Tools to Be Used

Google CoLab

Github – Tutorial

Pytorch

Qiskit

https://jqub.ece.gmu.edu/categories/QFV/

Tutorial on QuantumFlow+VACSEN: A Visualization System for Quantum Neural Networks on Noisy Quantum Devices

Shaolun Ruan, Yong Wang, Betis Baheri, Qiang Guan, Zhepeng Wang, Weiwen Jiang

SMU | GUANS Lab @ KSU | JQub @ Mason

09/23/2022

Agenda

- **Session 1: Opening <u>(10:00 - 10:15)</u>**
- **Session 2: VACSEN: A Visualization Tool for Noise in Quantum Computing <u>(10:15 - 11:30)</u>**
- **Session 3: QuantumFlow Co-Design Framework <u>(13:00 - 14:00)</u>**
- **Session 4: Quantum Neural Network Compression <u>(14:00 - 14:30)</u>**

Tutorial on QuantumFlow+VACSEN: A Visualization System for Quantum Neural Networks on Noisy Quantum Devices

Session 1: Opening

Weiwen Jiang, Ph.D.
Assistant Professor
Electrical and Computer Engineering
George Mason University
wjiang8@gmu.edu
https://jqub.ece.gmu.edu

Our Goals on Quantum Learning

- For Quantum Neural Network Researchers

  Q: What's a practical way to approaching to quantum advantage?

  A: Algorithm-Compiler-Device Co-Design

- For Quantum Computer Users

  Q: How to make users be aware of the status of quantum devices?

  A: Visualization

- For Everyone

  Q: How to enable everyone can use quantum machine learning?

  A: Quantum learning demonization!

What is Classical AI Democratization & What is the Challenge?

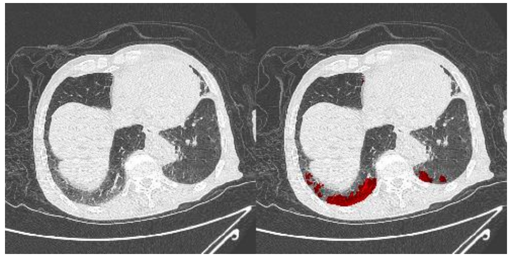
"It's here to collaborate, to augment, to <u>enhance human lives and productivity</u> and <u>make everybody's life better</u>. And related to that, is to **democratize A.I.** in a way that everybody gets benefit. Not just a few, or a selected group." **Fei-Fei Li, 2017**

Medical AI Scenario

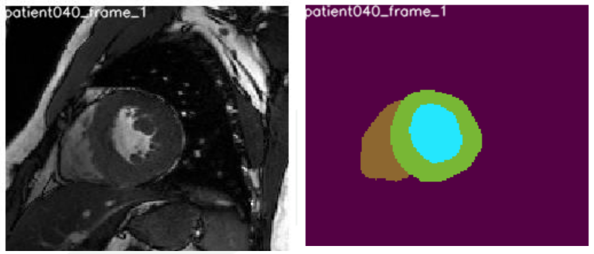
AR/VR in Surgery

Medical Diagnosis

AI Can Perform Medical Tasks



COVID CT Segmentation



Real-Time MRI Segmentation

Let Doctors Design Neural Networks?

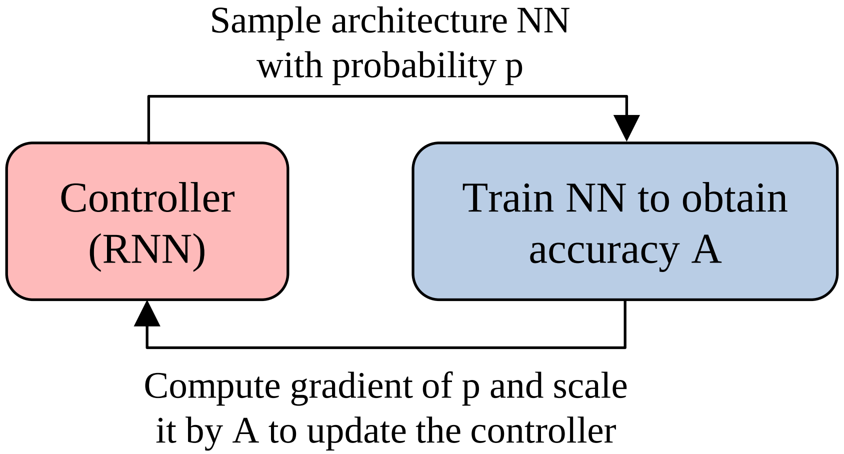
Progress of Classical AI Democratization

Google's Initial Contributions
(Neural Architecture Search)

Given: Dataset

Objective:
- Automated search for NN **(w/o human)**
- Maximize accuracy on the given dataset

Output: A neural network architecture



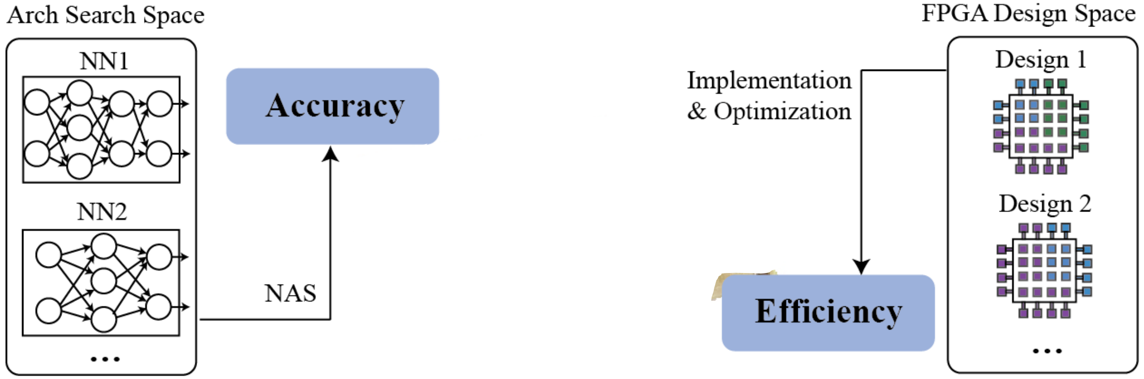
[ref] Zoph, Barret, and Quoc V. Le. "Neural architecture search with reinforcement learning." *ICLR 2017*

Our Contributions
(Network-Accelerator Co-Design)

Given: (1) Dataset; (2) Target hardware, e.g., FPGA.

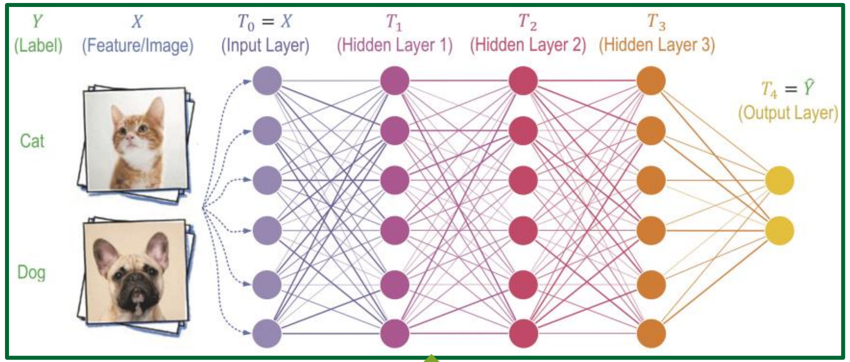
Objective:
- Automated search for NN and HW design
- Maximize accuracy on the given dataset
- Maximize hardware efficiency

Output: A pair of neural network and hardware design

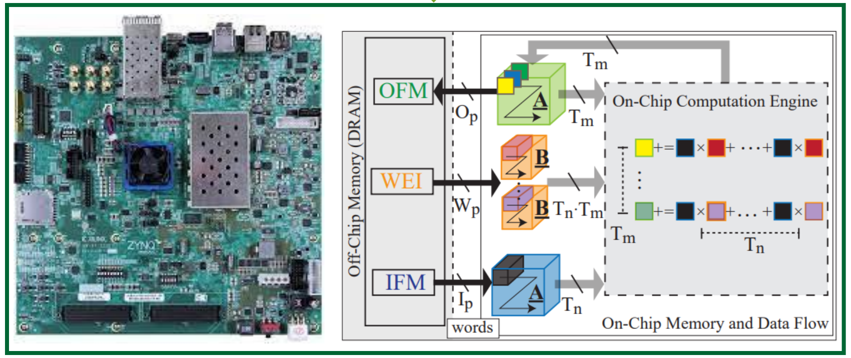


[ref] Jiang, Weiwen, et al. "Accuracy vs. efficiency: Achieving both through fpga-implementation aware neural architecture search." *DAC 2019*. (BEST PAPER NOMINATION)
[ref] Jiang, Weiwen, et al. "Hardware/software co-exploration of neural architectures", TCAD 2020 (BEST PAPER AWARD)

Co-Design Stack of Neural “Architectures”



Co-Design

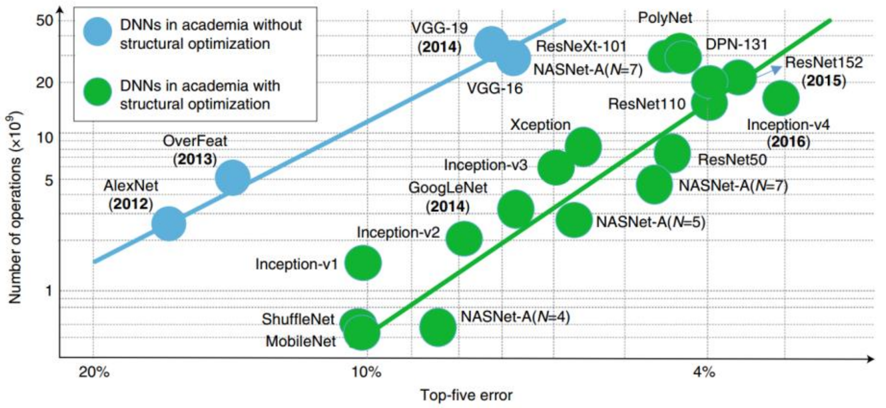


- What is the best **Neural Network Architecture** for FPGAs
- Model optimization (pruning and quantization)?
- Library

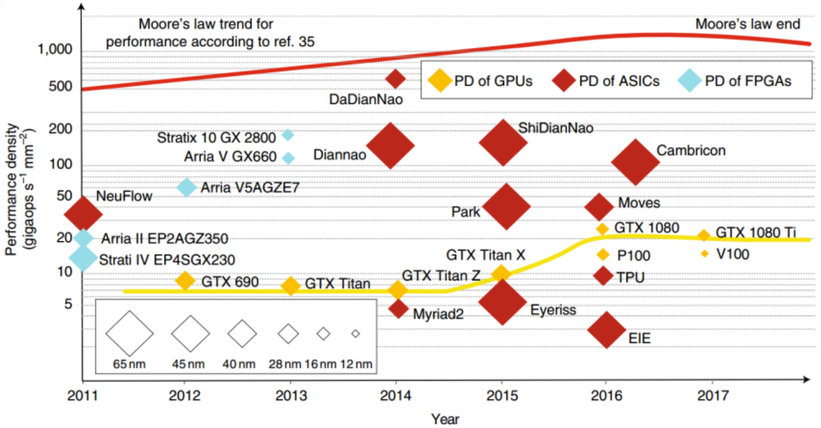
Co-Design Framework (e.g., Our FNAS)		
	Network exploration	NAS (Google)
	Network compression	Deep Comp (Stanford)
	Programming library	DNNBuilder (UIUC)
	Hardware accelerator	DNN on FPGA (UCLA)

- Mapping and scheduling?
- What is the best **FPGA Architecture** for neural networks

Bottlenecks in Classical Computing



Deep neural network grows exponentially



Perf. of classical computing stops increasing

Medical AI Scenario: (Input size exponentially grows from Radiology to Pathology Imaging)

Radiology Imaging

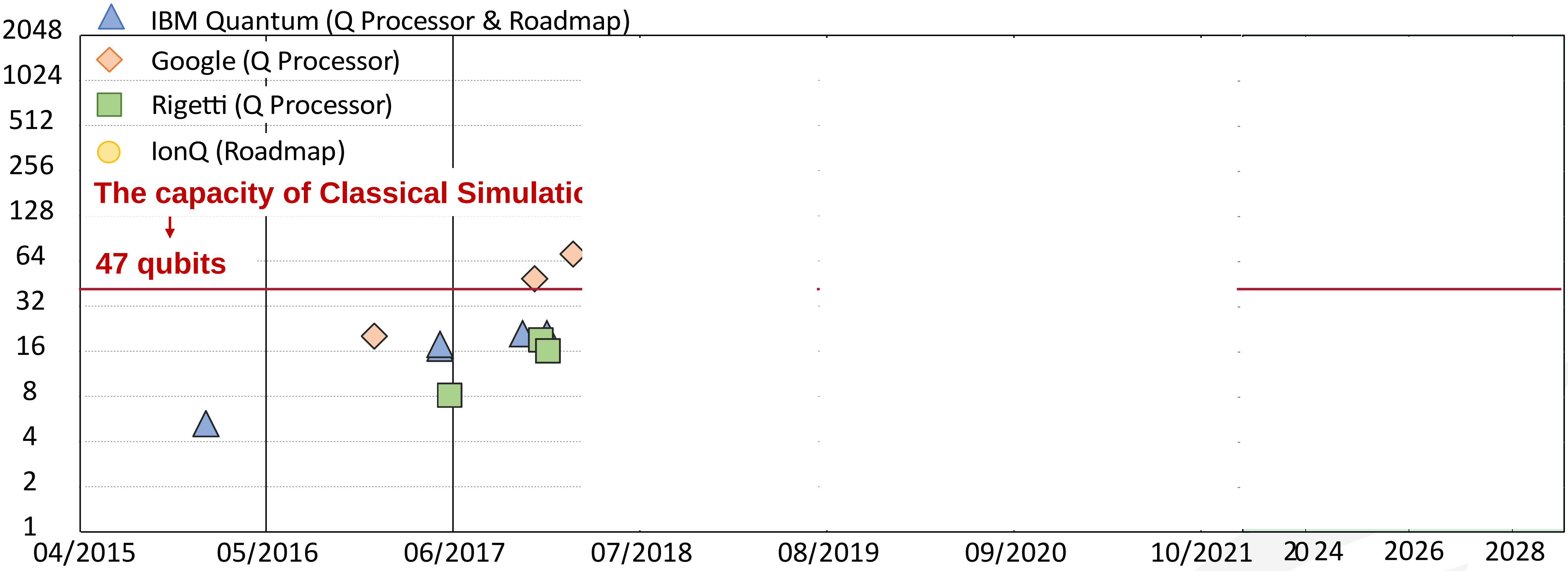
Radiology Modality	Avg. Size (MB)
CT Scan	153.4
MRI	98.6
X-ray angiography	157.5
Ultrasound	69.2
Breast imaging	38.8

Pathology Imaging

Biopsy Type	Compressed Size(MB)/Study	Original Size (GB)
Dermatopathology	1,392 (20x compression)	**27**
Head and neck	1,965 (20x compression)	**38**
Hematopathology	40,300 (40x compression)	**1574**
Neuropathology	1,872 (20x compression)	**37**
Thoracic pathology	3,240 (20x compression)	**63**

[ref] Lauro, Gonzalo Romero, et al. "Digital pathology consultations—a new era in digital imaging, challenges and practical applications." *Journal of digital imaging* 26.4 (2013).

Impossible in Classical But Possible in Quantum Computing



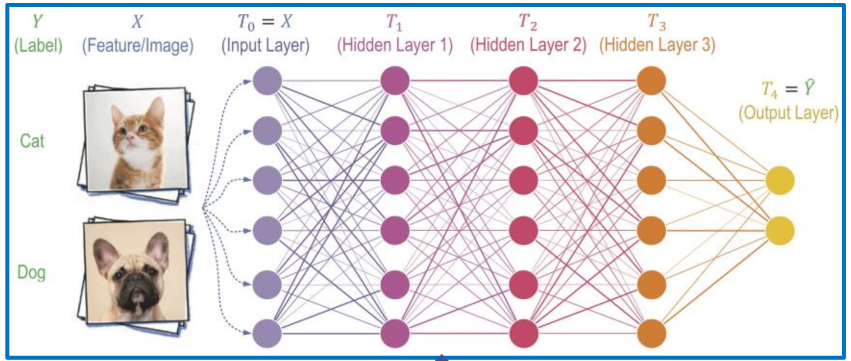
The maximum qubits that supercomputers can simulate for arbitrary circuits is less than 47 qubits.

(1) Summit w/ 2.8 PB memory for **47 qubits**; (2) Sierra w/ 1.38 PB memory for **46 qubits**;

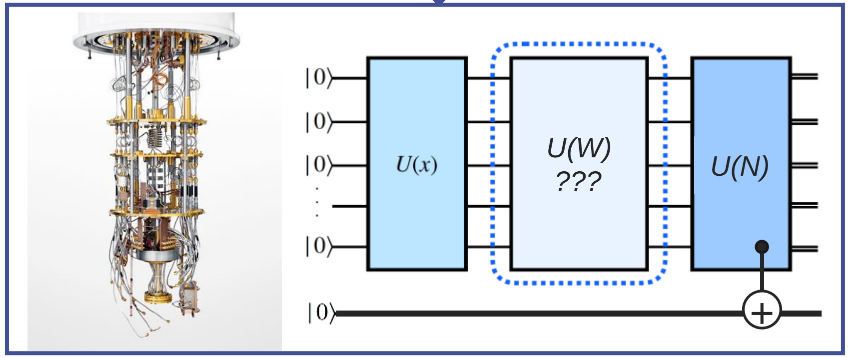
(3) Sunway TaihuLight w/ 1.31 PB memory for **46 qubits**; (4) Theta w/ 0.8 PB memory for **45 qubits**.

[ref] Wu, Xin-Chuan, et al. "Full-state quantum circuit simulation by using data compression." Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis. 2019.

Co-Design of Neural Networks and Quantum Circuit



Co-Design

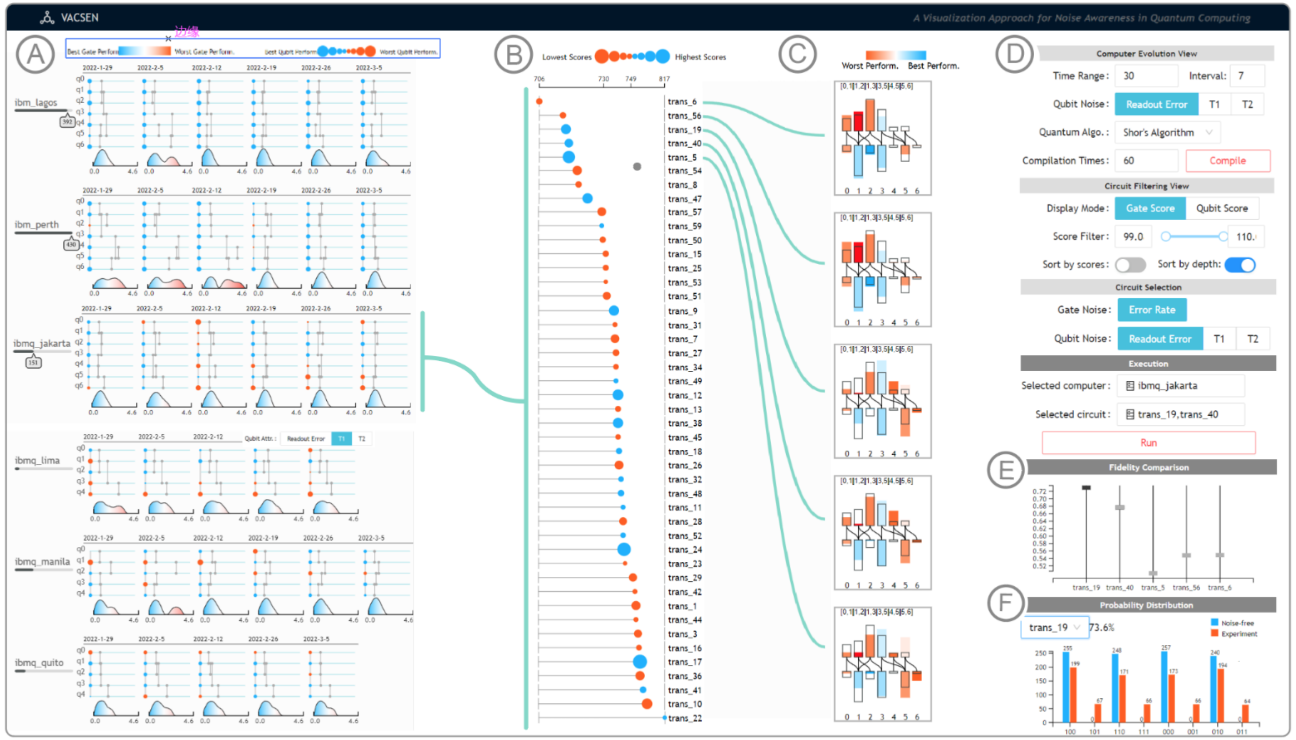


- **What is the best Neural Network Architecture for QC?**
- **Can we compress the quantum neural network?**
- **Library**

Co-Design Framework QuantumFlow	Network exploration	**QF-Mixer**
	Network compression	**CompVQC**
	Programming library	**QFNN**
	Device-aware design	**QF-RobustNN**

- **……**
- **What is the best QC design for neural networks?**

Session 2: VACSEN: A Visualization Tool for Noise in Quantum Computing



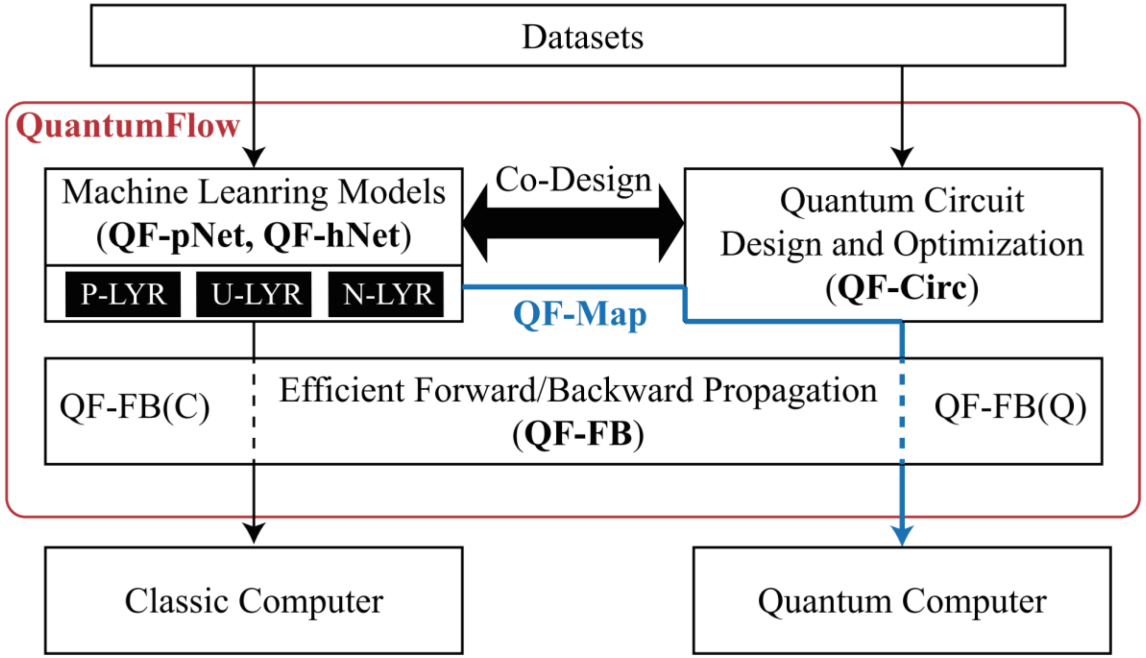
VIS 2022

the premier forum for advances in visualization and visual analytics

October 16, 2022

VACSEN introduces a novel visualization technique to achieve noise-aware quantum computing, detailed comparison on the filtered compiled circuit view, and user-friendly interaction to achieve better fidelity.

Session 3: QuantumFlow Co-Design Framework

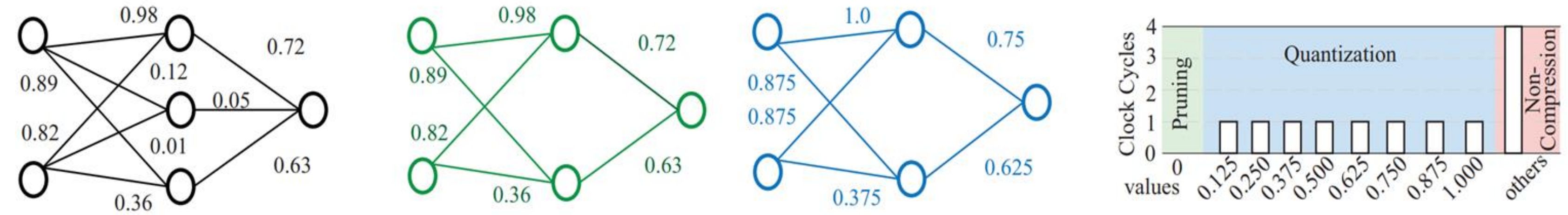


https://www.nature.com/articles/s41467-020-20729-5
https://github.com/JQub/QuantumFlow_Tutorial

- Correctly implement binary neuron on quantum computers.
- Reduce complexity from O(n) in classical computers to O(polylog(n)) in quantum computers.
- On MNIST, achieve same accuracy with a cost reduction of 10.85× over classical computers.

Session 4: Quantum Neural Network Compression

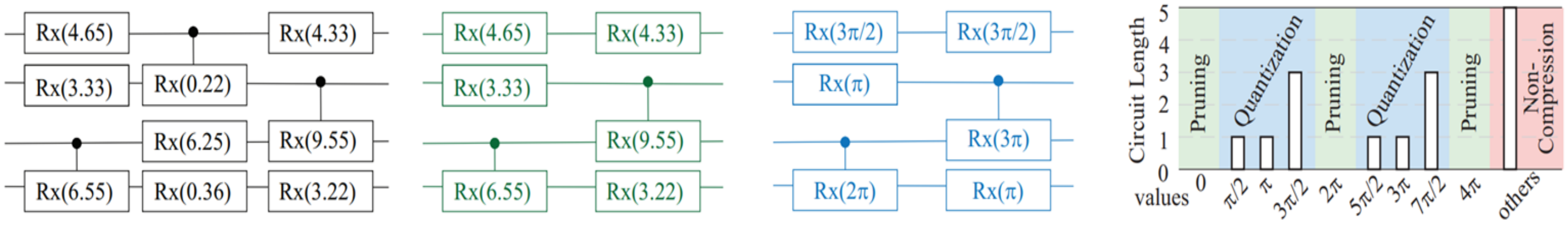
- Pruning and Quantization in Classical ML



(a) Non-Compression Classical NN (b) Classical NN with Pruning (c) Pruned NN with Quantization (d) Cost of Different Levels in Classical NN

IEEE/ACM 2022 INTERNATIONAL CONFERENCE ON COMPUTER-AIDED DESIGN ICCAD 41st Edition

- Pruning and Quantization in Quantum ML



(e) Non-Compression QNN (f) QNN with Pruning (g) Pruned QNN with Quantization (h) Cost of Different Levels in RX Gate in QNN

October 30, 2022

Reduction on the compiled circuit length for more than 2X with <1% accuracy loss.

electronics
An open access journal by MDPI

IMPACT FACTOR 2.690

CITESCORE 3.7

Electronics (ISSN 2079-9292) is an international, peer-reviewed, open access journal on the science of electronics and its applications.

Editor-in-Chief
Prof. Dr. Flavio Canavero
Politecnico di Torino, Italy

First decision to author **16.6** days
Median Submission to Publication **35** days

Semi-Monthly Released

No Copyright Constraints

Electronics 2022 Best Paper Award
Electronics 2022 Young Investigator Award

More information can be found at https://www.mdpi.com/journal/electronics/awards

- Email: electronics@mdpi.com
- www.mdpi.com/journal/electronics
- Twitter: @ElectronicsMDPI

Topics

- Microelectronics
- Optoelectronics
- Power Electronics
- Bioelectronics
- Microwave and Wireless Communications
- Computer Science & Engineering
- Networks
- Systems & Control Engineering
- Circuit and Signal Processing
- Semiconductor Devices
- Artificial Intelligence
- Electrical and Autonomous Vehicles
- Quantum Electronics
- Artificial Intelligence Circuits and Systems (AICAS)
- Industrial Electronics
- Flexible Electronics
- Electronic Multimedia
- Electronic Materials

electronics
An open access journal by MDPI

IMPACT FACTOR 2.690

CITESCORE 3.7

Special Issue:

Quantum Machine Learning: Theory, Methods and Applications

Guest Editors:

Dr. Weiwen Jiang
George Mason University, Fairfax, VA 22030, USA

Dr. Ying Mao
Fordham University, New York, NY 10458, USA

Dr. Samuel Yen-Chi Chen
Computational Science Initiative, Brookhaven National Laboratory, New York, NY 11973-5000, USA

Deadline for manuscript submissions:
20 November 2022

Topics are welcome to contribute:

- Quantum machine learning
- Quantum neural network
- Quantum supervised learning
- Quantum unsupervised learning
- Quantum reinforcement learning
- Quantum learning theory
- Variational quantum circuits
- Noisy intermediate-scale quantum devices (NISQ)

https://www.mdpi.com/journal/electronics/special_issues/quantum_machine_learning