Analysis of cough sound measurements including COVID-19 positive cases: A machine learning characterization

Julio J. Valdés
Digital Technologies
National Research Council Canada
Ottawa, Canada
julio.valdes@nrc-cnrc.gc.ca

Pengcheng Xi
Digital Technologies
National Research Council Canada
Ottawa, Canada
pengcheng.xi@nrc-cnrc.gc.ca

Madison Cohen-McFarlane
Department of Systems and Computer Engineering
Carleton University
Ottawa, Canada
madison.mcfarlane@sce.carelton.ca

Bruce Wallace
Bruyère Research Institute
SAM³ Innovation Hub
Carleton University, Canada
wally@sce.carleton.ca

Rafik Goubran
Department of Systems and Computer Engineering
Carleton University
Ottawa, Canada
goubran@sce.carleton.ca

Frank Knoefel
Bruyère Research Institute
Elizabeth Bruyère Hospital
Carleton University, Ottawa, Canada
FKnoefel@bruyere.org

2021 IEEE International Symposium on Medical Measurements and Applications (MeMeA) | 978-1-6654-1914-7 /21/$31.00 ©2021 CRWN | DOI: 10.1109/MEMEA52024.2021.9478714

***Abstract*—Remote monitoring and measurement are valuable tools for medical applications and they are particularly important in the context of pandemic outbreaks, like the current COVID-19. This paper presents an analysis of sound measurements of cough events from the point of view of their predictive content with respect to identification of different types of cough, including positive COVID-19 cases. The data consisted of a collection of audio samples collected from different sources including dry, wet, whooping and COVID-19 coughs. Unsupervised and supervised machine learning techniques were used to reveal the underlying structure of the data, described by dissimilarity spaces constructed from pair-wise dynamic time warping measures derived from the original sound measurements. Intrinsic dimensionality, nonlinear mappings to low-dimensional spaces and visual cluster assessment techniques allowed a representation of the cough types distribution. Supervised classification techniques were used to obtain models identifying cough classes and high performance classifiers were obtained for most of them, including COVID-19. These results are preliminary and there is potential to improve, as they were obtained directly from a small dataset, without signal preprocessing (trimming, filtering, etc.), hyperparameter tuning, ensemble models, and class imbalance handling approaches.**

***Index Terms*—cough events, sound measurements, dynamic time warping, dissimilarity spaces, machine learning, COVID-19**

I. Introduction

The use of remote monitoring tools has been more readily adopted in recent years. This is especially true in the current climate of physical distancing due to the COVID-19 pandemic. From a medical perspective remote monitoring and measurement can provide essential information to physicians, especially now that remote appointments are being used more frequently. Measurements that have been investigated recently include; respiratory measurements [1] [2] [3] [4] [5], pre-existing condition monitoring [6] [7] [8] [9], and sleep monitoring [10] [11] [12].

With the rise of COVID-19, remote monitoring tools that can monitor or possibly identify respiratory conditions would be very beneficial. One of the most common symptoms of respiratory disease is cough, whose characteristics can provide medical insight to health care professionals when identifying the underlying cause (e.g. virus, bacteria, or acute/chronic condition) [13]. Cough characteristics are typically explored by a physician during an appointment by *1)* asking the patient to describe the characteristics of the cough, e.g. frequency and productivity, *2)* listening to involuntary coughs during the encounter, *3)* listening to the lungs of the patient during deep breathing and *4)* potentially asking the patient to produce a voluntary cough. However, voluntary coughs may have different characteristics when compared to an involuntary cough [14]. Self-report measures are well known to be inconsistent, which is especially true in the older adult population.

The automatic characterization of cough based on some of the current medical ontology (wet cough, dry cough, whooping cough) may prove beneficial as an additional source of information for physicians when making medical decisions. Furthermore, characterization of specific underlying condition associated with a particular cough (e.g. COVID-19) may have implication in disease identification, monitoring and contact tracing in public areas.

Studies have indicated that cough caused by COVID-19 is likely to have distinct latent features. In [15], the authors analyzed the pathomorphological changes caused by the COVID-19 in the respiratory system from the studies examining X-rays and CT scans of COVID-19 patients. Their study also included the autopsy report studies of deceased patients. Their findings suggest that cough sound signatures with COVID-19 are likely to have some idiosyncrasies stemming from the distinct underlying pathomorphological alterations. In an-

other study [16], the authors discovered vocal biomarkers of COVID-19 based on the coordination of subsystems of speech production involving respiration, phonation, and articulation. Their study hints that biomarkers derived from measures of vocal subsystem coordination provide an indicator of COVID-19 impact on respiratory function.

Recent studies indicate the feasibility of AI-enabled diagnosis of COVID-19 patients from cough samples. In [15], an AI engine could distinguish between COVID-19 patient coughs and several types of non-COVID-19 coughs with over 90% accuracy. The performance can be further improved if more and better data become available. In another study [17], the authors conducted analysis on a large-scale crowdsourced dataset of respiratory sounds collected to aid diagnosis of COVID-19. Their results indicate that even a simple binary machine learning classifier is able to classify correctly health and COVID-19 sounds (cough or breathing sound). Their models achieved an AUC (Area Under the Curve) of above 80% across all tasks. Very recent studies with large datasets incorporating laboratory molecular-test and clinically validated samples processed with deep learning techniques, have achieved very good results [18]. These recent research developments have therefore motivated us to develop machine learning approaches for distinguishing cough sounds including COVID-19 positive cases.

The novelty of our paper includes: *i)* the classification of between different types of coughs (dry, wet, whoop and COVID-19), and *ii)* an analysis using unsupervised machine learning methods for characterizing the cough data and its internal structure. The paper is organized as follows: Section II describes the data, Section III discusses the signal processing techniques used to construct the datasets processed, Section IV describes the unsupervised and supervised machine learning techniques (ML) used in the study. Section V presents the results and Section VI the conclusions and final remarks.

II. Cough Data

The dataset used in this study consisted of a collection of 215 cough audio samples compiled from three sources. The wet and dry coughs were labeled as such by a medical professional and were originally used in [4]. The whoop cough sounds were downloaded from [5] and were originally used to create a diagnosis tool for pertussis (whooping cough) . Only a subset of the links were still available on YouTube (23/38), which were manually segmented and labeled by a technician familiar with cough sound analysis. Finally COVID-19 positive cough events were used as the final class and were obtained from the NoCoCoDa database [13].

In the specific case of COVID-19 samples, the NoCoCoDa database contains cough events (in the form of wav format files), obtained from online interviews with COVID-19 positive individuals. All interviews were found online and were published by news sources. The dataset used in this study contains instances of four types of cough and the classes and their distribution are shown in Table I. Examples of cough audio signals (one from each class) are shown in Fig. 1.

TABLE I
COUGH CLASSES AND SIZES (215 OBJECTS)

index	Class	Size	Size (%)
1	Dry	19	8.33
2	Wet	27	11.84
3	Whoop	96	42.11
4	COVID-19	73	32.02

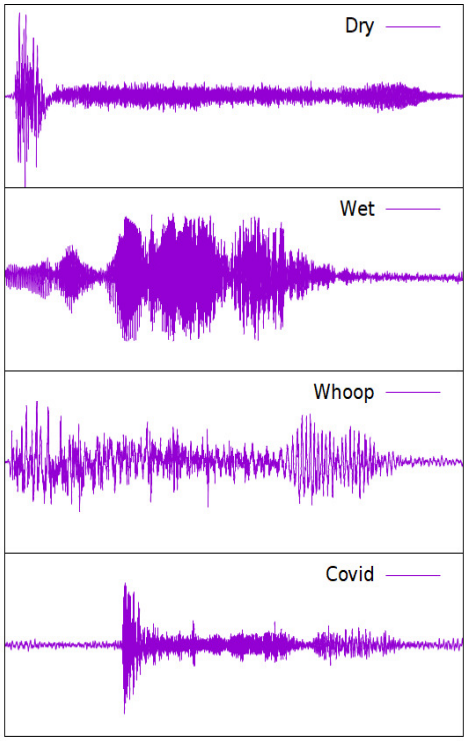

Fig. 1. Examples of cough audio signals for different classes.

III. Signal Processing

The signals from the 215 cough event recordings from the collated database were used in their original (raw) form. Typically, signals go through a preprocessing stage which entails trimming, noise and artifact removal (with different digital filtering techniques), transformations, etc. The present study aims at making first assessments of the data structure and the predictive potential of the original information. It is acknowledged that avoiding the always important preprocessing stage implies a risk with respect to the results, but it would reduce the degrees of freedom and biases of the data analysis process, which involves parameter-dependent algorithms. Accordingly, the results obtained should be considered as preliminary and baseline-defining.

A. Dynamic Time Warping

Dynamic Time Warping (DTW) is a classical method to compute a measure between two time series with different length, used to compare their similarity [19], [20]. It is based on the notion of optimal match between two time sequences, where each index from the first series is matched with one

or more indices from the other one and conversely, following some monotonicity rules between the indices of the two series. The DTW measure is not based on features extracted from the signals (as required by other measures), but directly on the signal values in a holistic manner. DTW is not a distance in the metric sense as it does not guarantee the triangular inequality. However, it is a dissimilarity which has been extensively and successfully used in time series analysis, signal processing and information retrieval. In the present data, some signals have initial delays at the beginning of their cough events, but they have near zero values which have little effect on DTW dissimilarities. In this paper, the only operation performed on the original signals in order to create the data used by machine learning methods of the next sections was the computation of a pairwise DTW matrix.

IV. Machine Learning Techniques

A. Unsupervised Methods

1) Intrinsic Dimension: Data objects are typically represented in feature spaces, but they can also be described by a structure composed of pairwise distances or dissimilarities. The dimensions of these descriptor spaces can be large, affecting the performance of statistical and machine learning procedures (the curse of dimensionality). In reality, the data often exist in low dimension non-linear manifolds within the higher dimensional descriptor spaces. The intrinsic dimension (IDim) is defined as the dimension of the embedded manifolds actually containing the data and revealing these manifolds helps to understand the internal structure of the data and to improve the performance of machine learning techniques. Four different approaches for estimating the intrinsic dimensionality were used in this paper [21], [22], [23] [24].

A low dimensional transformation of the original space provides an environment representing the data where the information could be visualized. If the intrinsic dimension is sufficiently low, the new space would accurately represent the structure of the information contained in the data. Otherwise, it would provide an approximation.

2) Low Dimensional Mapping: Uniform Manifold Approximation and Projection (UMAP) is a general-purpose manifold learning and dimension reduction algorithm [25], [26], that can be used for nonlinear mapping, non-linear dimension reduction and visualization purposes. The algorithm is founded on three assumptions about the data: *i)* The data is uniformly distributed on a Riemannian manifold (one with a collection of inner products for every point on a tangent space); *ii)* The Riemannian metric is either locally constant or can be approximated as such, and *iii)* The manifold is locally connected. From these assumptions the manifold is modeled with a fuzzy topological structure. From it, the mapping is found by searching for a low dimensional projection of the data that has the closest possible equivalent fuzzy topological structure.

This approach was used for approximating the manifold defined by the DTW distance matrix as a 3D space, based on the $IDim$ obtained with the methods from Section. IV-A1. The final transformation is given by the composition $\varphi = (\mathcal{U} \circ \mathcal{P})$, where $\mathcal{U}$ is the nonlinear UMAP mapping and $\mathcal{P}$ is a principal components transformation applied to the nonlinear space with a number of components equal to the number of target dimensions. As a result, φ is a canonical mapping where the axis are ordered according to a monotonic decrease of variance in the nonlinear space.

3) Cluster Tendency Assessment: Clustering is a very important component of the exploratory data analysis process, crucial for the understanding of the data, prior to subsequent stages involving modeling, prediction, etc. The clustering tendency assessment problem consists of estimating whether there are (natural) clusters in the data and their quantity. The *VAT* family of algorithms [27], [28], [29], [30], [31] are among the few oriented to this problem (VAT stands for visual assessment of (cluster) tendency). The general idea is to reorder a pairwise dissimilarity matrix and then display it as a dissimilarity image, where possible clusters in the data can be identified as dark blocks aligned along the main diagonal. In particular, the iVAT (improved VAT) member was used [31], which uses geodesic distances derived from the original dissimilarities to obtain an improved assessment of the presence of clustering structure within the data. An example of the iVAT's way of assesing clustering structure is shown in Fig. 2. In this paper, iVAT was used to make an initial assessment of the structure of the DTW matrix derived directly from the raw signals.

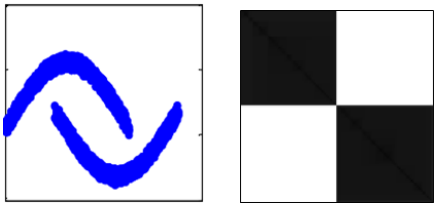

Fig. 2. Example of a synthetic dataset containing two irregular clusters and the corresponding iVAT representation of the distance structure.

B. Dissimilarity Spaces

A Dissimilarity Space [32] is a mapping defined by a dissimilarity measure between N objects of any kind from a given set ($\mathcal{O}$). The measure does not necessarily need to be a metric in the triangular inequality sense. Typically, a subset of prototype objects are set forth $\mathcal{P} \subseteq \mathcal{O}$, $P = card(\mathcal{P})$ and the dissimilarities are considered only with respect to the objects in $\mathcal{P}$ ($card(X)$ is the cardinality of set X). The dissimilarity space is then given by the matrix

$$D_s(\mathcal{O}, \mathcal{P}) = \begin{bmatrix} \delta_{o_1,p_1} & \cdots & \delta_{o_1,p_P} \\ & \ddots & \\ \delta_{o_N,p_1} & \cdots & \delta_{o_N,p_P} \end{bmatrix}_{N \times P} \quad (1)$$

where δ_{o_i,p_k} is the dissimilarity between objects $o_i \in \mathcal{O}$ and $p_k \in \mathcal{P}$. In the present case, δ_{o_i,p_k} is the DTW distance between signals o_i and p_k and this matrix can be used as a data matrix by machine learning procedures. Multiple dissimilarity spaces can be constructed by incorporating additional dissimilarity measures between signals.

C. Supervised Methods

The k-nearest neighbors (k-NN) is a classical non-parametric, instance-based machine learning algorithm, extensively used for both classification and regression problems. In the former case, the decision for a given object is the result a majority vote among the classes corresponding to the k-most similar objects.

Extreme Gradient Boosting (XGBoost) [33] is a tree-based algorithm where ensemble techniques are used so that previous model errors are resolved in the new models (gradient boosting). It is considered one of the most successful ML algorithms with high performance in most situations. Random Forests (RF) is a classification algorithm operating with a collection of classification trees [34]. It samples objects from the training set at random (with replacing), growing a tree with the sample. In the same way, a subset of attributes of a given fixed size are randomly chosen using them for splitting the nodes of the trees, which are ensembled using a voting scheme. RF and XGBoost are closely related and work well in most problems, including those involving a large number of attributes (the present case). They were applied to the dissimilarity space data from Section IV-B.

The operation of the supervised classifiers used in this paper can be parametrized by the mapping $f : \mathcal{O} \rightarrow \mathcal{C}$, where $\mathcal{C} = \{c_1, \cdots, c_n\}$ is class set ($n = 4$ is its cardinality). f operates on objects $\mathcal{O} = \{O_1, \cdots, O_N\}$ of the dissimilarity space with $O_i \in \mathbb{R}^{N_{tr}}$, where N_{tr} are the objects in the training set (the prototypes). Since the dissimilarities are computed with respect to the objects in the training set only, there is no information leakage. All algorithms (kNN, XGBoost and Random Forests) operate on the dissimilarity space. The $N = 215$ objects were divided into training and testing sets using a $90\%/10\%$ random stratified split (193/22 objects), implying $N_{tr} = 193$, where the 10% testing was held out and used for validation of the classifiers. In addition, 10-fold crossvalidation (CV) was used to obtain appraisals of the generalization capability of the classifiers.

V. Results

A. Unsupervised Analysis

The first step is an assessment of the structure of the data and the mutual relationships between the objects in the original space. For that purpose, a 215×215 dynamic time warping distance matrix was computed for the cough audio signals. Based on the data structure determined by that matrix and regardless of the distribution of the classes of interest in the study, the next step is to investigate whether there are natural groupings (clusters) induced by the pairwise DTW distances. The clustering tendency assessment obtained with the iVAT procedure (Section IV-A3) is shown in Fig. 3.

The image representing the reordered DTW matrix does not reveal clearly defined clusters, even though iVAT uses a path-based distance transform to improve the effectiveness of the basic VAT process. The algorithm seems to be affected by noise and possibly also by bridge points between clusters in the dataset, making the result in-conclusive. However, it suggests the presence of a relatively higher density core region (upper left area in Fig. 3), progressively becoming more diffuse (lower-right of the figure).

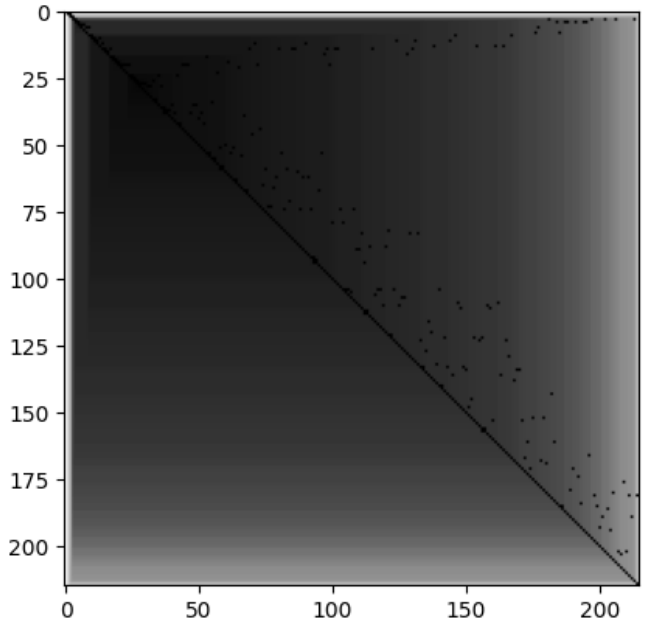

Fig. 3. iVAT representation of the DTW dissimilarity matrix between the 215 cough audio samples.

The 215 DTW attributes associated to each of the samples in the dissimilarity space could be considered as the dimension of that space to represent the data, but rarely it is the real dimension of the manifold that actually contains the objects. The intrinsic dimensionality estimates for the DTW distance matrix using the techniques presented in Section IV-A1 were: *i)* Correlation Integral: 2.64, *ii)* Maximum Likelihood: 3.63, *iii)* Nearest Neighbor Information: 2.57, *iv)* convergence property of U-stats: 1 , *v)* Packing Number method (Greedy alg.): 1.35.

Clearly, most estimates coincide in identifying a rather low dimensional manifold (around $[2, 4]D$), embedded in the original $215D$ space. An approximation to the low-D manifold could be obtained with nonlinear mappings and inspected with visual techniques. The 3D space obtained with the $(\mathcal{U} \circ \mathcal{P})$ transformation is shown in Fig. 4 , where class membership was incorporated a-posteriori as colors as an aid to interpretation (the mapping itself is unsupervised). Due to the canonical nature of the mapping, the smallest variance of the nonlinear space occurs along the Z-axis. The actual variance distribution is 82.5% along the X axis, 14.3% along Y and 3.2% along Z, resulting in a quasi-2D structure, in agreement with the intrinsic dimension findings.

The distribution of the objects in the 3D space does not exhibit sharply defined clusters. Rather, there are regions of relative higher density, particularly around the periphery, and a central, lower density area, coinciding with the pattern

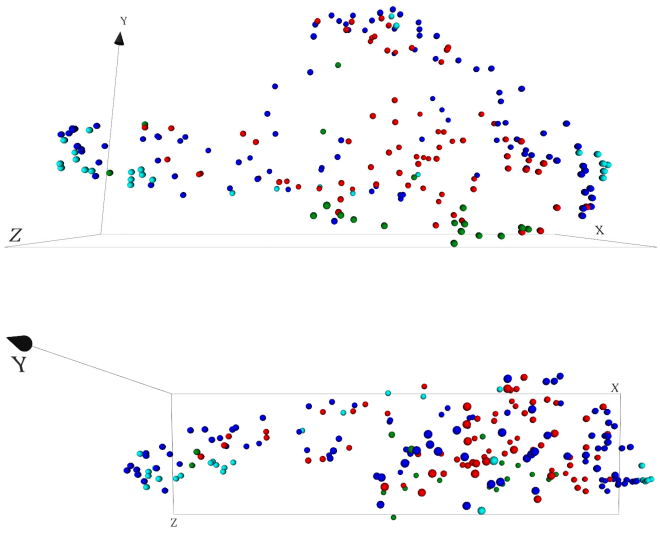

Fig. 4. ($\mathcal{U} \circ \mathcal{P}$) 3D nonlinear mapping of the DTW matrix. Top: General view, mostly showing the XY plane. Bottom: View from above, mostly showing the XZ plane. Colors indicate class membership: Green: Dry, Cyan: Wet, Blue: Whoop, Red: COVID-19.

suggested by the iVAT algorithm (Fig. 3). Roughly, the class distribution patterns are: *i)* The Dry class members (green) are mostly located at the lower-right periphery, *ii)* those of the Wet class (cyan) are mostly at the lower-left and lower-right, *iii)* the Whoop class (blue) predominates at the upper-edge and the lower-left regions and *iv)* the central-inner region is mostly occupied by the COVID-19 class, usually close to elements of the Whoop class. Most classes exhibit multimodal distributions and there are many cases of elements between the modes.

B. Supervised Methods

The kNN classifier was used with $k = 4$ neighbours, whereas XGBoost and Random Forest has the default parameters specified in [35]. The results obtained on the testing set are shown in Table II, summarized by the typical precision (Prec), recall and F1 measures ($2(Prec \times Recall)/(Prec + Recall)$). The row order correspond to the four classes (Dry, Wet, Whoop and COVID-19) considered in that order (COVID-19 results in bold).

TABLE II
CLASSIFICATION QUALITY MEASURES FOR THE TESTING SET (KNN, XGBOOST AND RANDOM FOREST). CLASSES ARE (FROM TOP TO BOTTOM) :$\{Dry, Wet, Whoop, Covid-19\}$. THE COVID-19 CLASS IS HIGHLIGHTED.

kNN			XGBoost			Random Forest		
Prec	**Rec**	**F1**	**Prec**	**Rec**	**F1**	**Prec**	**Rec**	**F1**
0.50	1.00	0.67	1.00	1.00	1.00	1.00	0.50	0.67
0.60	1.00	0.75	1.00	0.33	0.50	1.00	0.33	0.50
0.86	0.60	0.71	0.70	0.70	0.70	0.70	0.70	0.70
0.83	**0.71**	**0.77**	**0.67**	**0.86**	**0.75**	**0.60**	**0.86**	**0.71**

For kNN the cross-validation mean accuracy on training was 56.45%, which improved for XGBoost (68.89%) and Random Forest (68.33%). On the testing set, kNN had a better recall for the Dry and Wet classes than for Whoop and COVID-19. However, precision was better for the latter two classes than for the first two. XGBoost and Random Forest exhibited a good recall for the Dry and COVID-19 classes, a moderate one for Whoop and a low performance for the Wet class. Notably, both recalled COVID-19 with 86% accuracy, even though precision-wise kNN was better. For that class, XGBoost compensated the drop in precision with respect to kNN with a high increase in recall. For the Wet class, kNN outperformed XGBoost and Random Forest recall-wise, which is a consequence of the highly multimodal distribution of that class, which tends to favour instance-based algorithms.

In the overall, considering the sum of the F1 index, XGBoost was slightly better than kNN, but the complexity of the class distribution makes no algorithm better when considering all classes and measures. With the exception of the Whoop class, for each cough type there was always a classifier that delivered high performance, precision of recall-wise. In the case of Whoop this behavior is also observed, but in a lesser degree, as the top recall was 0.7 and the top precision 0.86. Ensemble models perform better in such cases.

The availability of more data samples, particularly for the much harder Wet class would create better conditions for the machine learning classifiers. It is noteworthy that in the especially important case of the COVID-19 class it was possible to obtain high rank classification accuracies directly from the audio cough signals. Clearly, the aggregation of other sources of information like clinical data, X-rays and measurements from other medical information acquisition devices would improve the preliminary results obtained. In the same way, the introduction of dedicated data preprocessing strategies (e.g. appropriate signal filtering) also represents a source of potential improvement. The machine learning techniques should include hyper-parameter optimization, class imbalance handling, and particularly ensemble methods, working with multiple dissimilarity spaces and other strategies.

VI. CONCLUSIONS

Sound recordings of cough events contain predictive information with respect to the identification of different types of cough, including COVID-19. Unsupervised machine learning methods revealed a complex structure for the dynamic time warping distance between the audio signals without neatly defined clusters. However, there are class distribution patterns within the signal space with enough structure as to enable supervised classifiers to achieve high performance for the Dry, Whoop and COVID-19 cough types (the Wet was not so effectively retrieved, in the framework of individual classifiers).

The results presented here should be considered as baseline since they were obtained directly from the raw audio signals of a small dataset, without dedicated preprocessing like signal trimming, filtering, etc. Moreover, no hyperparameter tuning was applied to the machine learning techniques used, which did not include ensemble models, nor elaborated schemes for class imbalance. For classifying a new sample signal, a feature vector is constructed by computing the dissimilarities between

the new signal and the training set prototypes. Then, the classifier is applied to the feature vector to obtain the predicted class. There is potential to improve the characterization and the classification results obtained, including a more comprehensive analysis of the methods. These should be aspects to be considered in future studies.

ACKNOWLEDGEMENT

This work has been supported by the Pandemic Response Challenge Program at National Research Council of Canada. The authors would like to thank Jean-François Houle, Robert DiRaddo and Andrew Scheidl, program and pillar leads, for their support. We are thankful to Elena Di Francesco, program challenge officer, for her help and assistance. We also thank Di Jiang and Andrew Law for their help during project planning.

REFERENCES

[1] H. Azimi, S. S. Gilakjani, M. Bouchard, S. Bennett, R. A. Goubran, , and F. Knoefel, "Breathing signal combining for respiration rate estimation in smart beds," in *Proc. IEEE Int. Symp. Med. Meas. Appl. (MeMeA)*, IEEE, Ed., Jul. 2017, pp. 303—-307.

[2] C. Uysal, A. Onat, and T. Filik, "Non-contact respiratory rate estimation in real-time with modified joint unscented kalman filter," *IEEE Access*, vol. 8, pp. 99 445—-99 457, 2020.

[3] H. Zhao, H. Hong, L. Sun, Y. Li, C. Li, , and X. Zhu, "Noncontact physiological dynamics detection using low-power digital-if doppler radar," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 7, pp. 1780—-1788, 2017.

[4] H. Chatrzarrin, A. Arcelus, R. Goubran, , and F. Knoefel, "Feature extraction for the differentiation of dry and wet cough sounds," in *Proc. IEEE Int. Symp. Med. Meas. Appl. (MeMeA)*. IEEE, 2011, pp. 162—-166.

[5] R. X. A. Pramono, S. A. Imtiaz, and E. Rodriguez-Villegas, "A cough-based algorithm for automatic diagnosis of pertussis," *PLoS One*, vol. 11, no. 9, pp. 1—-20, 2016.

[6] M. Cohen-McFarlane, J. R. Green, R. A. Goubran, , and F. Knoefel, "Smart monitoring of fluid intake and bladder voiding using pressure sensitive mats," in *Proc. IEEE Eng. Med. Biol. Soc.* IEEE, 2016, p. 4921–4924.

[7] W. Xie, P. Gaydecki, , and A. L. Caress, "An inhaler tracking system based on acoustic analysis: Hardware and software," *IEEE Trans. Instrum. Meas.*, vol. 68, no. 11, pp. 4472—-4480, 2019.

[8] P. Casti, A. Mencattini, M. C. Comes, G. Callari, D. Di Giuseppe, S. Natoli, M. Dauri, E. Daprati, and E. Martinelli, "Calibration of vision-based measurement of pain intensity with multiple expert observers," *IEEE Transactions on Instrumentation and Measurement*, vol. 68, no. 7, pp. 2442–2450, 2019.

[9] B. Wallace, F. Knoefel, R. Goubran, P. Masson, A. Baker, B. Allard, V. Guana, and E. Stroulia, "Detecting cognitive ability changes in patients with moderate dementia using a modified "whack-a-mole" game," *IEEE Transactions on Instrumentation and Measurement*, vol. 67, no. 7, pp. 1521–1534, 2018.

[10] J. Luo, H. Liu, X. Gao, B. Wang, X. Zhu, Y. Shi, X. Hei, and X. Ren, "A novel deep feature transfer-based osa detection method using sleep sound signals," *A novel deep feature transfer-based OSA detection method using sleep sound signals*, vol. 41, no. 7, p. 075009, 2020.

[11] F. Deng, J. Dong, X. Wang, Y. Fang, Y. Liu, Z. Yu, J. Liu, and F. Chen, "Design and implementation of a noncontact sleep monitoring system using infrared cameras and motion sensor," *IEEE Transactions on Instrumentation and Measurement*, vol. 67, no. 7, pp. 1555–1563, 2018.

[12] I. Yilmaz, "Angiotensin-converting enzyme inhibitors induce cough," *Turkish Thoracic Journal*, vol. 20, no. 1, pp. 36—42, 2019.

[13] M. Cohen-McFarlane, R. Goubran, and F. Knoefel, "Novel coronavirus cough database: NoCoCoDa," *IEEE Access*, vol. 30, pp. 1–8, 2020.

[14] C. Magni, E. Chellini, F. Lavorini, G. A. Fontana, , and J. Widdicombe, "Voluntary and reflex cough: Similarities and differences," *Pulm. Pharmacol. Ther.*, vol. 24, no. 3, pp. 308—-311, 2011.

[15] A. Imran, I. Posokhova, H. N. Qureshi, U. Masood, M. S. Riaz, K. Ali, C. N. John, M. I. Hussain, and M. Nabeel, "Ai4covid-19: Ai enabled preliminary diagnosis for covid-19 from cough samples via an app," *Informatics in Medicine Unlocked*, vol. 20, p. 100378, 2020. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S2352914820303026

[16] T. F. Quatieri, T. Talkar, and J. S. Palmer, "A framework for biomarkers of covid-19 based on coordination of speech-production subsystems," *IEEE Open Journal of Engineering in Medicine and Biology*, vol. 1, pp. 203–206, 2020.

[17] C. Brown, J. Chauhan, A. Grammenos, J. Han, A. Hasthanasombat, D. Spathis, T. Xia, P. Cicuta, and C. Mascolo, "Exploring automatic diagnosis of covid-19 from crowdsourced respiratory sound data," in *Proc. KDD '20*, ser. KDD '20. New York, NY, USA: Association for Computing Machinery, 2020, p. 3474–3484. [Online]. Available: https://doi.org/10.1145/3394486.3412865

[18] J. Andreu-Perez, H. Perez-Espinosa, E. Timonet, M. Kiani, M. I. Giron-Perez, A. B. Benitez-Trinidad, D. Jarchi, A. Rosales, N. Gkatzoulis, O. F. Reyes-Galaviz, A. Torres, C. Alberto Reyes-Garcia, Z. Ali, and F. Rivas, "A Generic Deep Learning Based Cough Analysis System from Clinically Validated Samples for Point-of-Need Covid-19 Test and Severity Levels," *IEEE Transactions on Services Computing*, p. 1, 2021.

[19] H. Sakoe and S. Chiba, "Dynamic programming algorithm optimization for spoken word recognition," *IEEE Trans. on Acoust., Speech, and Signal Process.*, vol. ASSP 26, pp. 43–49, 1978.

[20] E. Keogh, "Exact indexing of dynamic time warping," in *Proc. 28th International Conference on Very Large Data Bases*, Hong Kong, 2002, pp. 406–417.

[21] P. Grassberger and I. Procaccia, "Measuring the strangeness of strange attractors," *Physica*, vol. D, no. 9, pp. 189–208, 1983.

[22] J. Costa and A. Hero, "Geodesic entropic graphs for dimension and entropy estimation in manifold learning," *IEEE Transactions on Signal Processing*, vol. 52, no. 8, pp. 2210–2221, 2004.

[23] E. Levina and P. Bickel, "Maximum likelihood estimation of intrinsic dimension," *Advances in Neural Information Processing Systems*, vol. 17, pp. 777–784, 2005.

[24] H. Hino, *Package ider*, 2017, https://cran.r-project.org/web/packages/ider/ider.pdf. [Online]. Available: https://cran.r-project.org/web/packages/ider/ider.pdf

[25] L. McInnes, J. Healy, and J. Melville, "UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction," *ArXiv e-prints*, Feb. 2018.

[26] L. McInnes, J. Healy, N. Saul, and L. Grossberger, "Umap: Uniform manifold approximation and projection," *The Journal of Open Source Software*, vol. 3, no. 29, p. 861, 2018.

[27] J. Bezdek and R. J. Hathaway, "VAT: A tool for visual assessment of (cluster) tendency," in *Proc. IJCNN 2002*. IEEE, 2002, pp. 2225–2230.

[28] T. C. Havens, J. C. Bezdek, J. M. Keller, and M. Popescu, "Clustering in ordered dissimilarity data," *Int. J. Intell. Systems*, vol. 24, no. 5, pp. 504–528, 2006.

[29] I. Sledge, J. C. B. T. C. Havens, J. M. Huband, and J. M. Keller, "Finding the number of clusters in ordered dissimilarities," *Soft Computing*, vol. 13, no. 12, pp. 1125–1142, 2009.

[30] L. Wang, U. Nguyen, J. Bezdek, C. Leckie, and K. Ramamohanarao, "iVAT and aVAT: Enhanced visual analysis for cluster tendency assessment," *PAKDD 2010. LNCS*, vol. 6118, pp. 16–27, 06 2010.

[31] T. Havens and J. Bezdek, "iVAT an efficient formulation of the improved visual assessment of cluster tendency (iVAT) algorithm," *IEEE TKDE*, vol. 24, no. 5, pp. 813–822, 2012.

[32] E. Pekalska and R. Duin, *The dissimilarity representation for pattern recognition. Foundations and applications.* Singapore: World Scientific, 2005.

[33] C. T. and G. C., "XGBoost: A scalable tree boosting system," in *Proc. of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ser. KDD '16. New York, NY, USA: ACM, 2016, pp. 785–794. [Online]. Available: http://doi.acm.org/10.1145/2939672.2939785

[34] L. Breiman, "Random forest," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

[35] "Xgboost package," 2020, https://xgboost.readthedocs.io/en/latest/parameter.html.