (%) for cough between Kaldi and Sphinx4

Table 3. The data set used in training and testing

Reference	Year	Methods	Results
[24]	2023	CNN	88.8% accuracy 91.2% sensitivity 86.5% specificity 86.5% precision
[25]	2023	DWT DCNN	DWT achieves better accuracy
[26]	2023	SVM KNN	99.6%
[27]	2023	SVM	96%
[28]	2022	ANN CNN	98.1% 98.5%
[29]	2022	MFCC kNN	80%
[30]	2022	GMM-HMM MFCC	91.57%
[31]	2020	ANNs	94.7% of sensitivity 95% of specificity
[32]	2020	CNN MFCC	Accuracy 70.58% Sensitivity 80.95%
Proposed work	2023	GMM-HMM MFCC PLP	Cough: 99.16% (MFCC features, 600 GMMs) Laughter: 95% (MFCC features, 600 GMMs) Speech: 82.16% (PLP features, 600 GMMs)

We have conducted a comparison of our proposed approach with other existing works. As presented in Table 2, our approach differs significantly from the methods employed by other researchers in several aspects. One notable difference was the remarkably high accuracy score achieved by our approach, which was measured at 99.16%. Moreover, our approach utilized a unique method by combining Hidden Markov Models and Gaussian Mixture Models. Additionally, we employed two distinct feature extraction techniques, namely MFCC and PLP. Overall, our comparison results validate the efficacy of our proposed approach (See Table 3).

9 CONCLUSION

Our study proposes a novel method for detecting coughs using speech-recognition algorithms. This approach was developed using the Kaldi open-source platform and employed a hybrid GMM-HMM model along with a simple monophone training model. In order to assess the effectiveness of the system, we employed two distinct feature extraction methods, namely PLP and MFCC. We evaluated the performance of these methods using a comprehensive corpus of speech, coughs, and laughter with varying characteristics. The results of our study demonstrate the potential of our approach in accurately detecting coughs. We found that both PLP and MFCC feature extraction methods yielded satisfactory performance levels in detecting coughs, with MFCC achieving slightly better results, with a recognition rate of 99.16%, compared to 98.33% for PLP. Our study builds upon the concept initially introduced in a previous publication [33]. In our future work, our primary objective is to improve our cough detection system through the implementation of hybrid and deep learning techniques. We aim to incorporate these approaches to enhance the accuracy and reliability of our system.

10 REFERENCES

[1] Lakshmanarao, A., Babu, M. R., & Kiran, T. (2021). An efficient covid19 epidemic analysis and prediction model using machine learning algorithms. *International Journal of Online & Biomedical Engineering*, *17*(11), 176–184. https://doi.org/10.3991/ijoe.v17i11.25209

[2] Van, N. T., Irum, S., Abbas, A. F., Sikandar, H., & Khan, N. (2022). Online learning—two side arguments related to mental health. *International Journal of Online & Biomedical Engineering*, *18*(9), 131–143. https://doi.org/10.3991/ijoe.v18i09.32317

[3] Ali, L. R., Jebur, S. A., Jahefer, M. M., & Shaker, B. N. (2022). Employing transfer learning for diagnosing COVID-19 disease. *International Journal of Online & Biomedical Engineering*, *18*(15), 31–42. https://doi.org/10.3991/ijoe.v18i15.35761

[4] Yanagihara, N., Von Leden, H., & Werner-Kukuk, E. (1966). The physical parameters of cough: the larynx in a normal single cough. *Acta Oto-Laryngologica*, *61*(1–6), 495–510. https://doi.org/10.3109/00016486609127088

[5] Miranda, I. D., Diacon, A. H., & Niesler, T. R. (2019, July). A comparative study of features for acoustic cough detection using deep architectures. In *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 2601–2605). IEEE. https://doi.org/10.1109/EMBC.2019.8856412

[6] Takahashi, S. Y., Morimoto, T., Maeda, S., & Tsuruta, N. (2004). Cough detection in spoken dialogue system for home health care. In *Eighth International Conference on Spoken Language Processing*. https://doi.org/10.21437/Interspeech.2004-605

[7] Drugman, T., Urbain, J., Bauwens, N., Chessini, R., Aubriot, A. S., Lebecque, P., & Dutoit, T. (2020). Audio and contact microphones for cough detection. *arXiv preprint arXiv:2005.05313*.

[8] Pramono, R. X. A., Imtiaz, S. A., & Rodriguez-Villegas, E. (2019, July). Automatic cough detection in acoustic signal using spectral features. In *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 7153–7156). IEEE.

[9] Amoh, J., & Odame, K. (2016). Deep neural networks for identifying cough sounds. *IEEE Transactions on Biomedical Circuits and Systems*, *10*(5), 1003–1011. https://doi.org/10.1109/TBCAS.2016.2598794

[10] Botha, G. H. R., Theron, G., Warren, R. M., Klopper, M., Dheda, K., Van Helden, P. D., & Niesler, T. R. (2018). Detection of tuberculosis by automatic cough sound analysis. *Physiological Measurement*, *39*(4), 045005. https://doi.org/10.1088/1361-6579/aab6d0

[11] Yin, Y., Ji, N., Wang, X., Shen, W., Dai, B., Kou, S., & Liang, C. (2023). An investigation of fusion strategies for boosting pig cough sound recognition. *Computers and Electronics in Agriculture*, *205*, 107645. https://doi.org/10.1016/j.compag.2023.107645

[12] Ayappan, G., & Anila, S. (2023). Mayfly optimization with deep belief network-based automated COVID-19 cough classification using biological audio signals. *Cybernetics and Systems*, *54*(6), 767–786. https://doi.org/10.1080/01969722.2023.2166244

[13] Monge-Álvarez, J., Hoyos-Barceló, C., San-José-Revuelta, L. M., & Casaseca-de-la-Higuera, P. (2018). A machine hearing system for robust cough detection based on a high-level representation of band-specific audio features. *IEEE Transactions on Biomedical Engineering*, *66*(8), 2319–2330. https://doi.org/10.1109/TBME.2018.2888998

[14] Rana, A., Dhiman, Y., & Anand, R. (2022, January). Cough detection system using TinyML. In *2022 International Conference on Computing, Communication and Power Technology (IC3P)* (pp. 119–122). IEEE. https://doi.org/10.1109/IC3P52835.2022.00032

[15] Wardhany, V. A., Hidayat, A., Utami, S. W., & Bastiana, D. S. (2022, December). Arduino nano 33 BLE sense performance for cough detection by using NN classifier. In *2022 6th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE)* (pp. 455–458). IEEE. https://doi.org/10.1109/ICITISEE57756.2022.10057829

[16] Harvard Medical School. https://www.health.harvard.edu/staying-healthy/that-nagging-cough, accessed 02-12-2022.

[17] Povey, D., Ghoshal, A., Boulianne, G., Burget, L., Glembek, O., Goel, N., ... & Vesely, K. (2011). The Kaldi speech recognition toolkit. In *IEEE 2011 Workshop on Automatic Speech Recognition and Understanding* (No. CONF). IEEE Signal Processing Society.

[18] Baum, L. E., & Petrie, T. (1966). Statistical inference for probabilistic functions of finite state Markov chains. *The Annals of Mathematical Statistics*, *37*(6), 1554–1563. https://doi.org/10.1214/aoms/1177699147

[19] Rabiner, L., & Juang, B. (1986). An introduction to hidden Markov models. *IEEE ASSP Magazine*, *3*(1), 4–16. https://doi.org/10.1109/MASSP.1986.1165342

[20] Fine, S., Singer, Y., & Tishby, N. (1998). The hierarchical hidden Markov model: Analysis and applications. *Machine Learning*, *32*, 41–62. https://doi.org/10.1023/A:1007469218079

[21] Beal, M., Ghahramani, Z., & Rasmussen, C. (2001). The infinite hidden Markov model. *Advances in Neural Information Processing Systems*, *14*. https://papers.nips.cc/paper_files/paper/2001/file/e3408432c1a48a52fb6c74d926b38886-Paper.pdf

[22] Hermansky, H. (1990). Perceptual linear predictive (PLP) analysis of speech. *The Journal of the Acoustical Society of America*, *87*(4), 1738–1752. https://doi.org/10.1121/1.399423

[23] Ibrahim, A. B., Seddiq, Y. M., Meftah, A. H., Alghamdi, M., Selouani, S. A., Qamhan, M. A., ... & Alshebeili, S. A. (2020). Optimizing arabic speech distinctive phonetic features and phoneme recognition using genetic algorithm. *IEEE Access*, *8*, 200395–200411. https://doi.org/10.1109/ACCESS.2020.3034762

[24] Peng, P., Jiang, K., You, M., Xie, J., Zhou, H., Xu, W., & Xu, Y. (2023). Design of an efficient CNN-based cough detection system on lightweight FPGA. *IEEE Transactions on Biomedical Circuits and Systems, 17*(1), 116–128. https://doi.org/10.1109/TBCAS.2023.3236976

[25] Mariappan, R. (2023, January). Early detection of covid using spectral analysis of cough and deep convolutional neural network. In *Distributed Computing and Intelligent Technology: 19th International Conference, ICDCIT 2023, Bhubaneswar, India, January 18–22, 2023, Proceedings* (pp. 197–207). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-24848-1_14

[26] Bensid, K., Lati, A., Benlamoudi, A., Ghouar, B. E., & Senoussi, M. L. (2023). Efficient Covid-19 disease diagnosis based on cough signal processing and supervised machine learning. *Diagnostyka*, *24*(1), 2023103. https://doi.org/10.29354/diag/156751

[27] Jyothi, N. M., & Madhusudhanan, S. (2023). Cough audio signal-based clinical emergency classification of corona variant infected patients using multiclass SVM. In *Intelligent Systems and Applications: Select Proceedings of ICISA 2022* (pp. 333–350). Singapore: Springer Nature Singapore. https://doi.org/10.1007/978-981-19-6581-4_26

[28] Tawfik, M., Nimbhore, S., Al-Zidi, N. M., Ahmed, Z. A., & Almadani, A. M. (2022, January). Multi-features extraction for automating COVID-19 detection from cough sound using deep neural networks. In *2022 4th International Conference on Smart Systems and Inventive Technology (ICSSIT)* (pp. 944–950). IEEE. https://doi.org/10.1109/ICSSIT53264.2022.9716529

[29] Jayadi, A., Prasetio, B. H., Akbar, S. R., Widasari, E. R., & Syauqy, D. (2022, November). Embedded flu detection system based cough sound using MFCC and KNN algorithm. In *2022 International Conference of Science and Information Technology in Smart Administration (ICSINTESA)* (pp. 1–5). IEEE. https://doi.org/10.1109/ICSINTESA56431.2022.10041610

[30] Hamidi, M., Zealouk, O., Satori, H., Laaidi, N., & Salek, A. (2022). COVID-19 assessment using HMM cough recognition system. *International Journal of Information Technology*, 1–9. https://doi.org/10.1007/s41870-022-01120-7

[31] Drugman, T., Urbain, J., Bauwens, N., Chessini, R., Aubriot, A. S., Lebecque, P., & Dutoit, T. (2020). Audio and contact microphones for cough detection. *arXiv preprint arXiv:2005.05313*.

[32] Bansal, V., Pahwa, G., & Kannan, N. (2020, October). Cough classification for COVID-19 based on audio MFCC features using convolutional neural networks. In *2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON)* (pp. 604–608). IEEE. https://doi.org/10.1109/GUCON48875.2020.9231094

[33] Barkani, F., Satori, H., & Hamidi, M. (2020, October). Cough detection system based on ASR-HMM. In *2020 Fourth International Conference On Intelligent Computing in Data Sciences (ICDS)* (pp. 1–7). IEEE. https://doi.org/10.1109/ICDS50568.2020.9268765

11 AUTHORS

Fatima Barkani is a PHD student at the Faculty of Sciences, Dhar El Mahraz, Sidi Mohammed Ben Abdellah University, Fez – Morocco. She obtained her Bachelor's degree from the Multidisciplinary Faculty of Nador (FPN), Mohamed I University in 2015, and her Master's degree from the Faculty of Sciences, Dhar El Mahraz, Sidi Mohammed Ben Abdellah University, Fez in 2017 (E-mail: fatimabarkani4@gmail.com).

Mohamed Hamidi is a Professor at the Department of Computer Science, Multidisciplinary Faculty of Nador (FPN), Mohamed First University – Oujda, Morocco. He received his PhD in Computer Science in 2020 from the Faculty of Sciences, Dhar El Mahraz, Sidi Mohammed Ben Abdellah University, Fez. He obtained his Bachelor's degrees from the Multidisciplinary Faculty of Nador (FPN), Mohamed I University in 2011, and his Master's degrees from the Faculty of Sciences Tetouan, Abdelmalek Essaadi University – Morocco in 2013 (E-mail: mohamed.hamidi.5@gmail.com).

Ouissam Zealouk received his PhD degree in Computer Science in 2020 from the Faculty of Sciences, Dhar El Mahraz, Sidi Mohammed Ben Abdellah University, Fez. He obtained his Bachelor's degree from the Multidisciplinary Faculty of Nador (FPN), Mohamed I University in 2012, and his Master's degree from the Faculty of Sciences Tetouan, Abdelmalek Essaadi University – Morocco in 2014 (E-mail: ouissam zealouk@gmail.com).

Hassan Satori is a Professor at the Department of Computer Science, Faculty of Sciences, Dhar El Mahraz, Sidi Mohammed Ben Abdellah University – Fez – Morocco. He received his PhD in Speech Recognition and Signal Processing in 2009. He obtained his MSc in Physics from Mohamed Premier University, and his PhD in Nanotechnology and Computer Simulation from the same university in 2001 (E-mail: hssatori@gmail.com).