Early COVID-19 Diagnosis from Cough Sound Using Random Forest and Low-Level Descriptors

Skander Hamdi
Department of computer science
University of Ferhat Abbes
Setif, Algeria
skander.hamdi@univ-setif.dz

Abdelouahab Moussaoui
Department of computer science
University of Ferhat Abbes
Setif, Algeria
abdelouahab.moussaoui@univ-setif.dz

Mourad Oussalah
Department of computer science and engineering
University of oulu
Oulu, Finland
mourad.ousalah@oulu.fi

Mohamed Saidi
Department of computer science
University of Ferhat Abbes
Setif, Algeria
mohamed.saidi@univ-setif.dz

***Abstract*—COVID-19 pandemic has raised the need to develop strategies for conducting large scale testing and isolate the infected and suspicious cases. The interest for building AI-based solutions to early diagnose COVID-19 cases from images, laboratory or any user's physiological input has been voiced by the research community. In this paper, we proposed an ensemble learning model that uses patient sound data when coughing to distinguish COVID-19 infected patients. The approach has been tested on a publicly available large-scale dataset called COUGHVID. We obtained 90.42% as classification rate, 99.65% as precision, 81.21% as sensitivity and 0.9046 for Area Under Curve (AUC).**

***Index Terms*—COVID-19, Machine learning, Cough sound, Audio features, Low-Level Descriptors, Random Forest**

I. Introduction

World Health Organization (WHO), on March, 2020, declared a global pandemic called COVID-19 (COrona VIrus Disease 2019) which is caused by Severe Acute Respiratory Syndrome (SARS-CoV2). Until June 4, 2021, statistics provided by the WHO [1] show that the pandemic has caused the death of 3,698,621 and the infection of 171,782,908 people around the world and it still rapidly spread due to the absence of any symptoms of infection until the virus long incubation period is reached. Fig. 1 shows the distribution of confirmed infection cases across different regions around the world where Americas present the highest number of cases (more than 68 millions) followed by Europe and South-East Asia with more than 54 millions and 32 millions, respectively. Early diagnosis is an essential step to reduce the infection spread, as infected people can be quickly identified and isolated. However, to reach this, a large number of test kits, Reverse Transcription Polymerase Chain Reaction (RT-PCR), which is the only golden option [1], must be performed. However, this depends mainly on the country's health system capacity to conduct large scale RT-PCR testing. Is it possible to provide a sufficient number of test kits in all hospitals and health centers to cover all citizens?

Identifying reliable and cost effective early diagnosing toolkit becomes a challenge raised by the research community and health authorities to quickly find effective ways to detect COVID-19 infections instead of using RT-PCR whose results can only be available after several hours with a potential lack of testing kits as well. Several AI-based prototypes were proposed by researchers to speed up the diagnosing process using computerized tomography (CT scans) [2], [3], [4], [5] and X-Ray scans [4], [5], [7], [8], [9]. CT scan analysis proved its effectiveness against RT-PCR in a comparison conducted by Yang et al. in [10]. What allows the exploitation of CT and X-Ray scans for rapid diagnosis are the hidden patterns that distinguish an infected person from a non-infected one. However, this requires the patient to personally attend the hospital or medical center, which increases the possibility of infection as well, especially among members of the medical staff.

According to [11], there are unique early pulmonary pathological indicators in COVID-19 infected persons even before the start of COVID-19 symptoms such as dry cough, fever, and some trouble breathing. In fact, cough is a symptom of a variety of medical problems caused by bacterial or viral respiratory infections that are unrelated to COVID-19 [12]–[15] which could lead us to the conclusion that it is not possible to use audio data for diagnosing the disease. However, this contrasts with another research reported by Chatrzarrin et al. [16], which confirmed the clear distinction between dry and wet cough. This motivates us to waive at the prospect of relying on the cough sound in diagnosing the disease through training an AI-model that differentiates between the cough of a sick person and that of healthy person.

In this research, we use the publicly available cough sound dataset, called COUGHVID to build a machine learning model which will able to classify a cough sound as COVID-19 or healthy. Especially, we aim to propose a statrgy that ensures

i) data cleaning; ii) silence removal; iii) extracting a large set of relevant features; iv) data augmentation and handling class imbalance problem as a single step and then; v) a Random Forest method for classifying a cough sound to either a Likely-COVID-19 or Not-Likely-COVID-19. The rest of the paper is organized as follows. Section II presents some background of our research, where we discuss previous works. Section III outlines our methodology, where we present the used data and its meta-data, as well as the details of the proposed framework. Section IV discusses the obtained results, Section V reports our concluding statement and potential perspective work.

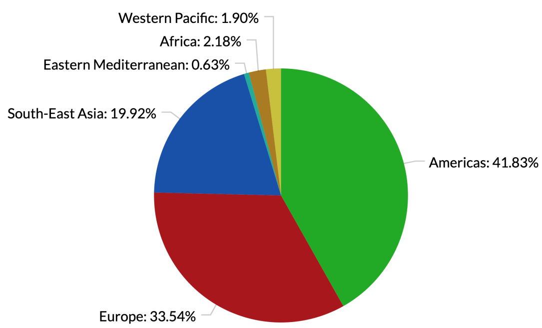

Fig. 1: COVID-19 confirmed cases across the world on June 04, 2021

II. Background

Many researches have been performed on diagnosing COVID-19 using various deep learning methods on several datasets. ResNet50 architecture, Transfer Learning and Convolutional Neural Networks (CNN) have been used in [17]. Long Short Term Memory (LSTM) using Recurrent Neural Networks (RRN) have been investigated in [18]. Deep Transfer Learning-based Multi-Class classifier, Deep Transfer Learning-based Binary-Class classifier and CNN model are studied in [19]. Agbley et al. [24] proposed a multi-modal classification based on audio signals and clinical data of COUGHVID dataset [25] where the original audio signals have been decomposed using Continous Wavelet Transform. A CNN architecture was used to train a transfer learning on the scalograms in order to initialize weights and an autoencoder for the clinical data. The two modalities have been merged and fed into a feed-forward neural network whose last layer uses a softmax activation function.

In machine learning context, the publicly Coswara and Sarcos datasets have been employed by Pahar in [20] with several machine learning and deep learning algorithms. Synthetic Minority Oversampling TEchnique (SMOTE) [21] was used to deal with dataset skew. Mel Frequency Cepstral Coefficients MFCCs, Log Energies, Zero-Crossing Rate (ZCR) and Kurtosis have been extracted as features to train machine learning models. Logistic Regression with 13 MFCCs, achieved an accuracy of 75.7%, 57% of specificity and 94% sensitivity. Support Vector Machine with 26 MFCCs gave an accuracy of 73.91%, 74% for both specificity and sensitivity. After applying Sequential Forward Search (SFS) for feature extraction hyperparameters, the proposed LSTM architecture outperforms other state of art algorithms by achieving 92.91% accuracy, 96% specificity, 91% sensitivity and 0.9375 for Area Under Curve (AUC).

Brown et al. [23] collected a database of coughing and breathing dataset via a website and a mobile application, then a feature extraction step was performed in two stages. First, a set features were manually extracted, which include 13 components of MFCCs, temporal differential of MFCC, the differential of the delta of MFCC, duration, onset, tempo, period, RMS Energy, Spectral Centroid, Roll-off frequency and ZCR. Next, a set of 11 statistical functions have been applied to the handcrafted feature to better capture the distribution beyond the mean value. The total number of handcrafted features is 477. Second, using transfer learning, by employing VGGish model to extract a 256-dimensional feature vector which result in 733 features. For classification purpose, Logistic Regression, Gradient Boosting Trees, Support Vector Machine with Radial Basis Function kernel were contrasted, where the support vector machine. The latter gave after data augmentation, 61% as precision, 81% as recall and 0.88 as AUC.

Another study from Melek [26] used MFCCs of Virufy and NoCoCoDa databases to build a COVID-19 detection system by experimenting various machine learning methods including SVM, Partial Least Squares Regression, Linear Discriminant Analysis and K-Nearest Neighbors with Euclidean and Chebyshev distances. With 13 values of MFCCs and 2048 as frame length, kNN with Chebyshev distance outperforms the rest of methods by achieving 93.89% as accuracy. After using the optimum frame length with different number of MFCCs, the results show that the optimum number of features (MFCCs) is 19. Another hyper-parameter: the number of segments (the number of frames) was made varying from 1 to 50. Experiments showed that a choice of 17 segments is optimal and yields for a k-NN classifier, an accuracy rate of 98.33%, sensitivity of Non-COVID-19 class of 100%, 97.20% of COVID-19 class sensitivity and 0.9860 as AUC.

III. Methodology

In this research, the large-scale cough sound dataset was used after processing steps to build a machine learning model that makes rapid diagnosis of COVID-19. In this section, we will describe the used dataset as well as our methodology.

A. COUGHVID Dataset

We used the COUGHVID dataset [25], which is publicly available online via Zenodo. It provides 27,550 recordings which made it one of the largest datasets in the context of cough analysis under COVID-19 disease. It also provides a wide range of clinical and meta-data provided by users and annotators experts such as: respiratory condition, fever, muscle pain, COVID-19 symptoms, age, gender and other subjects related to PCR test, which can be used as data modality besides audio data as reported in [27]. The data we worked on is the audio recordings. For this purpose, we focused on covid_status label which has three possible values: *COVID-19, Healthy* and *Symptomatic*. The table I presents the number of recordings in each class and some statistics about each one after eliminating the unlabeled samples (no status declared).

	COVID-19	Healthy	Symptomatic	Total
Samples count	1155	12479	2590	16224
Percentage	**7.12%**	**76.92%**	**15.96%**	100%
Total length (hours)	2.67	28.62	5.92	37.21
Minimum (seconds)	0.48	0.48	0.6	-
Maximum (seconds)	13.08	19.74	19.14	-
Mean (seconds)	8.32	8.25	8.23	-
Standard deviation	2.48	2.37	2.38	-

TABLE I: COUGHVID Dataset class labels distribution and audio files duration after eliminating the unlabeled samples.

B. Data cleaning and pre-processing

1) Non-labeled samples elimination: As mentioned before, COUGHVID dataset contains 27550 recordings including 11326 without status label. For that reason, we eliminate all non-labeled samples which results 16224 labeled samples.

2) Silence removal: An important preliminary step is to remove the silence from each recording (from the beginning, the end and between every informative parts of the sound). To perform this step, we used a Python library called Unsilence [1]. As there is some empty recordings, silence removal step reduced the total number of recordings to 16082.

3) Classes merging: To turn the problem into a binary classification, we chose to combine *COVID-19* and *Symptomatic* into one single class by considering the symptomatic cases as positive. The new class names and their statistics are described in Table II.

	Likely-COVID-19	Non-Likely-COVID-19
Samples count	3705	12377
Percentage	**23.03%**	**76.97%**
Total length (hours)	4.45	14.26
Minimum (seconds)	0.51	0.49
Maximum (seconds)	10.05	14.15
Mean (seconds)	4.33	4.14
Standard deviation	2.19	2.12

TABLE II: New COUGHVID Dataset class labels distribution and audio files duration after eliminating the unlabeled samples,removing silence, classes merging

C. Low-Level Descriptors and Feature extraction

As our methodology is a machine learning based, we need to extract a set of features for learning phase. Since the use of prosodic and spectral features has achieved great success in classifying audio data modality, as mentioned in related works, a majority of researchers utilized different numbers of MFCCs and ZCR as input vector for their learners. The mentioned features will be used with a large set of features called Low-Level Descriptors (LDD). Furthermore, the use of LDD has given good results in several works for sound classification tasks [22], [28]–[31] and [32], [33] for health-care and medical subjects. We used the largest set of features called *ComParE* which provides the third level of descriptors (LLD deltas) using OpenSMILE toolkit [2] [34]. ComParE is the result of calculation of different functionals over LLD contours yielded exactly **6373** features including *prosodic*, *spectral* and *cepstral* features shown in Table III and fully described in [35]. The whole set of 6373 features is used during the training phase.

Feature type	Feature list
Prosodic	- Sum of auditory spectrum (loudness) - Sum of RASTA-style filtered auditory spectrum - Zero-Crossing Rate - RMS Energy - F0 (SHS and viterbi smoothing)
Spectral	- RASTA-style auditory spectrum, bands 1–26 (0–8 kHz) - Spectral energy 250–650 Hz, 1 k–4 kHz - Spectral roll off point 0.25, 0.50, 0.75, 0.90 - Spectral flux, centroid, entropy, slope - Psychoacoustic sharpness, harmonicity - Spectral variance, skewness, kurtosis
Cepstral	1 to 14 MFCCs
Quality of sound	- Probability of voicing - Log. HNR, Jitter (local, delta), Shimmer (local)

TABLE III: ComParE feature set [34], [35]

D. Data augmentation and class imbalance

Table II shows that *Likely-COVID-19* class has only **23.03%** from the whole dataset which is considered as skew and can affect the classification model. We noticed that because of the large difference between the two classes, the model lacks relevant patterns to correctly identify *Likely-COVID-19* class. To solve this problem, we propose to use Synthetic Minority Over-sampling Technique (SMOTE) [21] which is considered as both class imbalance handling and data augmentation technique where it synthesizes new examples from the minority class, i.e, the class which contains less samples.
SMOTE starts by picking a minority class instance at random and looking for its k-closest minority class neighbors. The synthetic instance is then produced by randomly selecting one of the k-closest neighbors ”b” and connecting ”a” and ”b” in the feature space to form a line segment. The synthetic instances are created by combining the two chosen examples, ”a” and ”b”, into a convex combination [21], [36].

E. Random Forest classifier

We used Random Forest as a classifier [37], which is a powerful machine learning algorithm due to its inherent property of ensuring diversity by searching the most discriminant features. Random forest is created by an ensemble of decision trees, trained with **bagging** method that uses a combination of multiple learning models to perform a final powerful model. Table IV presents the tuned hyper-parameters. We used the Scikit-learn's [3] default values for the rest of parameters.

[1] https://github.com/lagmoellertim/unsilence

[2] https://github.com/audeering/opensmile-python

[3] https://scikit-learn.org/

Parameter	Value
Bootstrap	**False** The whole training set is used to build each tree
Number of trees	**200**
Maximum features	**Square root** Square root of the total number of features The number of features to consider when looking for the best split

TABLE IV: Tuned hyper-parameters of the classifier.

F. Data split and model validation strategy

The dataset was split into two parts: 80% for training and 20% for testing. We used k-fold cross-validation method for validating our model with ***k=5***. We shall mention that the cross-validation and the training phases are performed using the training set (considered as training and validation set) which means that the test set is considered as unseen data which will provide more validation to our methodology.

IV. Results and discussion

The following evaluation metrics are calculated for each fold as well as test phase: *accuracy*, *recall*, *precision*, *F1-Score*, *AUC* and *specificity*. Table V presents the obtained results for the k-fold cross-validation in terms of training time and shows the test results of our classifier on the 20% of data with respect to each of the aforementioned performance metrics.

We also include the Receiver Operating Characteristic (ROC) in Figure 2 to summarize the test's overall diagnostic accuracy. The value of test AUC **(0.9046)** is considered as good, which shows that our model has more than 90% chance to find a discrimination between *Likely-COVID-19* and *Non-Likely-COVID-19*. Figure 3 illustrates the testing performance according to the number of classified samples in each class, where only 7 samples belonging to *Non-Likely-COVID-19* are misclassified. For the precision, our model is able to recognize more than **99%** of *Non-Likely-COVID-19* cases. In terms of sensitivity metric which is very relevant for medical diagnosis, **81.21%** from *Likely-COVID-19* cases are found to be correctly classified.

V. Conclusion

With the emergence of deep learning technology, new prospects emerged in tackling COVID-19 diagnosis. However, machine learning methods are still able to give promoising results. In this respect, this paper advocates a machine learning based framework for identifying COVID-19 cases from cough sound data. It makes use of COUGHVID dataset, which is one of the largest dataset made available for COVID-19 detection purpose. Due to dataset class skew, we merged two classes into one single class and also used SMOTE for data augmentation and tackling class unbalance. The feature extraction phase was carried out using OpenSMILE to extract a large set of features called LLD (6373 features) titled ComParE. By utilizing random forest algorithm for binary classification, we obtained an overall test result of 90.42%, 99.65% and 81.21% for accuracy, precision and sensitivity, respectively. Furthermore,

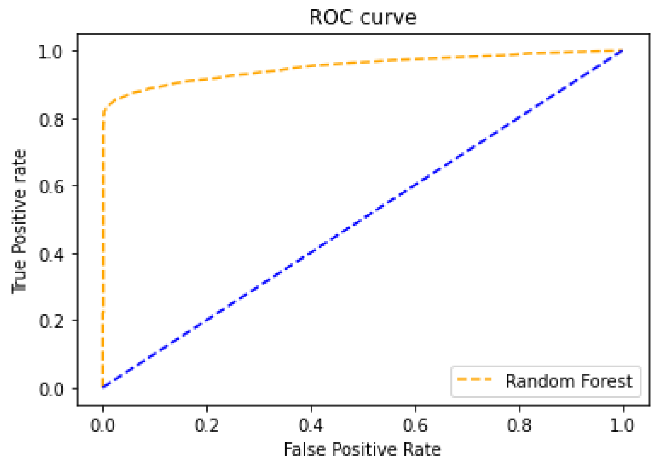

Fig. 2: Test ROC curve

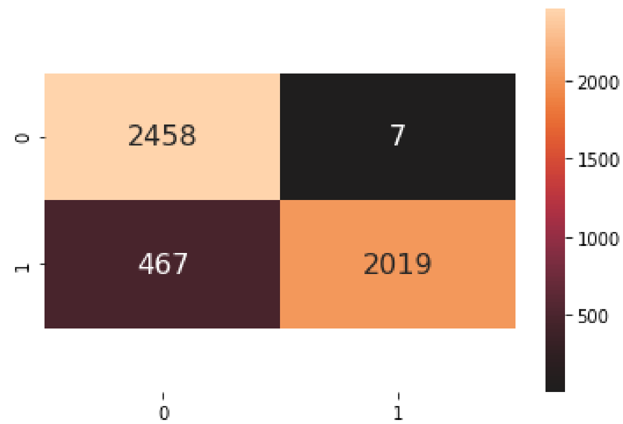

Fig. 3: Confusion matrix for the test set which contains 4951 samples (after applying SMOTE for augmentation and balancing) 1 = Likely-COVID-19, 0 = Non-Likely-COVID-19

we also achieved 90.46% for AUC and 99.71% for specificity, which demonstrate the feasibility of the developed framework for identifying COVID-19 cases from solely cough sound data.

References

[1] WHO Coronavirus (COVID-19) Dashboard https://covid19.who.int/ visited June 05, 2021

[2] P. Silva, E. Luz, G. Silva, G. Moreira, R. Silva, D. Lucio, D. Menotti, COVID-19 detection in CT images with deep learning: A voting-based scheme and cross-datasets analysis, Informatics Med. Unlocked. 20 (2020) 100427. https://doi.org/https://doi.org/10.1016/j.imu.2020.100427.

[3] V. Shah, R. Keniya, A. Shridharani, M. Punjabi, J. Shah, N. Mehendale, Diagnosis of COVID-19 using CT scan images and deep learning techniques, Emerg. Radiol. 28 (2021) 497–505. https://doi.org/10.1007/s10140-020-01886-y.

[4] V. Perumal, V. Narayanan, S.J.S. Rajasekar, Prediction of COVID-19 with Computed Tomography Images using Hybrid Learning Techniques, Dis. Markers. 2021 (2021) 5522729. https://doi.org/10.1155/2021/5522729.

[5] M. Berrimi, S. Hamdi, R.Y. Cherif, A. Moussaoui, M. Oussalah, M. Chabane, COVID-19 detection from Xray and CT scans using transfer

	Accuracy	Precision	Recall	F1-Score	AUC	Specificity	Training time
Fold-1	0.8977	0.9931	0.8024	0.8876	0.8984	0.9944	424
Fold-2	0.8861	0.9836	0.7868	0.8743	0.8868	0.9867	413
Fold-3	0.8992	0.9944	0.8044	0.8893	0.8999	0.9954	423
Fold-4	0.8909	0.9918	0.7897	0.8793	0.8915	0.9933	394
Fold-5	0.8934	0.9887	0.7972	0.8827	0.8940	0.9908	388
Avg $\pm$ sd	0.8935 $\pm$ 0.0047	0.9903 $\pm$ 0.0038	0.7961 $\pm$ 0.0068	0.8826 $\pm$ 0.0054	0.8941 $\pm$ 0.0047	0.9921 $\pm$ 0.0030	-
Test	**0.9042**	**0.9965**	**0.8121**	**0.8949**	**0.9046**	**0.9971**	-

TABLE V: Obtained k-fold cross-validation and average results of training random forest classifier on COUGHVID dataset, training time in seconds and test results on the unseen data.

learning, in: 2021 Int. Conf. Women Data Sci. Taif Univ. (WiDSTaif), 2021: pp. 1–6. https://doi.org/10.1109/WiDSTaif52235.2021.9430229.

[6] M.I. Elbakary, K. Iftekharuddin, COVID-19 detection using image analysis methods on CT images, in: Proc.SPIE, 2021. https://doi.org/10.1117/12.2581667.

[7] S. Asif, Y. Wenhui, H. Jin, Y. Tao, S. Jinhai, Classification of COVID-19 from Chest X-ray images using Deep Convolutional Neural Networks, MedRxiv. (2020). https://doi.org/10.1101/2020.05.01.20088211.

[8] T. Ozturk, M. Talo, E.A. Yildirim, U.B. Baloglu, O. Yildirim, U. Rajendra Acharya, Automated detection of COVID-19 cases using deep neural networks with X-ray images, Comput. Biol. Med. 121 (2020) 103792. https://doi.org/https://doi.org/10.1016/j.compbiomed.2020.103792.

[9] A. Makris, I. Kontopoulos, K. Tserpes, COVID-19 detection from chest X-Ray images using Deep Learning and Convolutional Neural Networks, MedRxiv. (2020). https://doi.org/10.1101/2020.05.22.20110817.

[10] Ai, T., Yang, Z., Hou, H., Zhan, C., Chen, C., Lv, W., Tao, Q., Sun, Z., & Xia, L. (2020). Correlation of Chest CT and RT-PCR Testing for Coronavirus Disease 2019 (COVID-19) in China: A Report of 1014 Cases. Radiology, 296(2), E32–E40. https://doi.org/10.1148/radiol.2020200642

[11] S. Tian, W. Hu, L. Niu, H. Liu, H. Xu, S.-Y. Xiao, Pulmonary Pathology of Early-Phase 2019 Novel Coronavirus (COVID-19) Pneumonia in Two Patients With Lung Cancer, J. Thorac. Oncol. 15 (2020) 700–704. https://doi.org/https://doi.org/10.1016/j.jtho.2020.02.010.

[12] A. Imran, I. Posokhova, H.N. Qureshi, U. Masood, M.S. Riaz, K. Ali, C.N. John, M.D.I. Hussain, M. Nabeel, AI4COVID-19: AI enabled preliminary diagnosis for COVID-19 from cough samples via an app, Informatics Med. Unlocked. 20 (2020) 100378. https://doi.org/https://doi.org/10.1016/j.imu.2020.100378.

[13] Irwin, R. S., & Madison, J. M. (2000). The diagnosis and treatment of cough. The New England journal of medicine, 343(23), 1715–1721. https://doi.org/10.1056/NEJM200012073432308

[14] Chang, A. B., Landau, L. I., Van Asperen, P. P., Glasgow, N. J., Robertson, C. F., Marchant, J. M., Mellis, C. M., & Thoracic Society of Australia and New Zealand (2006). Cough in children: definitions and clinical evaluation. The Medical journal of Australia, 184(8), 398–403. https://doi.org/10.5694/j.1326-5377.2006.tb00290.x

[15] Gibson, P. G., Chang, A. B., Glasgow, N. J., Holmes, P. W., Katelaris, P., Kemp, A. S., Landau, L. I., Mazzone, S., Newcombe, P., Van Asperen, P., Vertigan, A. E., & CICADA (2010). CICADA: Cough in Children and Adults: Diagnosis and Assessment. Australian cough guidelines summary statement. The Medical journal of Australia, 192(5), 265–271. https://doi.org/10.5694/j.1326-5377.2010.tb03504.x

[16] H. Chatrzarrin, A. Arcelus, R. Goubran, F. Knoefel, Feature extraction for the differentiation of dry and wet cough sounds, in: 2011 IEEE Int. Symp. Med. Meas. Appl., 2011: pp. 162–166. https://doi.org/10.1109/MeMeA.2011.5966670.

[17] J. Laguarta, F. Hueto, B. Subirana, COVID-19 Artificial Intelligence Diagnosis Using Only Cough Recordings, IEEE Open J. Eng. Med. Biol. 1 (2020) 275–281. https://doi.org/10.1109/OJEMB.2020.3026928.

[18] A. Hassan, I. Shahin, M.B. Alsabek, COVID-19 Detection System using Recurrent Neural Networks, in: 2020 Int. Conf. Commun. Comput. Cybersecurity, Informatics, 2020: pp. 1–5. https://doi.org/10.1109/CCCI49893.2020.9256562.

[19] A. Imran, I. Posokhova, H.N. Qureshi, U. Masood, M.S. Riaz, K. Ali, C.N. John, M.D.I. Hussain, M. Nabeel, AI4COVID-19: AI enabled preliminary diagnosis for COVID-19 from cough samples via an app, Informatics Med. Unlocked. 20 (2020) 100378. https://doi.org/https://doi.org/10.1016/j.imu.2020.100378.

[20] M. Pahar, M. Klopper, R. Warren, T. Niesler, COVID-19 Cough Classification using Machine Learning and Global Smartphone Recordings, ArXiv E-Prints. (2020) arXiv:2012.01926. https://ui.adsabs.harvard.edu/abs/2020arXiv201201926P.

[21] N. Chawla, K. Bowyer, L. Hall, W. Kegelmeyer, SMOTE: Synthetic Minority Over-sampling Technique, J. Artif. Intell. Res. 16 (2002) 321–357. https://doi.org/10.1613/jair.953.

[22] H. Skander, A. Moussaoui, M. Oussalah, M. Saidi, Gender Identification from Arabic Speech Using Machine Learning, in: 2020: pp. 149–162. https://doi.org/10.1007/978-3-030-58861-8_11.

[23] C. Brown, J. Chauhan, A. Grammenos, J. Han, A. Hasthanasombat, D. Spathis, T. Xia, P. Cicuta, C. Mascolo, Exploring Automatic Diagnosis of COVID-19 from Crowdsourced Respiratory Sound Data, in: Proc. 26th ACM SIGKDD Int. Conf. Knowl. Discov. & Data Min., Association for Computing Machinery, New York, NY, USA, 2020: pp. 3474–3484. https://doi.org/10.1145/3394486.3412865.

[24] B.L.Y. Agbley, J. Li, A. Haq, B. Cobbinah, D. Kulevome, P.A. Agbefu, B. Eleeza, Wavelet-Based Cough Signal Decomposition for Multimodal Classification, in: 2020 17th Int. Comput. Conf. Wavelet Act. Media Technol. Inf. Process., 2020: pp. 5–9. https://doi.org/10.1109/ICCWAMTIP51612.2020.9317337.

[25] L. Orlandic, T. Teijeiro, & D.A. Alonso (2020). The COUGHVID crowdsourcing dataset: A corpus for the study of large-scale cough analysis algorithms. ArXiv, abs/2009.11644.

[26] Maleki, M. (2021). Diagnosis of COVID-19 and Non-COVID-19 Patients by Classifying Only a Single Cough Sound. ArXiv, abs/2102.04880.

[27] B.L.Y. Agbley, J. Li, A. Haq, B. Cobbinah, D. Kulevome, P.A. Agbefu, B. Eleeza, Wavelet-Based Cough Signal Decomposition for Multimodal Classification, in: 2020 17th Int. Comput. Conf. Wavelet Act. Media Technol. Inf. Process., 2020: pp. 5–9. https://doi.org/10.1109/ICCWAMTIP51612.2020.9317337.

[28] M. Abdelwahab, C. Busso, Evaluation of syllable rate estimation in expressive speech and its contribution to emotion recognition, 2014. https://doi.org/10.1109/SLT.2014.7078620.

[29] Y. Gu, S. Chen, I. Marsic, Deep Multimodal Learning for Emotion Recognition in Spoken Language, in: 2018 IEEE Int. Conf. Acoust. Speech Signal Process., 2018: pp. 5079–5083. https://doi.org/10.1109/ICASSP.2018.8462440.

[30] S. Klaylat, Z. Osman, L. Hamandi, R. Zantout, Emotion recognition in Arabic speech, Analog Integr. Circuits Signal Process. 96 (2018) 337–351. https://doi.org/10.1007/s10470-018-1142-4.

[31] R. Zantout, S. Klaylat, L. Hamandi, Z. Osman, Ensemble Models for Enhancement of an Arabic Speech Emotion Recognition System, in: Lect. Notes Networks Syst., 2020: pp. 174–187. https://doi.org/10.1007/978-3-030-12385-7_15.

[32] L.A. Low, N.C. Maddage, M. Lech, L. Sheeber, N. Allen, Influence of acoustic low-level descriptors in the detection of clinical depression in adolescents, in: 2010 IEEE Int. Conf. Acoust. Speech Signal Process., 2010: pp. 5154–5157. https://doi.org/10.1109/ICASSP.2010.5495018.

[33] A. Pampouchidou, K. Marias, F. Yang, M. Tsiknakis, O. Simantiraki, A. Fazlollahi, M. Pediaditis, D. Manousos, A. Roniotis, G. Giannakakis, F. Meriaudeau, P. Simos, Depression Assessment by Fusing

High and Low Level Features from Audio, Video, and Text, 2016. https://doi.org/10.1145/2988257.2988266.
[34] F. Eyben, M. Wöllmer, B. Schuller, openSMILE – The Munich Versatile and Fast Open-Source Audio Feature Extractor, 2010. https://doi.org/10.1145/1873951.1874246.
[35] F. Weninger, F. Eyben, B. Schuller, M. Mortillaro, K. Scherer, On the Acoustics of Emotion in Audio: What Speech, Music, and Sound have in Common, Front. Psychol. 4 (2013) 292. https://doi.org/10.3389/fpsyg.2013.00292.
[36] H. He, Y. Ma, Imbalanced Datasets: From Sampling to Classifiers, in: Imbalanced Learn. Found. Algorithms, Appl., IEEE, 2013: pp. 43–59. https://doi.org/10.1002/9781118646106.ch3.
[37] [1] L. Breiman, Random Forests, Mach. Learn. 45 (2001) 5–32. https://doi.org/10.1023/A:1010933404324.