IEEE ICASSP Paper 1396

2021
TORONTO
Canada
June 6-11, 2021

Decentralizing Feature Extraction with Quantum Convolutional Neural Network for Automatic Speech Recognition

C.-H. Huck Yang
June 2021

Georgia Institute of Technology, GA, USA
Brookhaven National Laboratory, NY, USA
IBM Research, NY, USA and Kore University of Enna, Italy

Our Team

Huck Yang
Georgia Tech

Jun Qi
Georgia Tech

Sam Y.-C. Chen
Brookhaven National Lab

Pin-Yu Chen
IBM Research

Sabato Marco Siniscalchi
Kore University of Enna

Xiaoli Ma
Georgia Tech

Chin-Hui Lee
Georgia Tech

Paper - Decentralizing feature extraction with quantum convolutional neural network for automatic speech recognition, ICASSP 2021

Demo - ICASSP Show and Tell Session (June 9, 2021, from 8:00-10:00, EST)

Outline

Motivation

Quantum Computing and Federated Learning

Quantum Convolution for Speech Processing

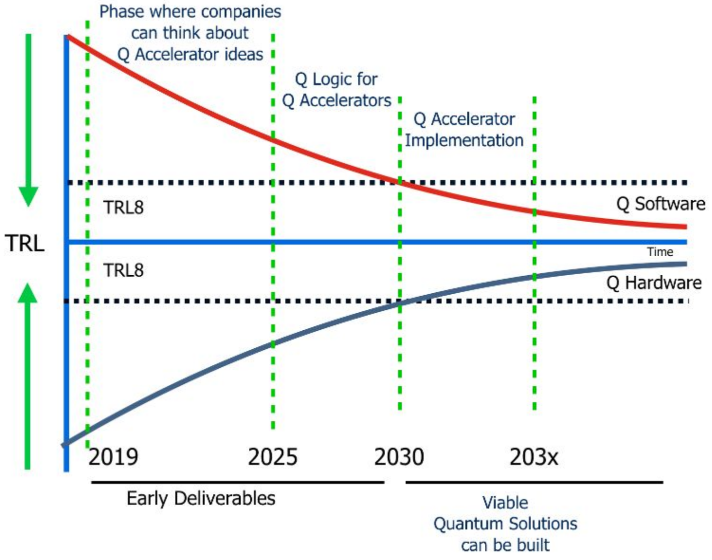
- QCNN-RNN Model
- Performance
- More Analysis

Overview and Future Discussion

- Implementation
- Quantum Hardware

Recent Development of Quantum Hardware - 1

Technology Readiness Level (TRL) on Quantum Computing Investigation [1].



(a) Development time frame

1. Fu, et al. "A heterogeneous quantum computer architecture." Proceedings of the ACM International Conference on Computing Frontiers. 2016.

Recent Development of Quantum Hardware - 2

Academic and Commercial Accessible Quantum Computers

1998	IBM, Oxford, UCB, Stanford, MIT	2 qubits
2000	TU Munich / Los Alamos National Lab	5 qubits / 7 qubits
2006	Perimeter Institute and MIT	12 qubits
2008	D-Wave Systems	28 qubits
2017	IBM, Oxford, UCB, Stanford, MIT / Intel	50 qubits / 49 qubits
2018	Google	72 qubits
2019	Rigetti / IBM	128 qubits / 65 qubits

More recent advances and on enhanced simulation (e.g., TPU Cirq released in 2021)

Definition of Quantum Machine Learning

Pros
(1) Large Representation Space
(2) State, Memory and Complexity

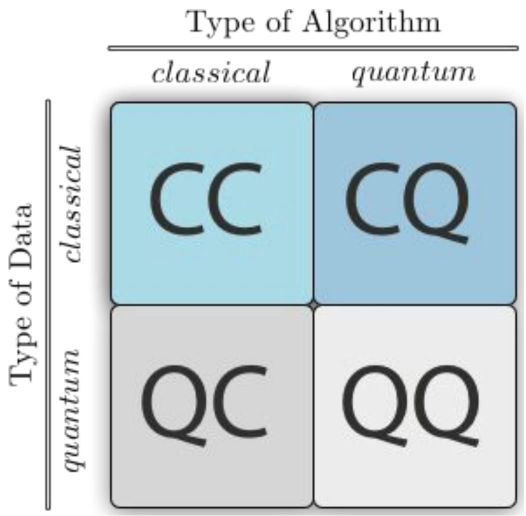


Table. 1 An overview of ML approaches and related key properties

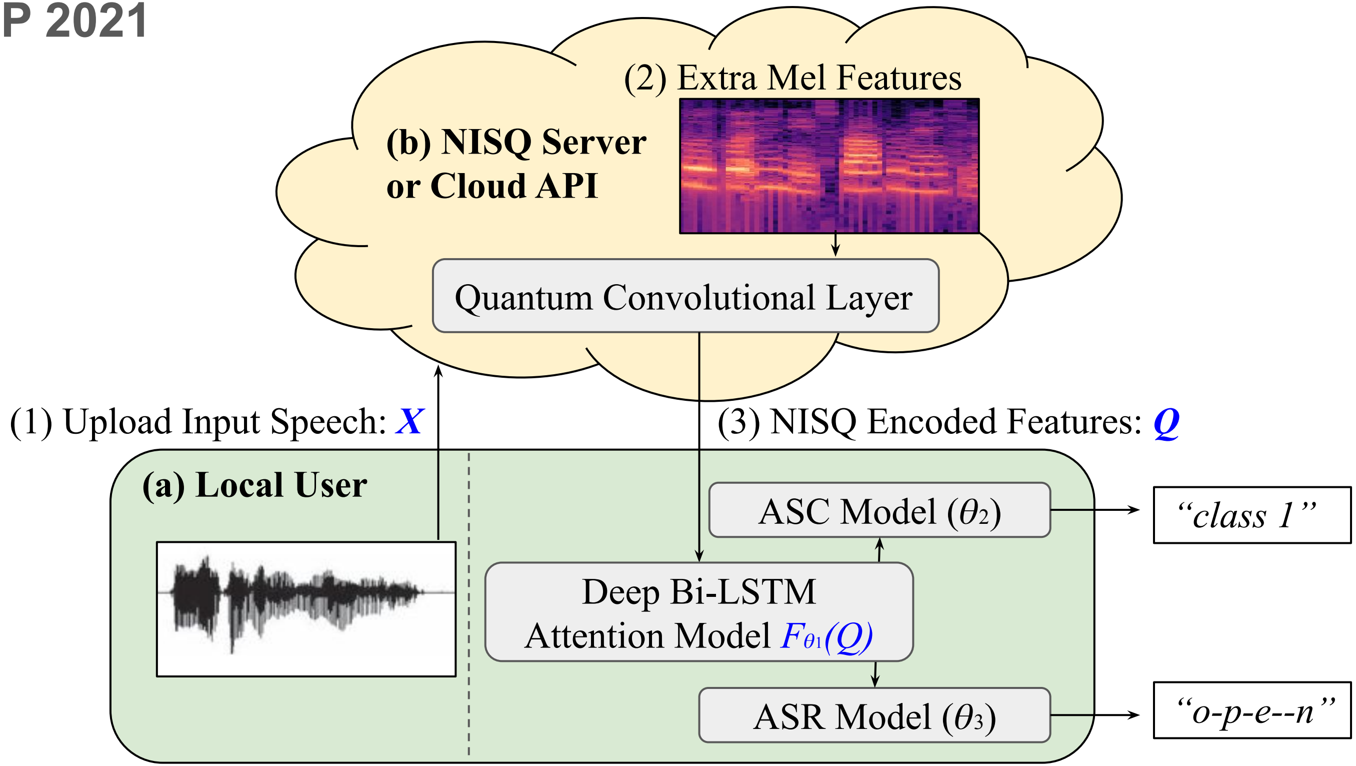
Approach	Input	Learning Model	Output	Properties
Classical	bits	DNN and more.	bits	Easy implementation
Quantum	qubits	VQC and more.	qubits	QA but limited resources
hybrid CQ	bits	VQC + DNN	bits	Accessible QA over VFL

Data Protection Issues such as GDPR and CCPA

Image Source Aïmeur, Esma; Brassard, Gilles; Gambs, Sébastien (2006-06-07). Machine Learning in a Quantum World. Advances in Artificial Intelligence. Lecture Notes in Computer Science

Proposed Architecture in "Vertical Federated Learning"

Figure 1. Quantum Convolution for Speech Command Recognition [5] Yang et al. ICASSP 2021



Speech Processing under Vertical Federated Learning - 1

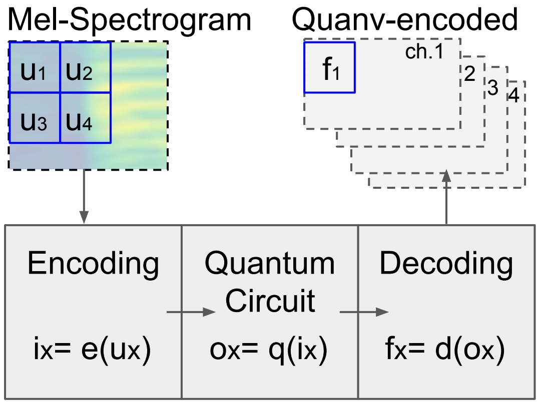
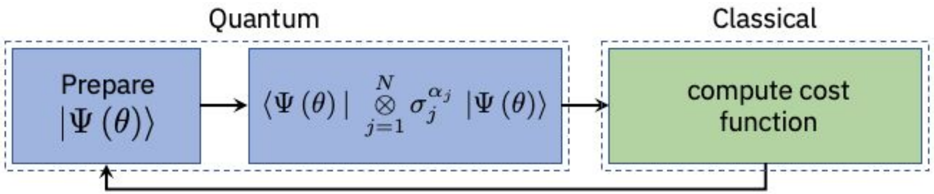


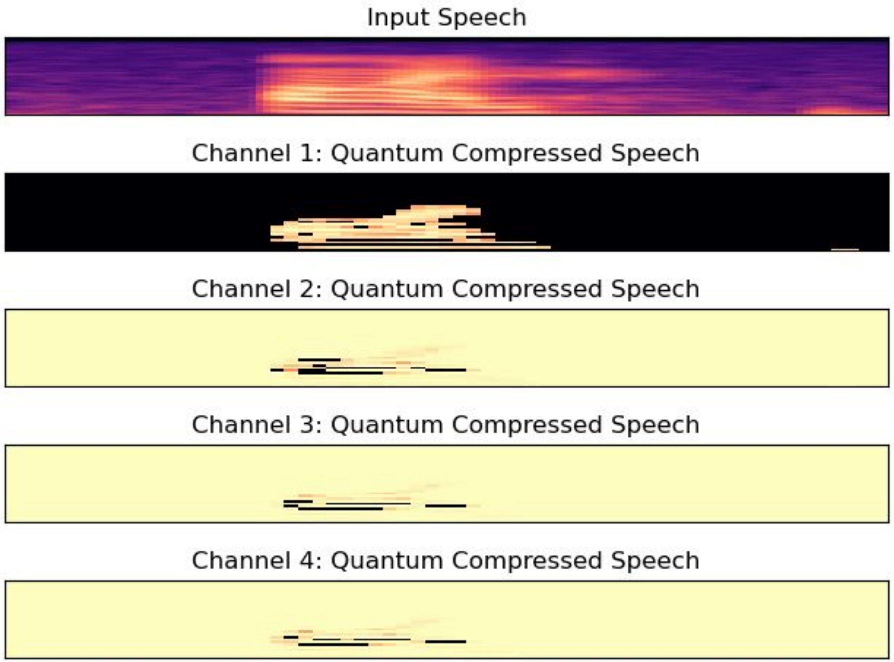
Figure 2. Mel spectrogram features are the input of a quantum circuit layer.



We use Random Parameter in this work; one could also adapt Gradient Propagation to Minimize a Target Loss Function

Image Source: IBM Qiskit Textbook Chap. 3.1

Speech Processing under Vertical Federated Learning - 2



Encoded (Compressed) by Random Quantum Circuit

Speech Processing under Vertical Federated Learning - 3

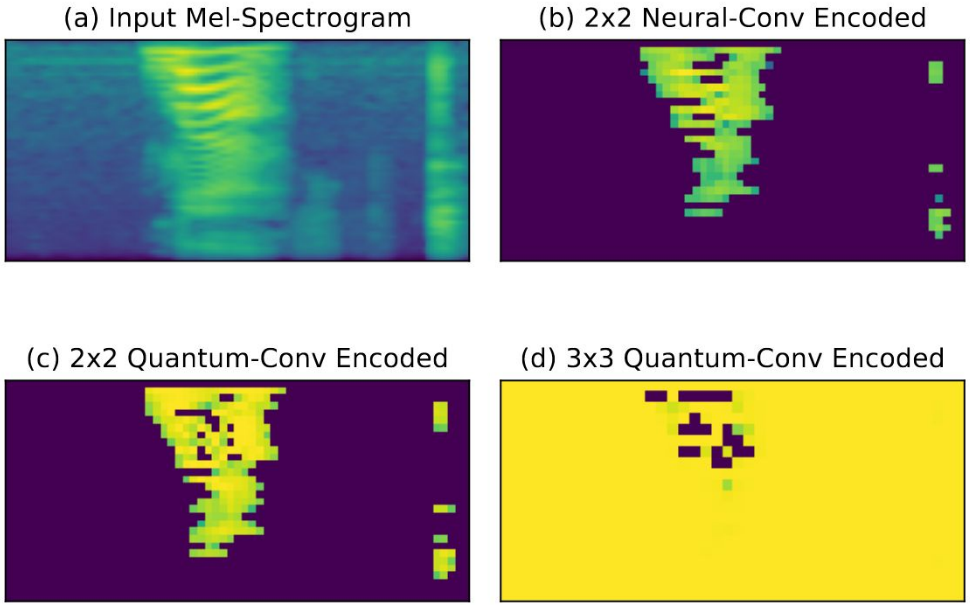


Figure 3. Visualization of the encoded features from different types of convolution layers. The audio transcription is ”yes.”

Proposed "Quantum Convolution and U-Att-RNN Model"

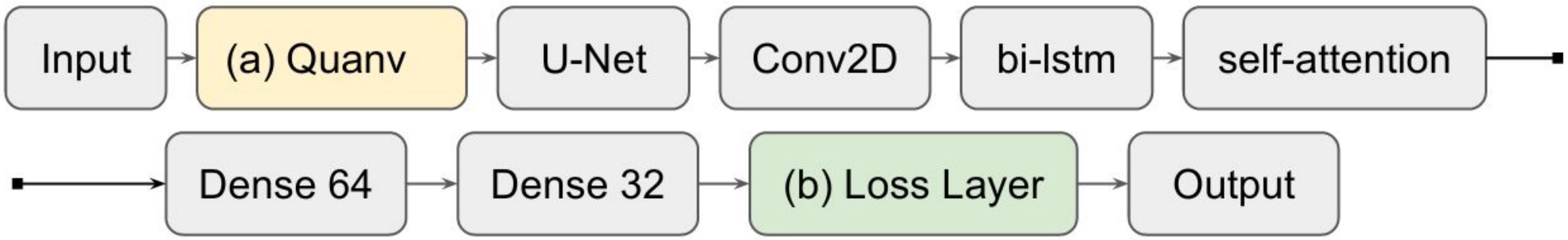


Figure 4. The proposed QCNN ***hybrid classical-quantum*** [2] architecture for speech processing and acoustic modeling tasks.

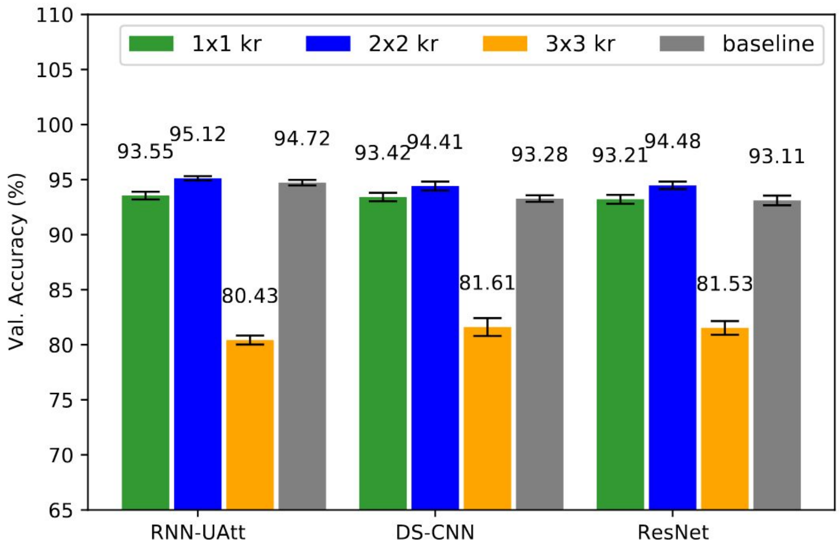
Performance Studies - 1

Table 2. Comparisons of spoken-term recognition on Google Speech Commands dataset.
We provide two baseline CNN-RNN Models from RNN_{At} [3] and its U-Net variant proposed in this work.

Model	Acc. (↑)	Parameters (Memory) (↓)
RNN_{At}	94.21±0.30	170,915 (32-bits)
Conv + RNN_{Att}	94.32±0.26	174,975 (32-bits)
Quanv + RNN_{Att}	94.75±0.17	174,955 (32-bits) + 4 (qubits)
RNN_{UAtt}	94.72±0.23	176,535 (32-bits)
Conv + RNN_{UAtt}	94.74±0.25	180,595 (32-bits)
Quanv + RNN_{UAtt}	**95.12**±0.18	180,575 (32-bits) + 4 (qubits)

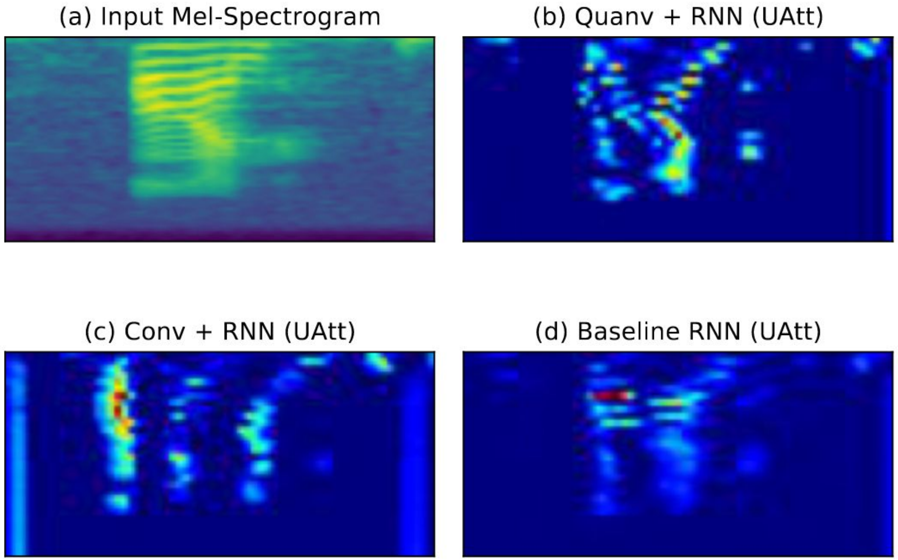
Performance Studies - 2

Figure 5. Performance studies of different quantum kernel size (dubbed kr) with DNN acoustic models for designing QCNN models.



Neural Saliency Studies - 3

Figure 6. Neural saliency results by class activation mapping [5].



Conclusion

- We provide a new architecture for quantum convolution enhanced speech processing.
- Proposed models show competitive results on speech command recognition.
- Preserve Quantum advantages on model parameter protection.
- Implementation is open source.

Future Directions

- Continuous Speech Recognition (CTC Model is provided)
- Differential Privacy Measurement (Also See Our Recent Work)

Feel free to contact us for collaboration. (huckiyang@gatech.edu)
ICASSP Show and Tell Session (June 9, 2021, from 8:00-10:00, EST)

References

1. K. Mitarai, M. Negoro, M. Kitagawa, and K. Fujii, “Quantum circuit learning,”Physical Review A, vol. 98, no. 3, p. 032309, 2018.

2. D. C. de Andrade, S. Leo, M. L. D. S. Viana, and C. Bernkopf,“A neural attention model for speech command recognition,” 2018

3. M. Henderson, S. Shakya, S. Pradhan, and T. Cook, “Quan-volutional neural networks: powering image recognition with quantum circuits,”Quantum Machine Intelligence, vol. 2,no. 1, pp. 1–9, 2020

4. V. Havlıcek et al. Supervised learning with quantum-enhanced feature spaces,”Nature, vol. 567, no. 7747, pp. 209–212, 2019.

5. Yang et al. Decentralizing Feature Extraction with Quantum Convolutional Neural Network for Automatic Speech Recognition. ICASSP 2021

6. Cong, I., Choi, S. & Lukin, M.D. Quantum convolutional neural networks. Nat. Phys. 15, 1273–1278 (2019).

7. P.Warden,“Speech commands: A dataset for limited-vocabulary speech recognition,” Google 2018

Appendix - More Training Time Results