**DIAGNOSTYKA**, 2023, Vol. 24, No. 1

**e-ISSN** 2449-5220
**DOI:** 10.29354/diag/156751

# EFFICIENT COVID-19 DISEASE DIAGNOSIS BASED ON COUGH SIGNAL PROCESSING AND SUPERVISED MACHINE LEARNING

**Khaled BENSID *, Abdelhai LATI , Azeddine BENLAMOUDI, Brahim Elkhalil GHOUAR, Mohammed Larbi SENOUSSI**

Faculty of New information and Communication Technologies
University Kasdi Merbah Ouargla, BP 511, 30000, Ouargla, Algeria
* Corresponding author, e-mail: bensid.khaled@univ-ouargla.dz

Abstract

The spread of the coronavirus has claimed the lives of millions worldwide, which led to the emergence of an economic and health crisis at the global level, which prompted many researchers to submit proposals for early diagnosis of the coronavirus to limit its spread. In this work, we propose an automated system to detect COVID-19 based on the cough as one of the most important infection indicators. Several studies have shown that coughing accounts for 65% of the total symptoms of infection. The proposed system is mainly based on three main steps: first, cough signal detection and segmentation; second, cough signal extraction; and third, three techniques of supervised machine learning-based classification: Support Vector Machine (SVM), K-Nearest Neighbours (KNN), and Decision Tree (DT). Our proposed system showed high performance through good accuracy values, where the best accuracy for classifying female coughs was 99.6% using KNN and 88% for males using SVM.

Keywords: voice disease, COVID-19, cough sounds, features extraction, classification.

## 1. INTRODUCTION

COVID-19, also known as the Coronavirus, has been one of the most common diseases in the last three years, infecting over 534 million people and killing over 6.31 million. According to the World Health Organization, it has been labelled a pandemic in a survey of at least 200 nations. COVID-19 is considered one of the epidemics that cause serious respiratory infections, causing the loss of many lives recently all over the world, especially in countries with limited economic resources where they cannot provide the necessary capabilities to confront this pandemic [1]. Cough is a common symptom of COVID-19. The cough caused by COVID-19 infection is dry and in the case of superinfection, it becomes moist with expectation. COVID-19 infection can give all types of coughs (dry and oily) but in most cases it is dry, the intensity, frequency, and duration of which varies from person to person. Also, coughs associated with COVID-19 are dry coughs without mucous. This implies that most coughs will be dry and hacking. Once a COVID-19 patient starts coughing, it might be difficult to stop. These qualities are different from other diseases' cough, which is an attempt by the body to get rid of an irritant. These coughs typically subside on their own and don't last all day. Therefore, the cough can be considered a useful feature when making a diagnosis. Hence, studying the cough signal through its occurrence and episodes and its frequency helps to accurately diagnose and understand the disease and its development [2]. In addition to the medical difference between the cough of Covid-19 and the cough of other diseases, this difference also appears on the audio signal.

Considering this, Fig. 1 shows the spectrograms for cough sounds for a COVID-19-positive user, non-COVID-19 (or healthy) user, and a user with asthma.

Fig. 1 shows that the COVID-19 cough is more consistent in nature and has a longer duration in comparison to the non-COVID-19 cough.

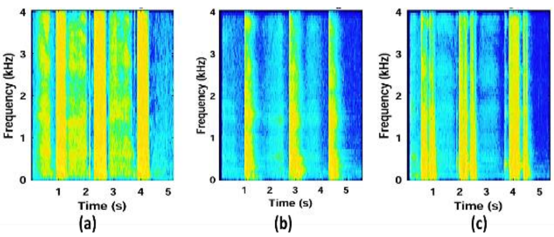


Fig. 1. Spectrograms for different diseases:
(a) COVID-19, (b) non-COVID-19, (c) asthma cough sound

However, manual processing of the cough signal in sounds is a very difficult task that leads to errors that make it difficult for experts to categorize their differing opinions. To avoid these problems, people are becoming more interested in studying and making algorithms that automatically detect coughs in audio signals [3]. Machine learning is one of the most developed scientific fields in recent years because of its great importance through its many uses in various fields. Moreover, the use of machine learning in diagnosing symptoms of the coronavirus is one of the most important research areas that has attracted the attention of many researchers with the aim of early detection of the disease to limit its spread.

In [4], the first system of COVID-19 detection based on sound signals was proposed by "N.Sharma" in 2020, where he conducted his study based on a Coswara database. The method presented by "N. Sharma" was a system based on a random forest classifier using a set of different short-time temporal and spectral acoustic features. This system has achieved an accuracy of 76.74%. In [5], "C. Brown" presented his proposed system for detecting asthma and COVID together, where his study was based on a large-scale crowdsourced dataset of respiratory sounds. Using spectrograms, spectral centroid, and MFCC, the proposed approach achieved an accuracy of 80% using the CNN model.

Despotovic, Vladimir, et al. [6] presented a new study based on the detection of COVID-19 based on a coughing and breathing database, namely, the CDCVA dataset. The presented method was based on an ensemble-boosted approach using spectrograms and wavelets to achieve an accuracy of 88.52%. In [7], researchers presented a new study about COVID-19 detection using cough audio signals collected in the VIRUFY database. Based on the CNN trained on features of zero-crossing rate, energy, energy entropy, spectral centroid, spectral entropy, spectral flux, spectral roll-offs, and MFCC, the proposed system achieved 88.52% accuracy.

Based on the Cambridge dataset [8], Rahman, Tawsifur et al. proposed an approach to detect COVID-19 using the sound of cough and breathing signals. The system achieved an accuracy of 96.5% for symptomatic and 98.85% for asymptomatic using spectrogram features and a CNN classifier.

In this paper, we contribute to detecting COVID-19 based on cough sounds by proposing a new algorithm to segment the cough signal from its original audio signal, as this algorithm isolates the cough signal from unwanted signals, mostly due to the background of the recording environment. To value this contribution, we classified the extracted cough segments based on three supervised machine learning classifiers, namely KNN, SVM, and DT, after extracting MFCC features from the cough sound.

The remainder of the paper is structured as follows:

In section 2, we talk about the COVID-19 cough sound. The proposed system is described in Section 3. In section 4, we produce the segmentation algorithm. Sections 5 and 6 describe voice feature extraction and classification, respectively. In section 7, we discuss the obtained results**.** Finally, section 8 concludes with the obtained results.

## 2. COVID-19 COUGH SOUND

As shown in Figure 2, the typical cough sound is generally divided into three phases: an explosive exhalation due to the glottis opening, an intermediate phase with attenuation of the cough, and a voiced phase due to the closure of the vocal cord. However, various cough patterns occur; for example, some cough sounds have only two, and the explosive phase is usually prolonged due to some illnesses [9]. Figure 2 also shows some differences between the cough sound signal of a healthy person and that of a person infected with COVID-19. For example, the explosive exhalation phase is longer in healthy people than in patients. In addition, the middle phase is longer in infected people than in healthy people due to inhalation difficulty caused by COVID-19 throat infections. Other differences can be detected by intelligent systems, such as those based on machine learning in classifying audio features extracted by signal processing methods.

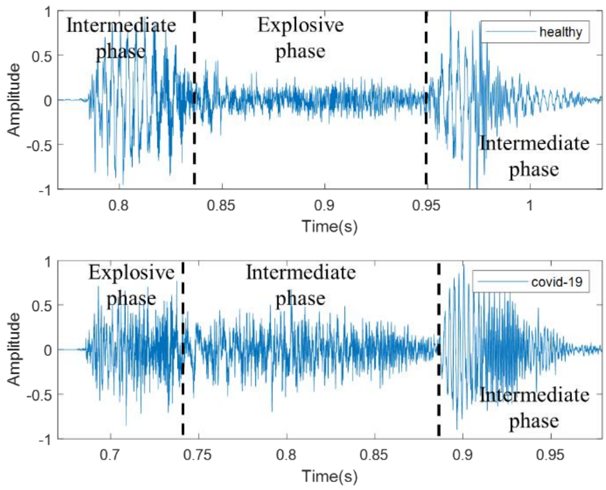


Fig. 2. Comparison of the cough sounds for a healthy subject and a COVID-19 [10]

## 3. SYSTEM DESIGN

The general procedure for the proposed approach to detect COVID-19 is described in the flowchart in Figure 3. In the first part, we take audio recordings containing cough signals using a database. Next, a time-domain algorithm extracts cough signals from audio recordings. The second part extracts features for each cough signal using MFCC and log energy. As a final step, subject machine learning classifiers (KNN, S VM, DT) are trained and evaluated for their performance in recognizing whether coughing indicates positive or passive COVID-19 infection.

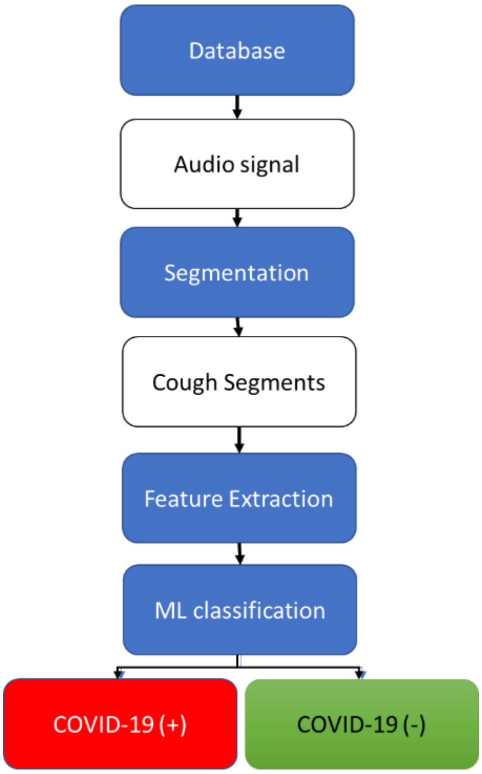


Fig. 3. System design

## 4. COUGH SEGMENTATION

In this work, we relied on a time-domain algorithm to segment cough sounds. This algorithm can extract at least one cough pulse and two at most. The main objective of using this cough segmentation algorithm is to use cough samples isolated from their original signal. The original cough signal contains unwanted parts from the ambient noise of the recording, which characterizes most of the signals in crowdsourcing databases. Figure 4 shows the stages of the proposed cough segmentation algorithm.

The proposed algorithm is based on the detection of the two main representative coughs.
To generalize the extraction algorithm, we apply a normalization phase on the original signal. The two significant coughs will be denoted respectively C1 and C2. They will be defined initially through their pic point by means of threshold comparison. Once done, each considered cough segment of fixed length L is extracted from the normalized signal. With length $L=m+M$, the segment is defined on the base of its corresponding threshold point. Here, $m$ points are taken before the pic point and the other $M$ points are taken just after this point.

**Cough Segmentation Algorithm**

Where:

$X = \frac{X}{\max(X)}$

$c = 1$

$k = 0$

$While\ k \leq length(x)$

$If\ |X(k)| \geq th\ \&\ C = 0$

$kth1 = k;$

$C1 = \sum_{i \in [kth1-m; kth1+M]} X(i) \quad ; (kth1 - m \geq 1)$

$C = C + 1$

$k = kth1 + M + 1$

$Else\ If\ |X(C)| \geq th\ \&\ C = 1$

$kth2 = k\ ;\ (kth2 + M < length(x))$

$C2 = \sum_{i \in [kth2-m; kth2+M]} X(i) \quad ; (kth2 - m > kth1 + M)$

$C = C + 1$

$k = kth2 + M$

$End\ if$

$If\ C \geq 2$

$Break$

$End\ if$

$End\ while$

$X$ : Audio signal
$th$ : Amplitude threshold detector
$m$ : Previous sample set
$M$ : Later sample set
C1 : The first cough signal segment
C2 : The second cough signal segment

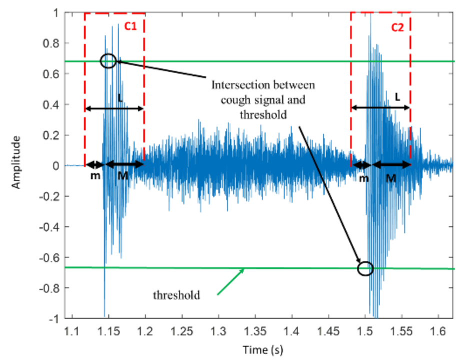


Fig. 4. Cough segmentation flowchart

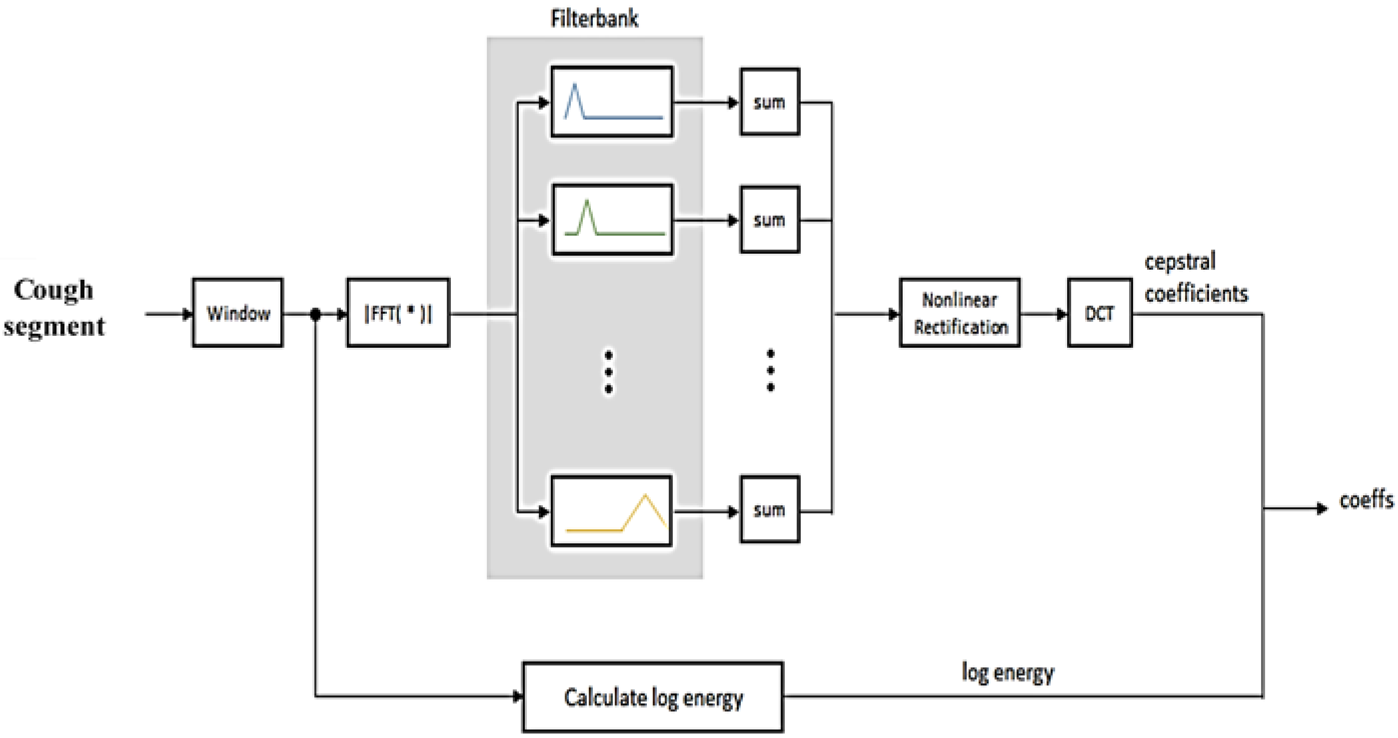


Fig. 5. MFCC Blocks

The proposed algorithm takes the cough signal from the total signal by detecting an amplitude peak above the threshold. In this work, we chose a threshold value equal to 0.85, considering that the value of the signal is confined between 1 and -1, that is, the normalization of the signal values, which is what is guaranteed by the first line of the algorithm. After detecting the first peak amplitude above the threshold, we take samples before and after it to capture a cough signal separate from the original signal. Then, the algorithm repeats the same steps to take another piece of the cough signal. Thus, we can extract two pieces of the cough signal from each original signal.

## 5. FEATURES EXTRACTION

In this work, we use the MEL-FREQUENCY CEPSTRAL COEFFICIENT (MFCC) algorithm with log energy to extract the discriminant characteristics of the cough signals. Where the MFCC algorithm is a simulation of the human ear, it is among the most prominent acoustic features that have given its efficacy in many studies related to voice recognition [10].

### 5.1. Mel-frequency cepstral coefficient (MFCC) with log energy

Mel frequency cepstrum coefficients are commonly used in voice recognition tasks. Cepstral coefficients are used to represent the filter in the source-filter
model of speech (vocal tract). The frequency response of the vocal tract is rather smooth, although the source of voiced speech can be described as an impulse train. As a result, the spectral envelope of a speech segment can be used to estimate the vocal tract. The goal behind Mel frequency cepstral coefficients is to reduce information about the vocal tract (smoothed spectrum) into a minimal number of coefficients based on understanding the cochlea. Although there is no hard standard for calculating the coefficients, the basic steps are outlined in Figure 5. The default Mel filter bank spaces out the first 10 triangular filters linearly and the rest of the filters logarithmically. The log energy is frequently used to supplement or completely replace the information found in the zeroth Mel frequency cepstral coefficient. The calculation of log energy is dependent upon the input domain. The following equation is used to get the logarithmic energy if the input (cough segment) is a time-domain signal:

$$\log E = \log(sum(x^2)) \qquad (1)$$

The following equation is used to determine the logarithmic energy if the input is a frequency domain signal:

$$\log E = \log\left(\frac{\mathrm{sum}(|x|^2)}{\mathrm{FFTLength}}\right) \qquad (2)$$

## 6. CLASSIFICATION

### 6.1. Support Vector Machine

Support Vector Machines (SVM) is a supervised machine learning method considered a discriminative technique that can find solutions to classification, regression, and detection issues.

SVM involves creating a hyperplane that divides two or more groups of points according to the circumstances and configuration points. The initial concept behind SVM was to best separate scheduled points into distinct classes by using a kernel function [11].

In a high-dimensional or infinite-dimensional space, SVMs create a hyperplane or group of hyperplanes. Using a training dataset, the hyperplane identifies which points are the best. The hyperplane with the greatest distance from any class's closest training data point achieves a decent separation (the functional margin) where the generalization error of the classifier decreases as the margin increases [12].

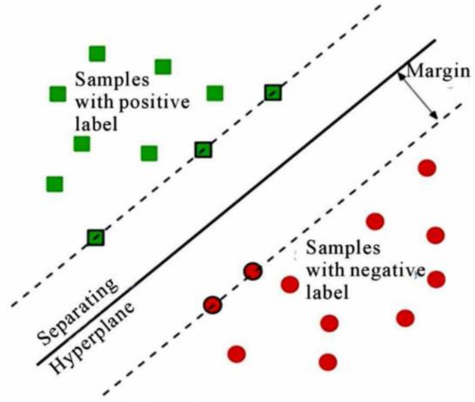


Fig. 6. Concept of SVM [13]

### 6.2. K-Nearest Neighbours (KNN)

K-Nearest Neighbours is a typical supervised statistical computer vision method that classifies images by comparing the "K" value of the training data with the test data. The "K" value is estimated using training feature extraction.

The Euclidean equation principle is used in a KNN classifier to find similar things. Then, it is used to classify things and figure out what happened. In both cases, the input is the k training samples from the dataset closest to each other. KNN is mainly used for classification or regression [13].

When using the KNN classifier, the result is a class member. The objects are grouped based on their neighbours' votes, and each object is put into the group with the most votes from its k nearest neighbours (K is a positive integer, usually small). If k = 1, the object is just put in the class of its single closest neighbour.

For classification, it can be helpful to give neighbours different weights, so neighbours closer to the mean contribute more to the mean than neighbours farther away.

One common way to weigh neighbours is to give each one a value of 1/d, where d is the distance to the neighbour. Neighbours are chosen from a set of objects whose classes (for KNN classification) are known.

This can be thought of as a training set for the algorithm, but no explicit steps are needed to train it. One thing about the KNN algorithm is that it considers how the data is organized locally. The KNN function uses the Euclidean distance, which can be found with the following equation: [14, 15].

$$D_m(x_i, x_j) = \left( \sum_{i,j=1}^{n} \left| x_i - x_j \right|^m \right)^{1/m} \qquad (3)$$

Where p and q are the things to be compared, and n is the number of things to be compared. There are other ways to figure out numbers. For example, the distance is the distance from Manhattan distance [16, 17].

Another idea is how to find the parameter K. The KNN algorithm picks how many neighbors to look at. The selection of the K value significantly influences the diagnostic performance of the KNN algorithm. While capital k lessens the impact of variance due to random errors, it also runs the danger of obscuring minute but significant patterns. That finds a balance between overfitting and underfitting while picking K writers [18][19]. The K parameter is set to the square root of the training dataset observations.

### 6.3. Decision tree

One of the potent techniques frequently employed in several fields, including machine learning, image processing, and pattern recognition, is decision trees [20].

A numerical feature is compared to a threshold value in each test, and the DT is a sequential model that effectively and cogently connects a set of fundamental tests. The conceptual principles in the neural network of connections between nodes are significantly simpler to create than the numerical weights. Therefore, DT is used mostly for grouping purposes. Additionally, DT is a classification model frequently used in data mining. Each tree is made up of nodes and branches. Each node represents a feature in a classification category, and each subset specifies a value the node may accept. Decision trees have several implementation domains due to their easy analysis and accuracy in various data formats [21-23].

## 7. EXPERIMENTAL RESULT

### 7.1. Exprimental dataset

The dataset used in this strain was collected by the government of Buenos Aires, Argentina, and has been clinically validated. This database is called IATos v1 [24] and comprises 5884 cough cases from 2821 individuals via WhatsApp used in a clinical setting (Pizzo et al., 2021). This data set was one of the few-balanced data sets, with 1409 individuals testing positive for COVID-19 and 1412 testing negatives. This database was recorded with a sampling frequency of 48 kHz.

### 7.2. Exprimental metrics

In this study, KNN, SVM, and DT classifiers trained by the MFCC with log energy features for both genders assessed the performance of the voice pathology detection system.

The classification results are provided in terms of accuracy, precision, specificity, and F1 score. These performance indicators are widely used to assess the effectiveness of various medical decision systems.

**Accuracy:** we can define accuracy as a measure of the capability of classifying the samples correctly, which is expressed as follows:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

**Sensitivity:** the probability that abnormal samples will be diagnosed as positive.

$$\text{Sensitivity} = \frac{TP}{TP + FN} \tag{5}$$

**Precision:** the probability of normal samples being correctly identified.

$$Precision = \frac{TP}{TP + FP} \tag{6}$$

**Specificity**: the probability of normal samples being incorrectly identified.

$$Specificity = \frac{TN}{FP + TN} \tag{7}$$

**F1 Score:** The F1-score measures a model's accuracy on a dataset. It evaluates binary classification systems, classifying examples as "positive" or "negative."

$$F_1 score = 2 \times \left(\frac{Precision \times Recall}{Precision + Recall}\right) \tag{8}$$

**True Positive (TP):** the pathogenic nature of the speech sample is recognized by the marker.

**True Negative (TN):** the voice sample is healthy, and the marker detects it.

**False Positive (FP):** the voice sample is normal, but the marker detects it as pathological.

**False Negative (FN):** the diseased voice sample is recognized by the marker as healthy.

### 7.3. Result and discussion

The early medical detection system is one of the best solutions to improve the performance of medical devices and reduce the spread of epidemics. In our study, we trained supervised machine learning models on MFCC features extracted from healthy cough segments and covid-19 cough segments. The performance of the proposed system in classifying cough segments into healthy or covid-19 for females and males was evaluated separately.

In the first step, we change the value of K in KNN classifier and evaluate the performance of our proposed system.

Table 1. The classification results with K value

| K | Accuracy | Sensitivity | Specificity | F1-score | Precision |
|---|---|---|---|---|---|
| 1 | 82 | 73.52 | 100 | 97 | 1 |
| 2 | 74 | 65.78 | 100 | 97 | 1 |
| 3 | 64 | 62.96 | 65.21 | 0.34 | 0.68 |
| 4 | 66 | 61.11 | 78.57 | 0.88 | 0.73 |

Table 1 presents the results of the male cough segments classification into healthy or COVID-19 using KNN classifier. The result shows that k value play an important role in improving the efficiency of our proposed system. And the best value of k is k=1 when give accuracy = 82% and f1-score = 97%.

Table 2 presents the results of the female cough segments classification into healthy or COVID-19.

**For Female:**

Table 2. The classification results of the cough segments into healthy or COVID-19 for females

| classifier | accuracy | Sensitivity | specificity | F1 score | Precision |
|---|---|---|---|---|---|
| SVM | 89.9 | 92.08 | 87.7 | 88 | 0.84 |
| KNN | 99.6 | 99.3 | 100 | 99 | 1 |
| DT | 99.3 | 98.6 | 100 | 99 | 1 |

Table 2 shows the efficiency of the suggested system for COVID-19 detection in females. The KNN algorithm provided the best accuracy for detection, with a value of 99.6%.

Table 3 shows the results of the Male cough segments classification into healthy or COVID-19.

**For Male:**

Table 3. The classification results of the cough segments into healthy or COVID-19 for males

| classifier | accuracy | Sensitivity | specificity | F1 score | Precision |
|---|---|---|---|---|---|
| SVM | 88 | 82.75 | 95.23 | 88 | 96 |
| KNN | 82 | 73.38 | 100 | 84 | 1 |
| DT | 84 | 75.75 | 100 | 86 | 1 |

As shown in table 4, We achieved the best classification accuracy with the SVM algorithm with a value of 99.6%.

**Table 4: The best results of our test**

| Gender | classifier | Accuracy | sensitivity | specificity | Precision |
|---|---|---|---|---|---|
| Male | SVM | 88 | 82.75 | 95.23 | 96 |
| Female | KNN | 99.6 | 99.3 | 100 | 1 |

The best performances of the proposed system are presented in Table 4, which shows the results obtained when training this system with different classifier feature sets. As can be seen, while KNN outperforms SVM in the female gender, it achieves an accuracy of 99%. However, on the other hand, in the male gender, the SVM outperforms the KNN classifier because its accuracy reaches 88%.

## 8. CONCLUSION

Technology is frequently employed in various sectors, including medicine, to identify diseases. The fundamental goal of medical detection techniques is to detect COVID-19 in the person being tested. This COVID screening method is a substitute for traditional methods of restorative assays. In this way, it aims to help curb this global spread. With that in mind, these are forward-looking approaches and have proved to be effective in terms of accuracy. However, our work in this area has limitations in the number of tests used to build prior knowledge proof. This can, of course, be used as confirmation, provides the concept of machine operability, and has well-founded learning methods to prepare for the diagnosis of COVID. The main goal of our study is to propose a caught signal for the COVID-19 diagnosis system using MFCC features and KNN, SVM, and DT classifiers. The obtained results show excellent performance for pathology and health and are very encouraging and show the effectiveness of both genders in developing an efficient detection system.

***Author contributions***: *research concept and design, K.B., A.B.; Data Collection and assembly, K.B., A.B., B.E.G.; Data analysis and interpretation, K.B., A.L.; Writing the article, B.K.G., M.L.S., K.B., A.B.; Critical revision of the article, A.L.; Final approval of the article, K.B., A.B.*

**Declaration of competing interest:** *The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.*

## REFERENCES

1. Pramono RX, Imtiaz SA, Rodriguez-Villegas E. Automatic cough detection in acoustic signal using spectral features. In2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) 2019:7153-7156. https://doi.org/10.1109/EMBC.2019.8857792
2. Alqudaihi KS, Aslam N, Khan IU, Almuhaideb AM, Alsunaidi SJ, Ibrahim NM, Alhaidari FA, Shaikh FS, Alsenbel YM, Alalharith DM, Alharthi HM. Cough sound detection and diagnosis using artificial intelligence techniques: challenges and opportunities. Ieee Access. 2021 15; 9:102327-44. https://doi.org/10.1109/ACCESS.2021.3097559
3. World Health Organization. Coronavirus disease 2019 (covid-19): situation report. 2020;51.
4. Sharma N, Krishnan P, Kumar R, Ramoji S, Chetupalli SR, Ghosh PK, Ganapathy S. Coswara–a database of breathing, cough, and voice sounds for COVID-19 Diagnosis. arXiv 2020; arXiv:2005.10548. https://doi.org/10.48550/arXiv.2005.10548
5. Brown C, Chauhan J, Grammenos A, Han J, Hasthanasombat A, Spathis D, Xia T, Cicuta P, Mascolo, C. Exploring automatic diagnosis of COVID-19 from crowdsourced respiratory sound data. In Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, Virtual. 2021:3474–3484. https://doi.org/10.48550/arXiv.2006. 05919
6. Despotovic V, Ismael M, Cornil M, Mc Call R, Fagherazzi G. Detection of COVID-19 from voice, cough and breathing patterns: Dataset and preliminary results. Comput. Biol. Med. 2021;138: 104944. https://doi.org/10.1016/j.compbiomed.2021.104944
7. Islam R, Abdel-Raheem E, Tarique M. A study of using cough sounds and deep neural networks for the early detection of COVID-19. Biomed. Eng. Adv. 2022;3:100025. https://doi.org/10.1016/j.bea.2022.100025
8. Rahman T, Ibtehaz N, Khandakar A, Hossain MSA, Mekki YMS, Ezeddin M, Bhuiyan EH, Ayari MA, Tahir A, Qiblawey Y. An intelligent application to detect COVID-19 patients using cough and breath sounds. Diagnostics. 2022; 12:920. https://doi.org/10.3390/diagnostics12040920
9. Morice AH, Fontana GA, Belvisi MG, et al. ERS guidelines on the assessment of cough, European Respiratory Journal. 2007;29(6):1256–1276. https://doi.org/10.1183/09031936.00101006
10. Islam Rumana, Abdel-Raheem Esam, Tarique Mohammed. Early detection of COVID-19 patients using chromagram features of cough sound recordings with machine learning algorithms. In: 2021 International Conference on Microelectronics (ICM). IEEE. 2021:82-85. https://doi.org/10.1109/ICM52667.2021.9664931
11. Hearst MA, Dumais ST, Osuna E, Platt J, Scholkopf B. Support vector machines. IEEE Intelligent Systems and their applications. 1998;13(4):18-28. https://doi.org/10.1109/5254.708428
12. Huynh TH, Tran VA, Tran HD. Semi-supervised tree support vector machine for online cough recognition. In Twelfth Annual Conference of the International Speech Communication Association 2011.
13. Hadjaidji E, Korba M, Khelil K. Spasmodic dysphonia detection using machine learning classifiers. In Inter Conference on Recent Advances in Mathematics and Infor (ICRAMI). 2021;1–5. https://doi.org/10.1109/ICRAMI52622.2021.9585920
14. Rajani Kumari LV, Padma Sai Y. Classification of arrhythmia beats using optimized K-nearest neighbor classifier. InIntelligent Systems. 2021:349-359. https://doi.org/10.1007/978-981-33-6081-5_31.
15. Brian Kulis. Metric Learning: A Survey. Foundations and Trends® in Machine Learning: 2013;5(4):287-364. https://doi.org/10.1561/2200000019
16. Khorshid SF, Abdulazeez AM. Breast cancer diagnosis based on k-nearest neighbors: A review. Journal of Archaeology of Egypt/Egyptology. 2021;15;18(4):1927-51. https://archives.palarch.nl/index.php/jae/article/view/6601
17. Sipper M, Moore JH. Conservation machine learning: a case study of random forests. Scientific Reports. 2021;11;11(1):1-6. https://doi.org/10.1038/s41598-021-83247-4
18. Zhang, Z. Introduction to machine learning: k-nearest neighbors. Annals of translational medicine. 2016;4(11). https://doi.org/10.21037/atm.2016.03.37
19. Usman M, Zubair M, Ahmad Z, Zaidi M, Ijyas T, Parayangat M, Wajid M, Shiblee M, Ali SJ. Heart rate detection and classification from speech spectral features using machine learning. Archives of Acoustics, 2021;46. https://doi.org/10.24425/aoa.2021.136559

20. Goswami S, Pramanick R, Patra A, Rath SP, Foltin M, Ariando A, Thompson D, Venkatesan T, Goswami S, Williams RS. Decision trees within a molecular memristor. Nature. 2021;597(7874):51-6. https://doi.org/10.1038/s41586-021-03748-0
21. Charbuty B, Abdulazeez A. Classification based on decision tree algorithm for machine learning. Journal of Applied Science and Technology Trends. 2021;24;2(01):20-28. https://doi.org/10.38094/jastt20165
22. Bilski P, Bobiński P, Krajewski A, Witomski P. Detection of wood boring insects' larvae based on the acoustic signal analysis and the artificial intelligence algorithm. Archives of Acoustics. 2016;42(1):61-70. https://doi.org/10.1515/aoa-2017-0007
23. Zhiyong G, Jiwu L, Rongxi W. Prognostics uncertainty reduction by right-time prediction of remaining useful life based on hidden Markov model and proportional hazard model. Eksploatacja i Niezawodność. 2021;23(1). https://doi.org/10.17531/ein.2021.1.16.
24. Pizzo DT, Esteban S, Scetta M. IATos: AI-powered pre-screening tool for COVID-19 from cough audio samples. arXiv preprint arXiv: 2021;2104:13247. https://doi.org/10.48550/arXiv.2104.13247

Received 2022-07-08
Accepted 2022-11-21
Available online 2022-12-01

**Khaled BENSID** received his BS and MSc degrees in the University kasdi merbah of ouargla, Algeria, in 2010 and 2012, respectively. The PhD was received in Electronic – Communication and signal processing in 2018, from Ouargla University, where he is actually an associate professor since 2019 in the Department of electronics and telecommunication. His research interests include computer vision, pattern recognition, signal and image processing, biometrics.

**Abdelhai LATI** received his engineering degree in electric and electronic engineering from IGEE (Institué de Génie Electrique et Electronique ex. INELEC) at Boumerdes University in Algeria, followed by his magister degree in automatics from EMP (Ecole Militaire Polytechniques) at Bordj Elbahri, Algeria, then a doctorate degree in electronics from University of Sciences and Technologies–Houari Boumedian (USTHB), Algiers, Algeria. Since 2019, he has been a professor researcher at the University Kasdi Merbah Ouargla, Algeria. Currently, his research interests include digital image processing and embedded systems.

**Azeddine BENLAMOUDI** received his BS and MSc degrees in the University Mohamed Khider of Biskra, Algeria, in 2010 and 2012, respectively. The PhD was received in Electronic – Communication and signal processing in 2018, from Ouargla University, where he is actually an associate professor since 2018 in the Department of electronics and telecommunication. His research interests include computer vision, pattern recognition, signal and image processing, biometrics and spoofing detection.

**Brahim Elkhalil GHOUAR** received his BS and MSc degrees in the University Kasdi Merbah of Ouargla, Algeria, in 2019 and 2022, respectively.

**Mohammed Larbi SENOUSSI** received his BS and MSc degrees in the University Kasdi Merbah of Ouargla, Algeria, in 2019 and 2022, respectively.