2022 13th International Conference on Information, Intelligence, Systems & Applications (IISA) | 978-1-6654-6390-4/22/$31.00 ©2022 IEEE | DOI: 10.1109/IISA56318.2022.9904377

Quantum Machine Learning for Audio Classification with Applications to Healthcare

Michael Esposito
SOLS, SenSIP Center
Arizona State University
Tempe, USA
mjesposi@asu.edu

Glen Uehara
School of ECEE, SenSIP Center
Arizona State University
Tempe, USA
guehara@asu.edu

Andreas Spanias
School of ECEE, SenSIP Center
Arizona State University
Tempe, USA
spanias@asu.edu

***Abstract*— Accessible rapid COVID-19 testing continues to be necessary and several studies involving deep neural network (DNN) methods for detection have been published. As part of a sponsored NSF I/UCRC project, our team explored the use of deep learning algorithms for recognizing COVID-19 related cough audio signatures. More specifically, we have worked with several DNN algorithms and cough audio databases and reported results with the VGG-13 architecture. In this paper, we report a study on the use of quantum neural networks for audio signature detection and classification. A hybrid quantum neural network (QNN) model for COVID-19 cough classification is developed. The design of the QNN simulation architecture is described and results are given with and without quantum noise. Comparative results between classical and quantum neural network methods for COVID-19 audio detection are also presented.**

Keywords—quantum computing, quantum machine learning, COVID-19, cough audio, spectral features, quantum noise

I. Introduction

Machine learning for audio detection and classification of breathing pathologies has been previously studied [1,2]. Recent studies on breathing pathologies concentrated on COVID-19 detection using audio cough patterns and machine learning [3-8]. More specifically, features obtained from spectrograms of COVID cough audio have been used in several recent studies. One of the early challenges was the availability of COVID coughing audio databases. Several initiatives were launched as early as the summer of 2020, and freely accessible databases for research became available. One of the early ones was the "Coswara" dataset [4] of cough and other sounds from positive and negative samples of COVID-19 audio. The "COUGHVID" dataset formed from crowd-sourced positive and negative cough samples across a wide range of demographic backgrounds was also published [5]. A more recent database for the DiCOVA challenge in Interspeech 2021 was described in [9].

The motivation to study audio-based COVID-19 features stems from the fact that COVID-19 causes disruptions in vocal tract tissues [8]. As a result, patients can generate specific patterns in recorded coughing audio. Several research groups began studying spectral features for cough audio signature characterization. Mel Frequency Cepstral Coefficients (MFCCs) [10], linear prediction models, and other statistical methods have previously been used as features for recognizing speech and audio sounds [11-13]. Our team also explored several audio features for neural network-based COVID detection. Our studies [6,7] determined that log-mel spectrogram representations provided the best results. Other studies [14] also determined that log-mel spectrograms generalize better than those based on traditional speech signal optimized features.

Our previous machine learning studies relied on classical neural networks and produced 60-82% accuracy results. In particular, our study that used log-mel spectrograms in conjunction with the VGG-13 neural network [6] provided more than 80% detection accuracy and was ranked relatively high in the DiCOVA 2021 challenge. Motivated by these results and also our recent efforts to develop hybrid quantum neural network models, we explore in this paper the use of QNN models for COVID-19 detection. The use of quantum computing for machine learning was the subject of several studies which were surveyed in [15].

One of the main challenges in quantum machine learning is access to actual quantum computers, which are currently expensive. However, several companies have developed quantum simulators [16-20] and provide access to quantum simulators through cloud or desktop computing. In addition, access to actual quantum computers is also provided, though current availability is very limited for large tasks. One of the greatest challenges is the management of quantum noise [21] and quantum precision, both of which will be examined in this paper specifically for COVID-19 cough audio classification.

Quantum machine learning simulation models require the design of quantum circuits. Data flow and qubit precision models must be formed in order to train and classify audio signatures. Audio processing and classification studies using QNNs have been published in [22,23]. Motivated by these studies, we began examining the feasibility of COVID-19 detection using QNNs. Challenges include quantum precision and handling quantum noise. In addition, access to simulators, design of quantum circuits, and latency in training QNNs are anticipated problems. Nevertheless, the potential in the future to take advantage of the benefits of fast and secure computation is motivating our study. Contributions of this study include a) the design of hybrid QNN algorithms for COVID-19 detection, b) characterization of the performance with different numbers of qubits, and c) effects of quantum noise. The rest of the paper is organized as follows. Section II presents the quantum circuit design process, Section III gives our results, and Section IV presents our conclusions.

II. Hybrid Quantum Neural Network Design

A. Challenges, Dataset and Preprocessing

Challenges in running the quantum machine learning simulations included long run times and careful pre-processing of data to extract robust features. To produce detection results comparable to previous classical computing simulations, we needed to optimize feature extraction and the QNN model. This involved identifying and centering of the cough audio signal within the log-mel spectrograms.

For this application, COVID-19 coughing audio from the DiCOVA [9] dataset, supplemented with samples from the COUGHVID [5] dataset, was used. The audio files were preprocessed using silence detection. A minimum cough length of 200ms was established during this segmenting. Additionally, 200ms of starting and ending silence was kept for each audio segment. Log-mel spectrograms were then generated from the audio files with a hop length of 128, a window length of 1024, and 60 mel bands.

B. Quantum Circuit and Feature Extraction

We explored several tools for designing quantum circuits including PennyLane [24]. PennyLane is a Python quantum computing library developed by Xanadu to simulate qubits, quantum circuits, and quantum machine learning architectures. This library allows for both coherent [25] and incoherent [26] quantum noise simulation. Coherent quantum noise is a more predictable, systematic noise that can result from quantum hardware that is not properly calibrated. Coherent noise, in general, may be filtered out using a standardized process across the hardware. Incoherent quantum noise tends to be less predictable than coherent quantum noise. The PennyLane quantum computing library can also simulate state vector qubits based on purely mathematical operations. A state vector simulator generally provides the ideal results in a quantum computing simulation.

Feature extraction was performed using two different quantum circuits, one constructed with two qubits and one with four qubits. The two-qubit circuit is shown in Fig. 1, and the four-qubit circuit is shown in Fig. 2. The quantum circuit features either two or four qubits depending on the simulation, each with an RY gate, then Unitary gates in which a set of gates with continuously changing weights performs rotations.

Fig. 1. A block diagram representation of the two-qubit circuit used for audio feature extraction.

Finally, the qubits are measured and mapped to classical bits. This particular quantum circuit was constructed using random circuit parameters [27] due to the robustness of random quantum circuits at relatively low complexity [28]. This particular instance of a random quantum circuit was generated using the PennyLane [24] RandomLayers template in which randomly chosen qubits are acted upon by layers of randomly chosen single-qubit rotations and 2-qubit entangling gates. We use a random circuits sampling process similar to that proposed in [29]. In our simulation, we used a range of one to four quantum circuit duplication layers during feature extraction to compare results. One quantum circuit layer corresponds to one set of RY, Unitary, and measurement gates. Duplication of this circuit combination refers to additional layers. Quantum convolution was performed during feature extraction of the audio data, in which the input audio was convolved with many applications of the same quantum circuit.

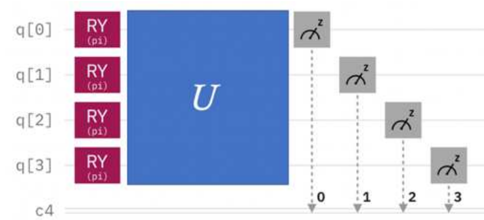

Fig. 2. A block diagram represents the four-qubit quantum circuit used for feature extraction in the audio classification task.

C. Neural Network Models and Training

Quantum-classical hybrid architectures [30-33] are evaluated in this study. Quantum convolutional ("quanvolutional") layers are combined with a classical neural network model for the classification of the audio data as seen in Fig. 3. Benefits of quantum-classical hybrid systems, when compared to fully quantum-based deep learning systems, include dramatically reduced training time and reduced memory requirements. We note that a recurrent neural network (RNN) [44] was previously shown [45] to perform well for audio classification tasks.

Three different models were trained and evaluated in this study: an RNN, a convolutional neural network (CNN), and a quanvolutional neural network (QNN). All three neural networks were trained using an 80/20 train/test split and were trained for 30 epochs. Test accuracy is tabulated in Table II.

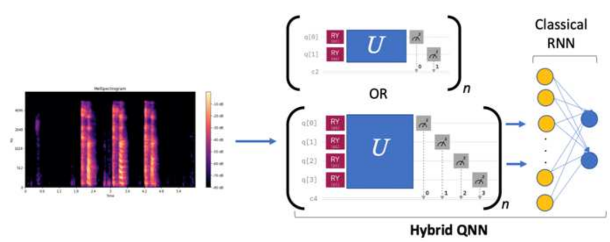

Fig. 3. A basic framework for the quantum ML data pipeline used in this study. Quantum convolution is performed in order to extract features from log-mel spectrograms and the features are then used to train a classical recurrent neural network.

III. Results

Our SenSIP team has previously engaged in several machine learning studies and presented results in speech, audio and various sensor and energy applications [34-42]. Quantum circuits for machine learning applications have been demonstrated in [30,31,42]. In the following, we present sound recognition results using quantum neural networks starting first with clean speech and then continuing with COVID-19 coughing sounds obtained from crowdsourced databases.

A. QNN Results with Clean Speech Signals

To validate the aforementioned quantum circuit designs, we first run the quantum simulations with clean speech signals. Our results with clean speech are tabulated in Table 1 and demonstrate that our quantum designs and our simulation software worked very well providing more than 90% accuracy.

TABLE 1. QNN RESULTS FOR CLEAN SPEECH.

Method	Test Accuracy (%)
QNN - 2 Qubits, 1 Layer	93.6
QNN - 4 Qubits, 1 Layer	96.4

From our simulation results listed in Table 1, we observe that even a single "quanvolutional" layer QNN gives a performance at the level of over 90%. Training set accuracy was better than 97% in both cases. Additional simulations with more layers may provide somewhat improved accuracy but only if the NN hyperparameters are carefully tuned or optimized [43] . In fact, our simulations with additional layers at this point, without lengthy optimization of hyperparameters, did not provide improved accuracy. For this reason, our simulations for coughing audio data presented next are also based on a single layer QNN.

B. Results Obtained for Coughing Sounds

In this section, we used coughing audio data for COVID-19 detection. As expected, the COVID data sets presented more challenges in terms of segmentation than clean speech. Spectrograms and features were re-computed for each simulation. In addition, we considered different types of quantum noise. The classical RNN achieved a test accuracy of 79.4% and the CNN model achieved 73.0%. Lastly, the QNN provided test accuracy ranging from 74.6% to 78.8% with no noise. Training set accuracy, as expected, was higher than 90% in some cases. Test set results (best case) are shown in Table 2.

TABLE 2. A SUMMARY OF THE RNN, CNN, AND QNN RESULTS. RESULTS FROM QNN SIMULATIONS WITH AND WITHOUT QUANTUM NOISE ARE SHOWN.

Method	Test Accuracy (%)
RNN	79.4
CNN	73.0
QNN - 2 Qubits, no noise	74.6
QNN - 4 Qubits, no noise	78.8
QNN - 2 Qubits with Quantum Noise, Bit Flip circuit, p = 0.01	60
QNN - 4 Qubits with Quantum Noise, Bit Flip circuit, p = 0.01	60.3
Classical VGG-13	78.3

We also present in Table 3 results with quantum noise models incorporated. We started with a simulation of noise with a single-qubit bit flip (Pauli X) error channel model. The simulations with quantum noise for 2 and 4 qubits produced around 60% accuracy. Hence we see a loss of accuracy of more than 14% with the bit flip noise model presented in Table 2. To examine further the effects of quantum noise, we designed another noisy circuit. Depolarization was used on the qubits with different probabilities of occurrence. This error circuit is a generalization of the bit flip and phase flip for the qubit channels. These results (best case) are captured in Table 3.

TABLE 3. A SUMMARY OF QNN WITH DIFFERENT PROBABILITY OF DEPOLARIZATION ERROR SIMULATION RESULTS.

Method, probability	Best Test Accuracy (%)
2 Qubits, p=0.001	71
2 Qubits, p=0.01	73
2 Qubits, p=0.1	71
2 Qubits, p=0.2	65
4 Qubits, p=0.001	73
4 Qubits, p=0.01	75
4 Qubits, p=0.1	67
4 Qubits, p=0.2	65

In general, we observed from the simulations of Table 3, that accuracy was slightly reduced when utilizing a quantum neural network relative to a classical NN. Accuracy decreased by approximately 10-12% when adding quantum noise. As expected, we also observed that the feature extraction process took a considerably longer time to execute than classical methods. In general, feature extraction time increased proportionally with the number of qubits used in the simulation. We note again that training set accuracy was higher than 87% in some cases.

IV. CONCLUSION

In this study, we demonstrated the ability of a quantum neural network to classify audio samples using features extracted from log-mel spectrograms. The quantum circuit was validated first with clean speech data and then examined with COVID-19 coughing audio data. Specifically, we utilized a curated dataset containing samples from the DiCOVA 2021 and COUGHVID datasets. Using an algorithm with a 2 qubit-1 layer quantum circuit placed before an RNN, we achieved a test accuracy of 74.6%. We also performed simulations using an algorithm with a 4 qubit-1 layer quantum circuit. This algorithm achieved a test accuracy of 78.8%. We note that our QNN simulations are currently taking several hours to execute, especially at the QNN training phase. Additionally, simulations were performed with added quantum noise. Both the 2 qubit-1 layer and 4 qubit-1 layer algorithms achieved an accuracy of around 60% with quantum noise added when tested using a single-qubit bit flip error channel. When tested using simulated depolarization error, the accuracies ranged from 65-73% for both the 2-qubit and 4-qubit algorithms. In future work, we will statistically validate these initial best case results using Monte Carlo simulations. We will also optimize our QNNs through careful tuning of hyperparameters and also by using quantum noise error mitigation techniques. We believe that these steps on QNN will provide accuracy in excess of 80% as was the case with the classical VGG-13 shown in Table 2.

ACKNOWLEDGMENT

Portions of this study have been supported by the SenSIP center and the REU award 1659871. Additional support was obtained from the SenSIP NSF I/UCRC project Award 1540040.

REFERENCES

[1] C. Bales et al., "Can Machine Learning Be Used to Recognize and Diagnose Coughs?," 2020 International Conference on e-Health and Bioengineering (EHB), 2020, pp. 1-4..

[2] R. X. A. Pramono, S. Bowyer, and E. Rodriguez-Villegas, "Automatic adventitious respiratory sound analysis: A systematic review," Plos One, vol. 12, no. 5, 2017.

[3] A. Imran et al., "Ai4covid-19: Ai enabled preliminary diagno- sis for covid-19 from cough samples via an app," Informatics in Medicine Unlocked, vol. 20, p. 100378, 2020.

[4] N. Sharma et al., "Coswara–a database of breathing, cough, and voice sounds for covid-19 diagnosis," arXiv preprint arXiv:2005.10548, 2020.

[5] L. Orlandic, T. Teijeiro, and D. Atienza, "The coughvid crowd- sourcing dataset: A corpus for the study of large-scale cough anal- ysis algorithms," arXiv preprint arXiv:2009.11644, 2020.

[6] M. Esposito, S. Rao, V. Narayanaswamy, and A. Spanias, "COVID-19 detection using audio spectral features and machine learning". *IEEE Asilomar Conf. on Circuits, Systems & Comp.,* Monterrey, Oct 2021.

[7] S. Rao, V. Narayanaswamy, M. Esposito, J. Thiagarajan, and A. Spanias, "Deep Learning with hyper-parameter tuning for COVID-19 Cough Detection," In *2021 12th IEEE IISA*, pp. 1-5). IEEE, July 2021

[8] T. F. Quatieri, T. Talkar and J. S. Palmer, "A Framework for Biomarkers of COVID-19 Based on Coordination of Speech-Production Subsystems," in *IEEE Open Journal of Engineering in Medicine and Biology*, vol. 1, pp. 203-206, 2020..

[9] Muguli, A., Pinto, L., R, N., Sharma, N., Krishnan, P., Ghosh, P.K., Kumar, R., Bhat, S., Chetupalli, S.R., Ganapathy, S., Ramoji, S., Nanda, V. (2021) DiCOVA Challenge: Dataset, Task, and Baseline System for COVID-19 Diagnosis Using Acoustics. *Proc. Interspeech* 2021

[10] Vergin, R., O'Shaughnessy, D., and Farhat, A., "Generalized mel frequency cepstral coefficients for large- vocabulary speaker-independent continuous-speech recognition," *IEEE Transactions on speech and audio processing*, 7(5), 525-532, 1999.

[11] A.S.Spanias,"Speech Coding: A tutorial review," *Proceedings of the IEEE*, vol. 82, no. 10, pp. 1541–1582, 1994.

[12] G. Wichern, J. Xue, H. Thornburg, B. Mechtley, and A. Spanias, "Segmentation, indexing, and retrieval for environmental and natural sounds," *IEEE Trans. on ASLP*, vol. 18, no. 3, pp. 688–707, 2010.

[13] P.C. Loizou and A.S. Spanias, "High-performance alphabet recognition," *IEEE Trans. on SAP*, pp. 430– 445, Nov 1996.

[14] M.Huzaifah,"Comparison of time-frequency representations for environmental sound classification using convolutional neural net- works," arXiv preprint arXiv:1706.07156, 2017.

[15] G.S.Uehara, A. Spanias and W. Clark, "Quantum Information Processing Algorithms with Emphasis on Machine Learning," 12th *IEEE IISA*, July 2021, pp. 1-11

[16] *CIRQ : Google quantum AI*, from https://quantumai.google/cirq

[17] A. Cross, "The IBM Q experience and QISKit open-source quantum computing software." In *APS March Meeting*. pp. L58-003. 2018.

[18] J. Pino et al, "Demonstration of the QCCD trapped-ion quantum computer architecture," *arXiv preprint arXiv:2003.01293*.

[19] Cloud computing services: Microsoft Azure, from https://azure.microsoft.com/en-us/

[20] E. A. Sete, W. J. Zeng and C. T. Rigetti, "A functional architecture for scalable quantum computing," *2016 IEEE ICRC,* 2016, pp. 1-6,.

[21] Clerk, A. A., Devoret, M. H., Girvin, S. M., Marquardt, F., & Schoelkopf, R. J. (2010). Introduction to quantum noise, measurement, and amplification. *Reviews of Modern Physics*, *82*(2), 1155.

[22] Chaharlang, Javad et al. "A Novel Quantum Steganography-Steganalysis System for Audio Signals. Multimedia Tools and Applications," 79. 10.1007/s11042-020-08694-z.

[23] J. Qi & J. Tejedor (2021). Classical-to-Quantum Transfer Learning for Spoken Command Recognition Based on Quantum Neural Networks. *arXiv preprint arXiv:2110.08689*.

[24] Bergholm, et al, (2018). Pennylane: Automatic differentiation of hybrid quantum-classical computations. *preprint arXiv:1811.04968*.

[25] Tsang, M., & Caves, C. M. (2010). Coherent quantum-noise cancellation for optomechanical sensors. *Physical review letters*, *105*(12), 123601.

[26] Boulant, N., Emerson, J., Havel, T. F., Cory, D. G., & Furuta, S. (2004). Incoherent noise and quantum information processing. *The Journal of chemical physics*, *121*(7), 2955-2961.

[27] Kalachev, G., Panteleev, P., Zhou, P., & Yung, M. H. (2021). Classical Sampling of Random Quantum Circuits with Bounded Fidelity. *arXiv preprint arXiv:2112.15083*.

[28] B. Fefferman, "*The power of random quantum circuits*," Simons Institute for the Theory of Computing. May 6, 2020.

[29] Bouland, et al. On the complexity and verification of quantum random circuit sampling. Nature Phys 15, 159–163 (2019).

[30] G. Uehara, V. Narayanaswamy, C. Tepedelenlioglu, A. Spanias, *"Quantum Machine Learning for Photovoltaic Topology Optimization," 2022 IEEE 13th IISA*, July 2022.

[31] G. Uehara, S. Rao, M. Dobson, C. Tepedelenlioglu and A. Spanias, "Quantum Neural Network Parameter Estimation for Photovoltaic Fault," *Proc. IEEE IISA 2021*, July 2021

[32] X. Hong and C. Maojun, "Hybrid quantum neural networks model algorithm and simulation," in *5th International Conference on Natural Computation, ICNC 2009*, vol. 1, no. 1, pp. 164–168.

[33] M. A. Metawei, H. Said, M. Taher, H. Eldeib, and S. M. Nassar, "Survey on Hybrid Classical-Quantum Machine Learning Models," in Proceedings of the *2020 IEEE International Conf. on Communications, Computing, Cybersecurity, and Informatics, CCCI 2020*, 2020, pp. 1–6.

[34] Spanias, T.Painter, V. Atti, Audio Signal Processing and Coding, Wiley, 2007.

[35] U. Shanthamallu and A. Spanias, Machine and Deep Learning Algorithms and Applications, Morgan & Claypool Publi, 2021.

[36] K. Jaskie and A. Spanias, Positive Unlabeled Learning, Morgan & Claypool Publishers, AI and Machine Learning, 2022.

[37] U. Shanthamallu, J. J. Thiagarajan, H. Song, A. Spanias, "GrAMME: Semi-Supervised Learning using Multi-layered Graph Attention Models," *IEEE Trans. NNLS,* V. 31, Oct. 2020.

[38] J. Thiagarajan, Narayanaswamy, V., Rajan, D.,Liang, J., Chaudhri, A., Spanias, A., (2021). Designing Counterfactual Generators using Deep Model Inversion. *Neurips* 2021

[39] K. Thopalli, S. Katoch, J. J. Thiagarajan, P. Turaga, and A. Spanias. "Multi-Domain Ensembles for Domain Generalization." In *NeurIPS 2021 Workshop on Distribution Shifts: Connecting Methods and Applications*. 2021.

[40] U. Shanthamallu, Thiagarajan, J. J., and Spanias, A. , "Uncertainty-Matching Graph Neural Networks to Defend Against Poisoning Attacks. In *Proc. of the AAAI,2021.,* Vol. 35, pp. 9524-9532, 2021.

[41] J. Thiagarajan, V. Narayanaswamy, R. Anirudh, P. Bremer, and A. Spanias. "Accurate and Robust Feature Importance Estimation under Distribution Shifts." *Proc. AAAI Conference,* v. 35, pp. 7891-98. 2021.

[42] M. Yarter, G. Uehara, A. Spanias, "Implementation and Analysis of Quantum Homomorphic Encryption," *2022 IEEE 13th IISA,* July 2022.

[43] M. Malu, G. Dasarathy, A. Spanias," Bayesian Optimization in High-Dimensional Spaces: A Brief Survey," *Proc. IEEE 12th IISA 2021*, July 2021.

[44] Medsker, L. R., & Jain, L. C. (2001). Recurrent neural networks. *Design and Applications*, *5*, 64-67.

[45] Graves, A., Mohamed, A. R., & Hinton, G. (2013, May). Speech recognition with deep recurrent neural networks. In *2013 IEEE ICASSP* (pp. 6645-6649). IEEE, 2013.