Quantum Transfer Learning for Wi-Fi Sensing

Toshiaki Koike-Akino
Perry Wang
Ye Wang

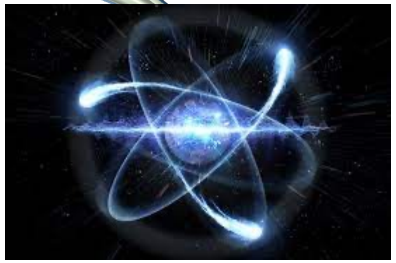
May 16, 2022

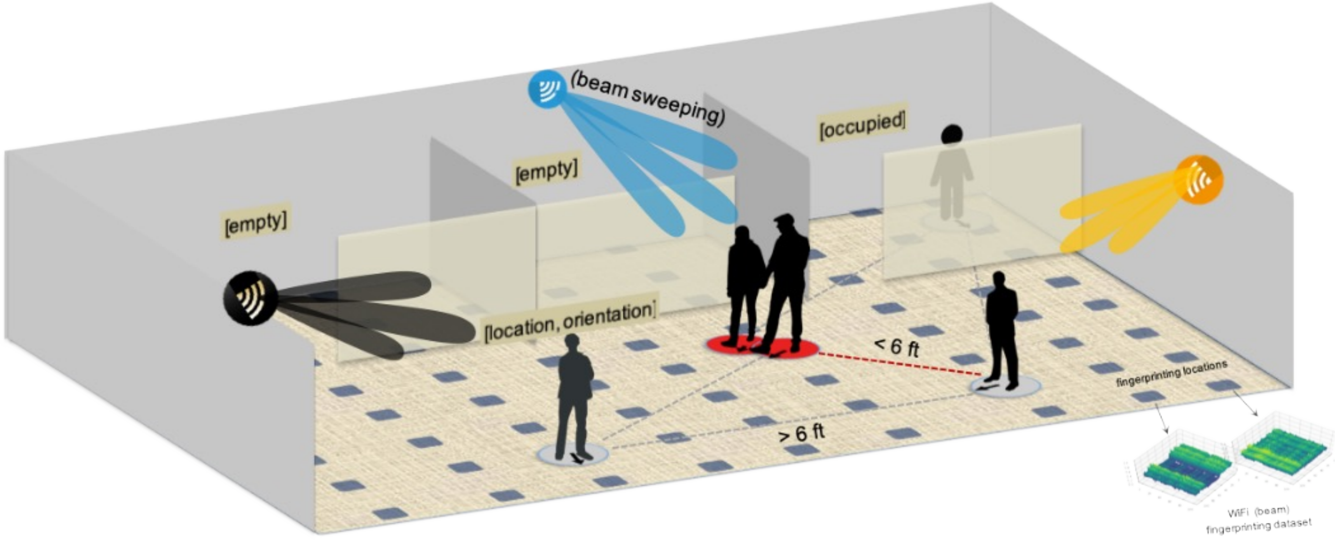
MITSUBISHI ELECTRIC RESEARCH LABORATORIES (MERL)
Cambridge, Massachusetts, USA
http://www.merl.com

Outline

- Trends of Machine Learning (ML)
- Quantum Machine Learning (QML)
- WiFi Sensing for Indoor Monitoring
  - Beam SNR measurement
  - Human pose monitoring
  - QML vs. DNN
  - Transfer learning performance
- Summary

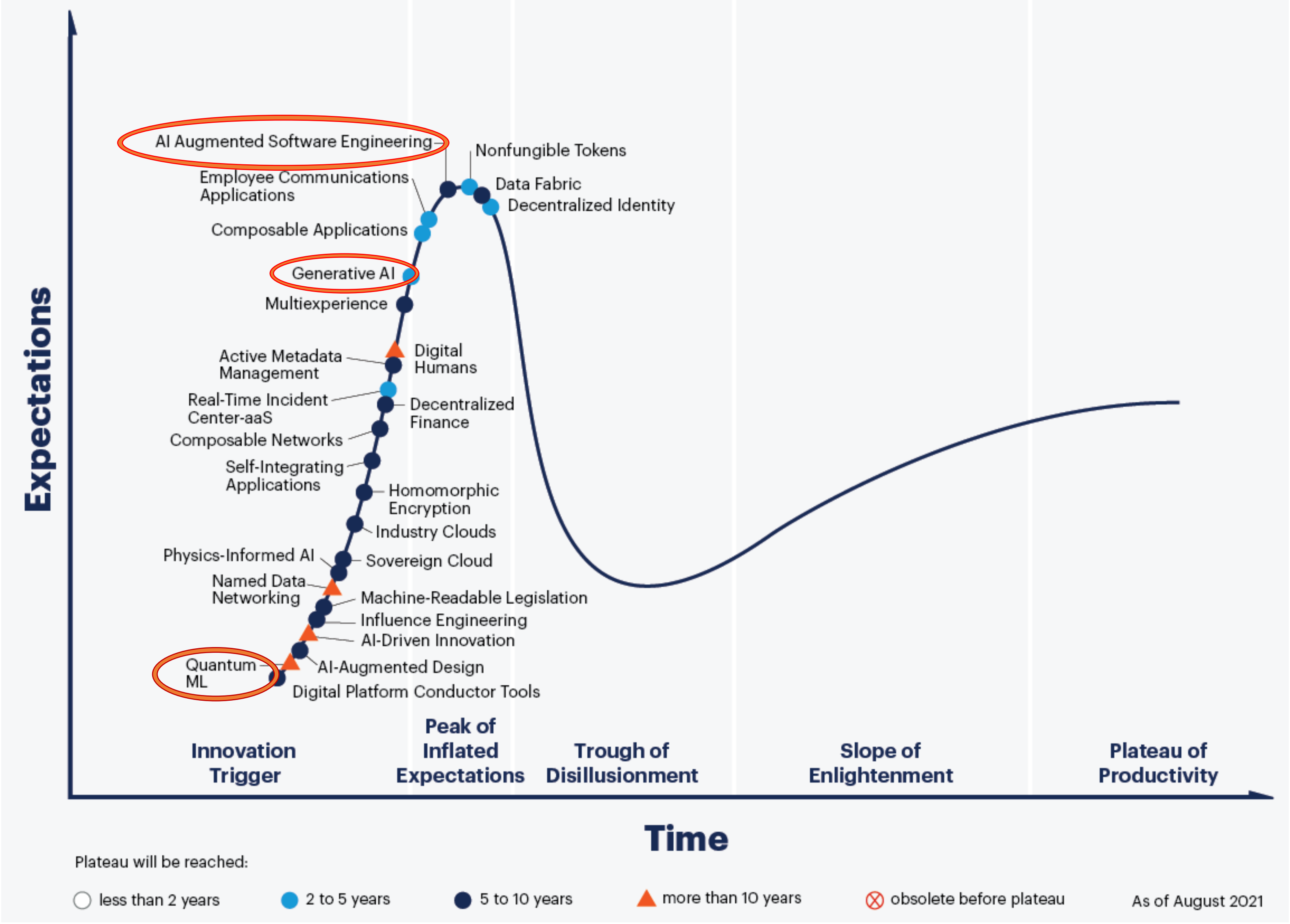
AI x Quantum





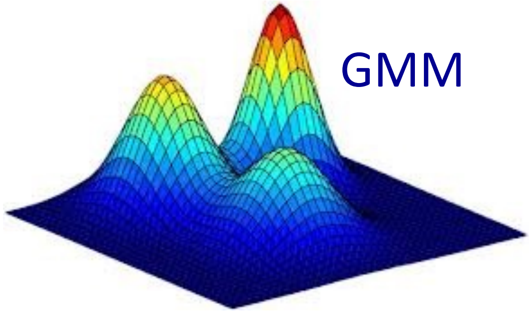
Emerging Technologies

- Gartnar’s Hype Cycle for Emerging Technologies (2021 August):
  - **AI**
  - **Generative AI**
  - **QML**



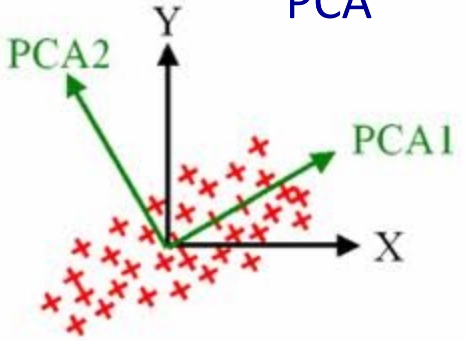
Artificial Intelligence (AI)

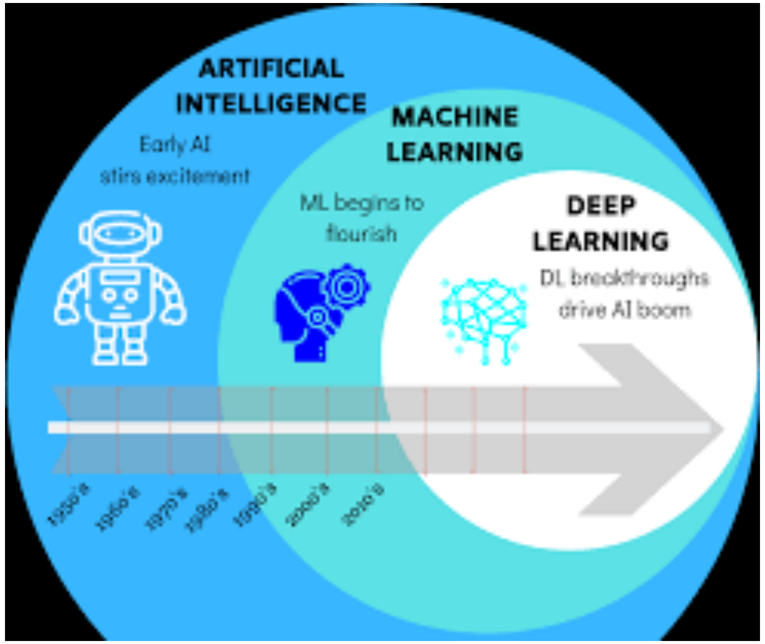
- K-means
- Gaussian mixture model (GMM)
- Principal component analysis (PCA)
- Independent component analysis (ICA)
- Logistic regression (LR)
- **Support vector machine (SVM)**
- Self-organizing map (SOM)
- Hidden Markov model (HMM)
- Artificial neural networks (ANN)
- **Deep learning (DL)**
- **QML …**

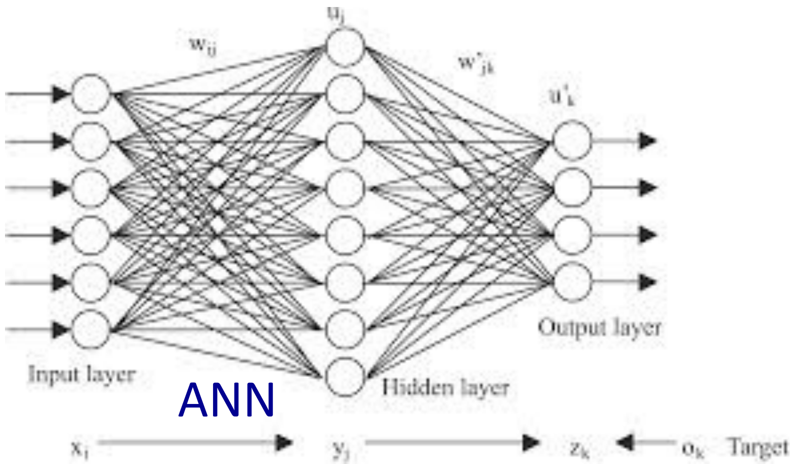


GMM

PCA

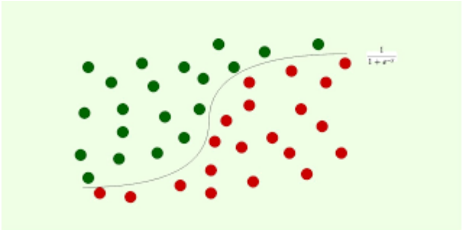


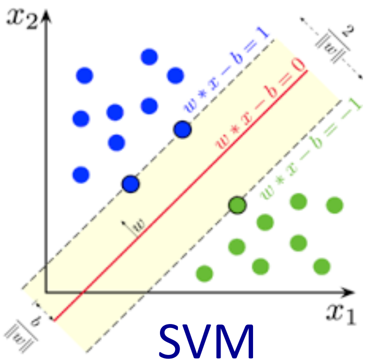




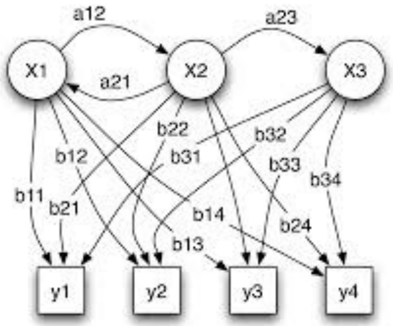
ANN

LR





SVM



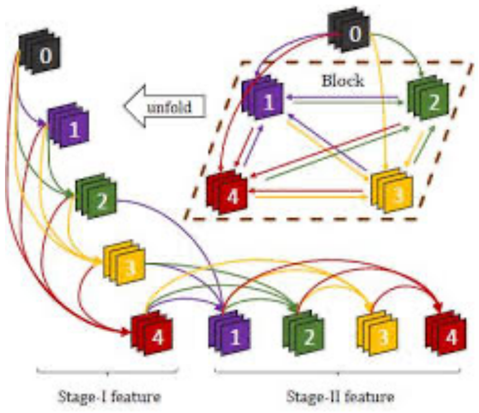
HMM

Deep Learning (DL) for Artificial Intelligence (AI)

- Deep learning = fancy name of multi-layer perceptron neural networks.
  - 2006 Hinton: Many layers, layer-wise pre-training, massive data sets
- Key enabling driver:
  - *Hardware* evolution: graphic processing units (**GPUs**), tensor processing units (TPUs), …
  - *Software* evolution: free libraries (**PyTorch, TensorFlow, …**)

DNN Architectures

- Residual Multi-Layer Perceptron (ResMLP)
- Residual Convolutional Neural Network (ResCNN)
- Bidirectional Long Short-Term Memory (BiLSTM)
- Transformer, U-net
- Implicit ODE, CVX, …



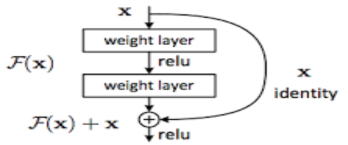
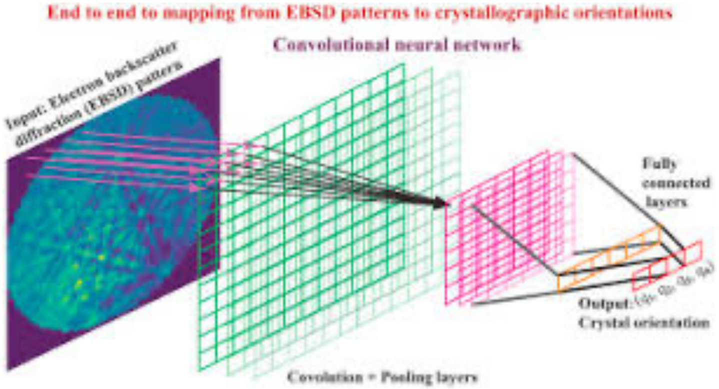
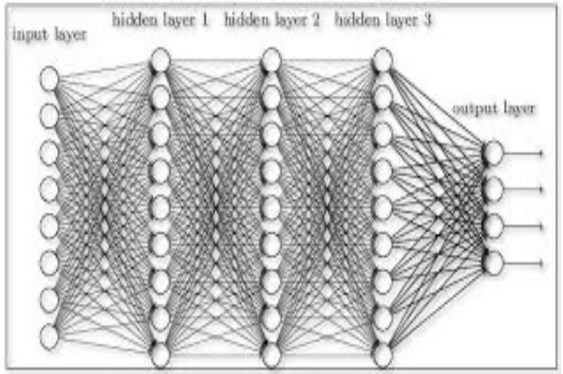
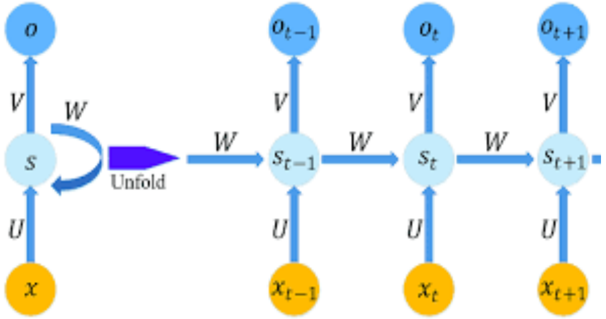
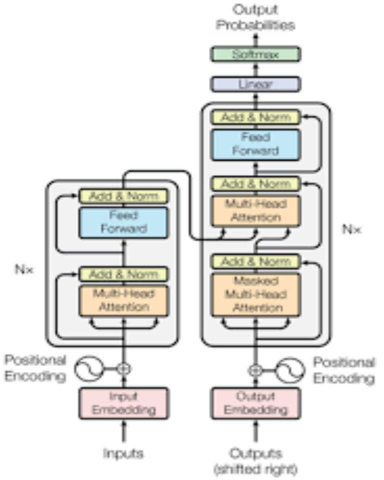


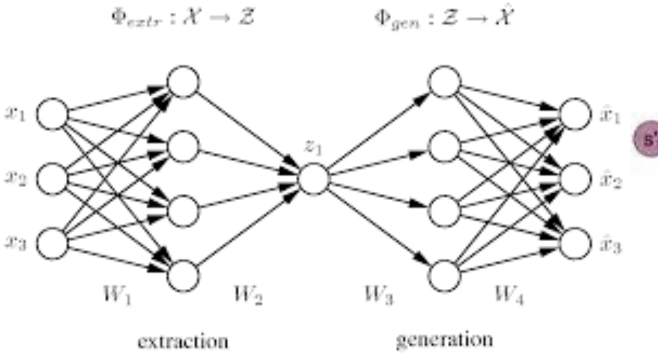
Figure 2. Residual learning: a building block.

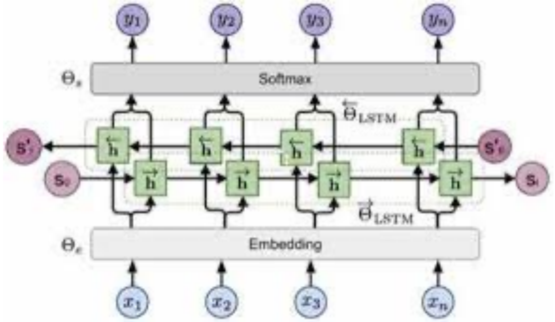


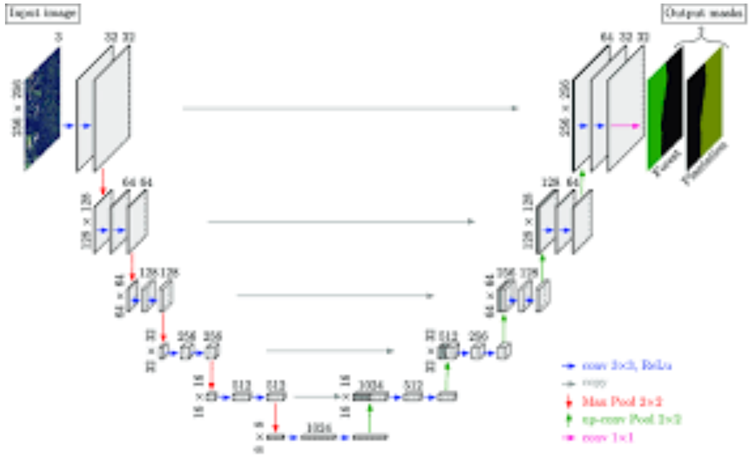


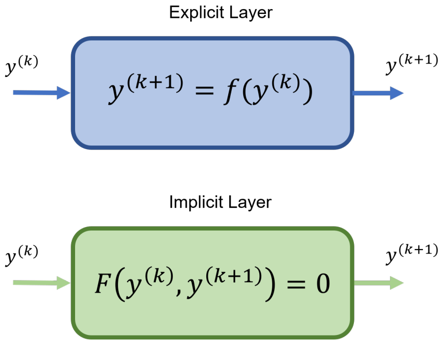






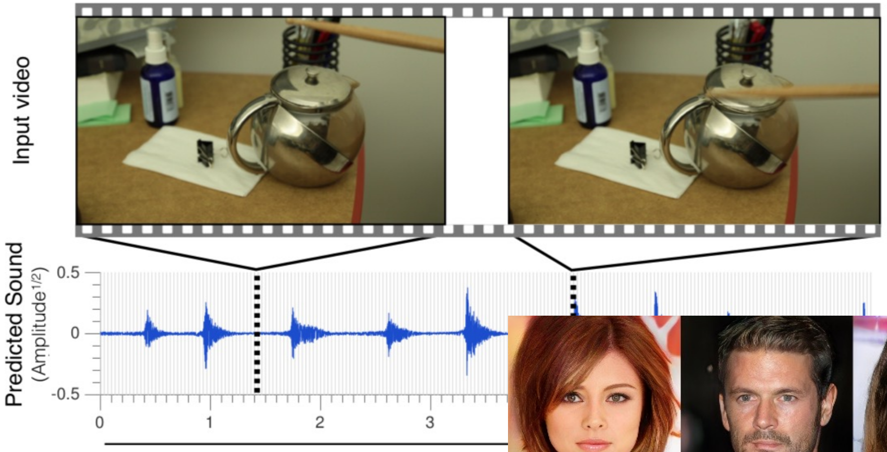


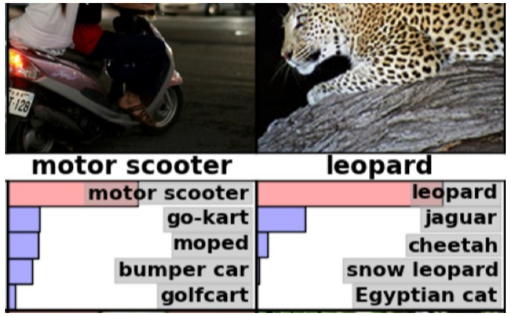




AI Success in Media (Audio & Visual) Signal Processing

- Denoising, segmentation, classification, translation, dialog, recognition, decomposition, generation, super-resolution, …

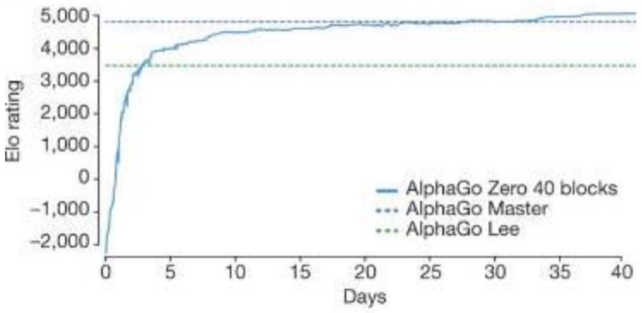


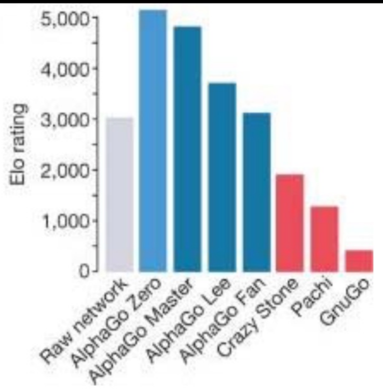



"man in black shirt is playing guitar."

AI Surpassed Human-Level Performance

- For some applications like gaming

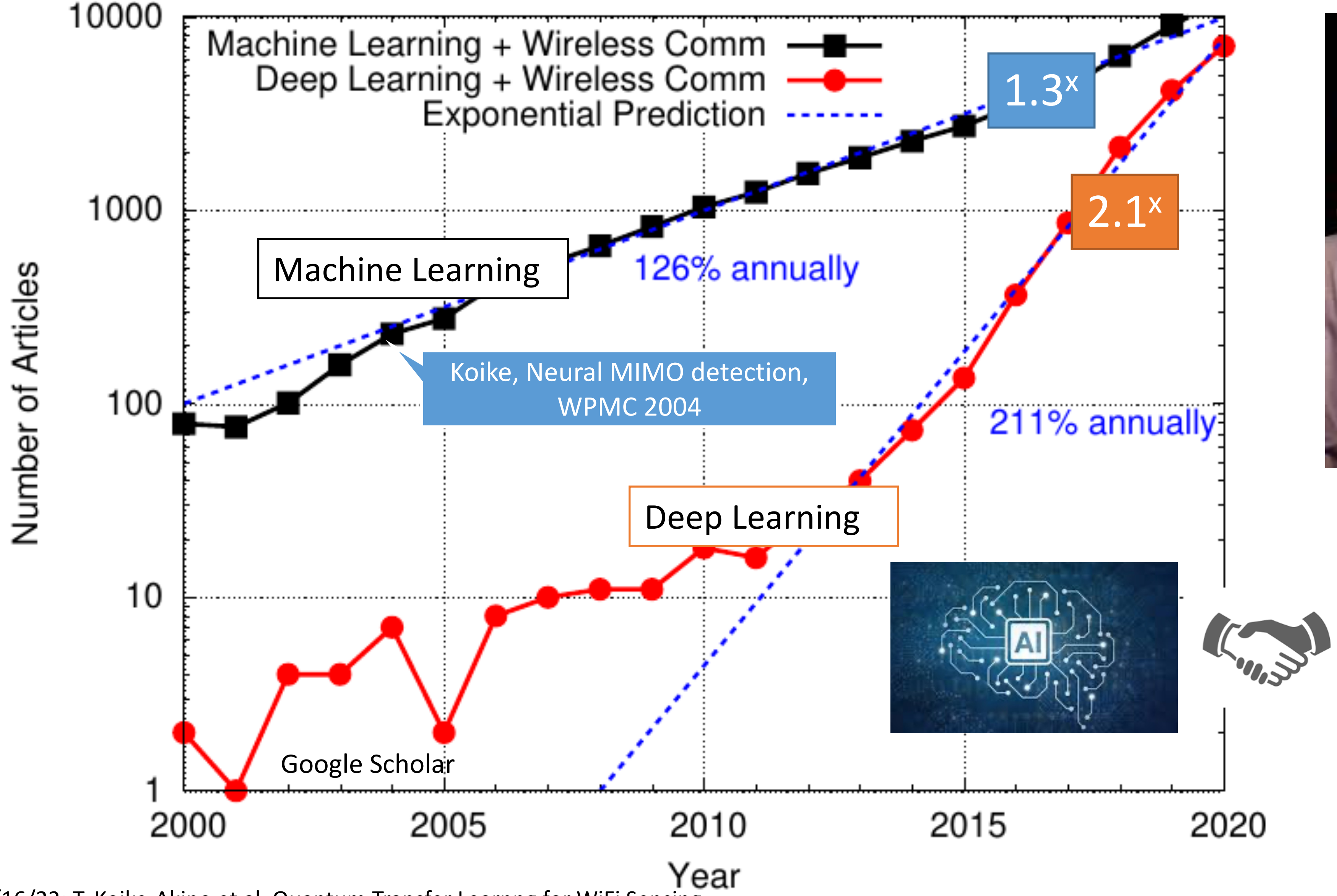


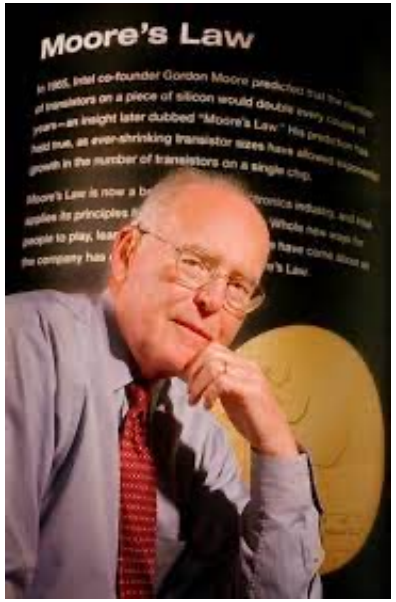


AI for Wireless Communications

Moore’s Law: Exponential Growth in Applications

- Hit count of articles per year in Google Scholar; ***Wireless Communication***





Deep Learning Crisis for Sustainable Growth

- Escalating power consumption of DNN training
  - [Strubell et al. Energy and policy considerations for deep learning in NLP. 2019]
  - 1-big DNN training with network architecture search (NAS) on GPUs requires **5-fold** higher carbon emission of single car lifetime!
- New computing modality alternative to CPU/GPU/TPU is desired
  - **Natural computing: Quantum computing**, DNA computing, etc.

Consumption	CO_2e (lbs)
Air travel, 1 passenger, NY↔SF	1984
Human life, avg, 1 year	11,023
American life, avg, 1 year	36,156
Car, avg incl. fuel, 1 lifetime	126,000
Training one model (GPU)	
NLP pipeline (parsing, SRL)	39
w/ tuning & experimentation	78,468
Transformer (big)	192
w/ neural architecture search	626,155

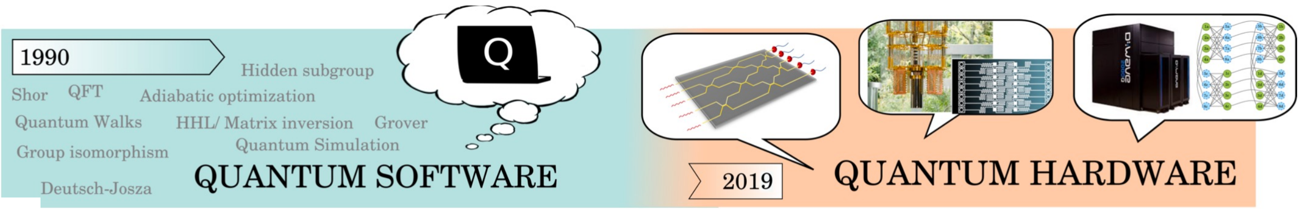
Table 1: Estimated CO_2 emissions from training common NLP models, compared to familiar consumption.[1]

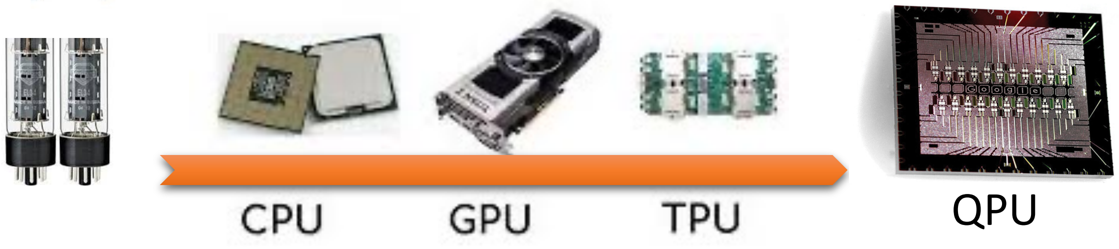
Model	Hardware	Power (W)	Hours	kWh·PUE	CO_2e	Cloud compute cost
$\text{Transformer}_{base}$	P100x8	1415.78	12	27	26	\$41–\$140
Transformer_{big}	P100x8	1515.43	84	201	192	\$289–\$981
ELMo	P100x3	517.66	336	275	262	\$433–\$1472
BERT_{base}	V100x64	12,041.51	79	1507	1438	\$3751–\$12,571
BERT_{base}	TPUv2x16	—	96	—	—	\$2074–\$6912
NAS	P100x8	1515.43	274,120	656,347	626,155	\$942,973–\$3,201,722
NAS	TPUv2x1	—	32,623	—	—	\$44,055–\$146,848
GPT-2	TPUv3x32	—	168	—	—	\$12,902–\$43,008

Table 3: Estimated cost of training a model in terms of CO_2 emissions (lbs) and cloud compute cost (USD).[7] Power and carbon footprint are omitted for TPUs due to lack of public information on power draw for this hardware.

Quantum Computing

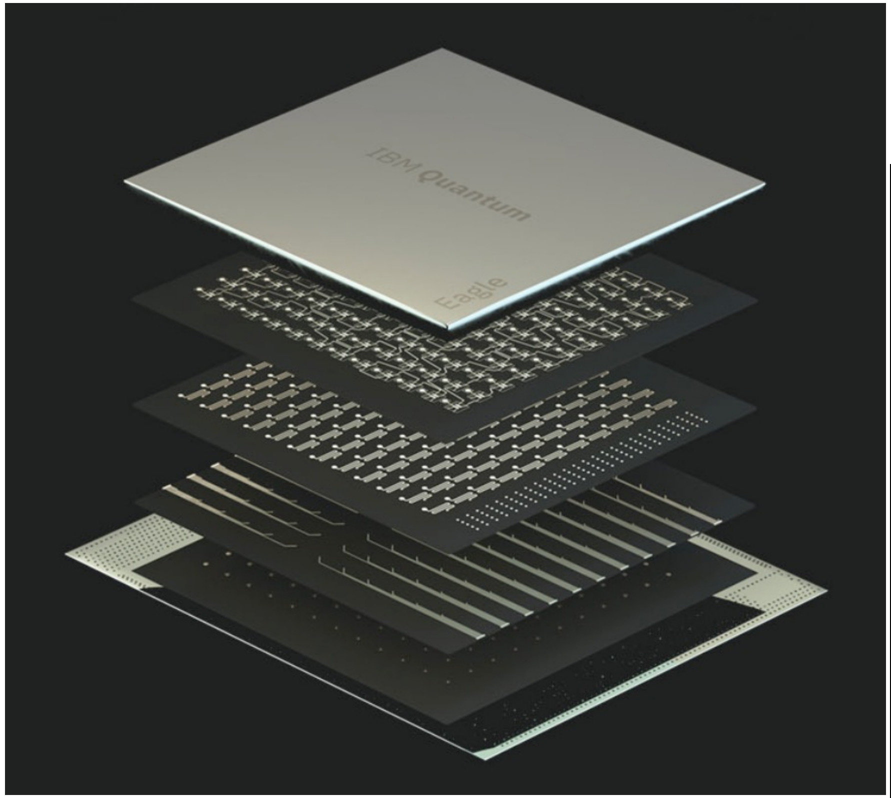
- Morgan Stanley: Quantum tech. can drive **4th industrial revolution**
- Escalating government funds: National Quantum Initiative **$1.2B**
- Quantum processing units (QPU) venders: **IBM**, **Google**, **Microsoft**, **Honeywell**, Intel, Nokia, AirBus, IONQ, rigetti, …
- Quantum cloud services: IBMQ, Amazon Bracket, Microsoft Azure, …
- Free libraries to evaluate quantum computing on realistic simulators or real devices



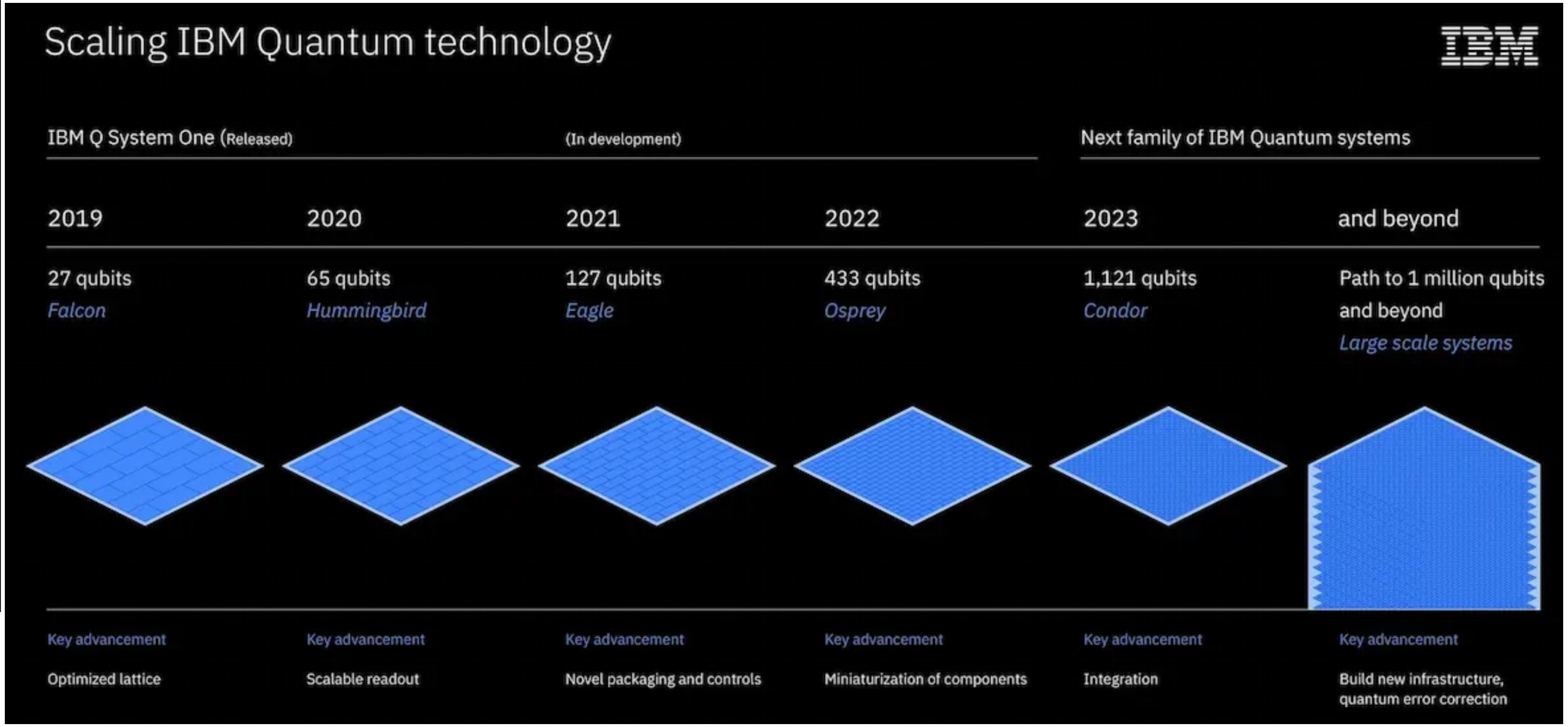


Evolution of Quantum Processing Unit (QPU)

- QPU development has been advancing rapidly to allow many qubits
  - IBM released **127-qubit** QPUs in Nov. 2021
  - IBM plans to release **1121-qubit** QPUs by 2023



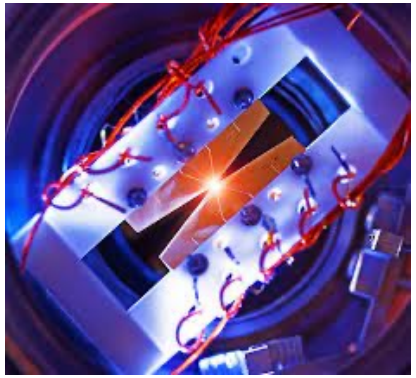
IBM 127-qubit QPU (Nov. 2021)

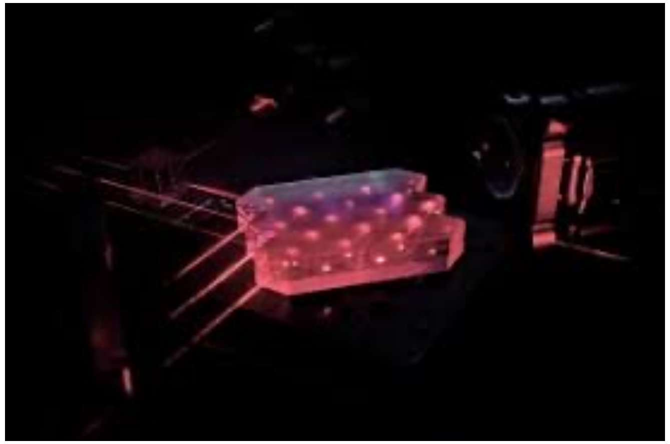


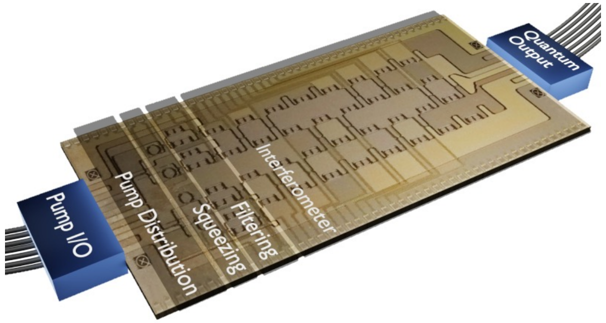
IBM QPU development roadmap (as of 2020)

QPU Hardware Realizations

- Superconducting
- Trapped ion
- Neutral atoms
- Nuclear magnetic resonance
- Quantum annealing
- Continuous wave
- Gaussian Boson sampling
- Tunable Kerr photonics
- Linear optical
- …

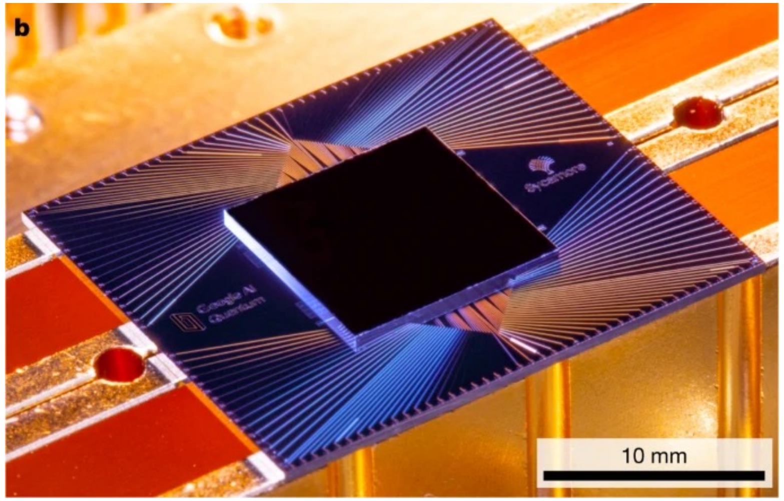


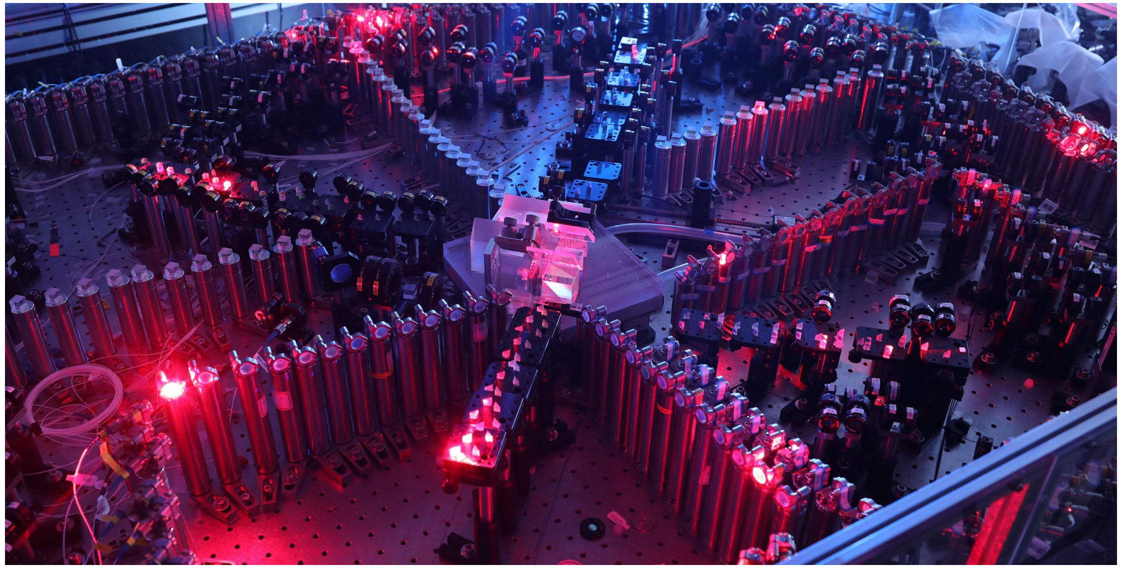




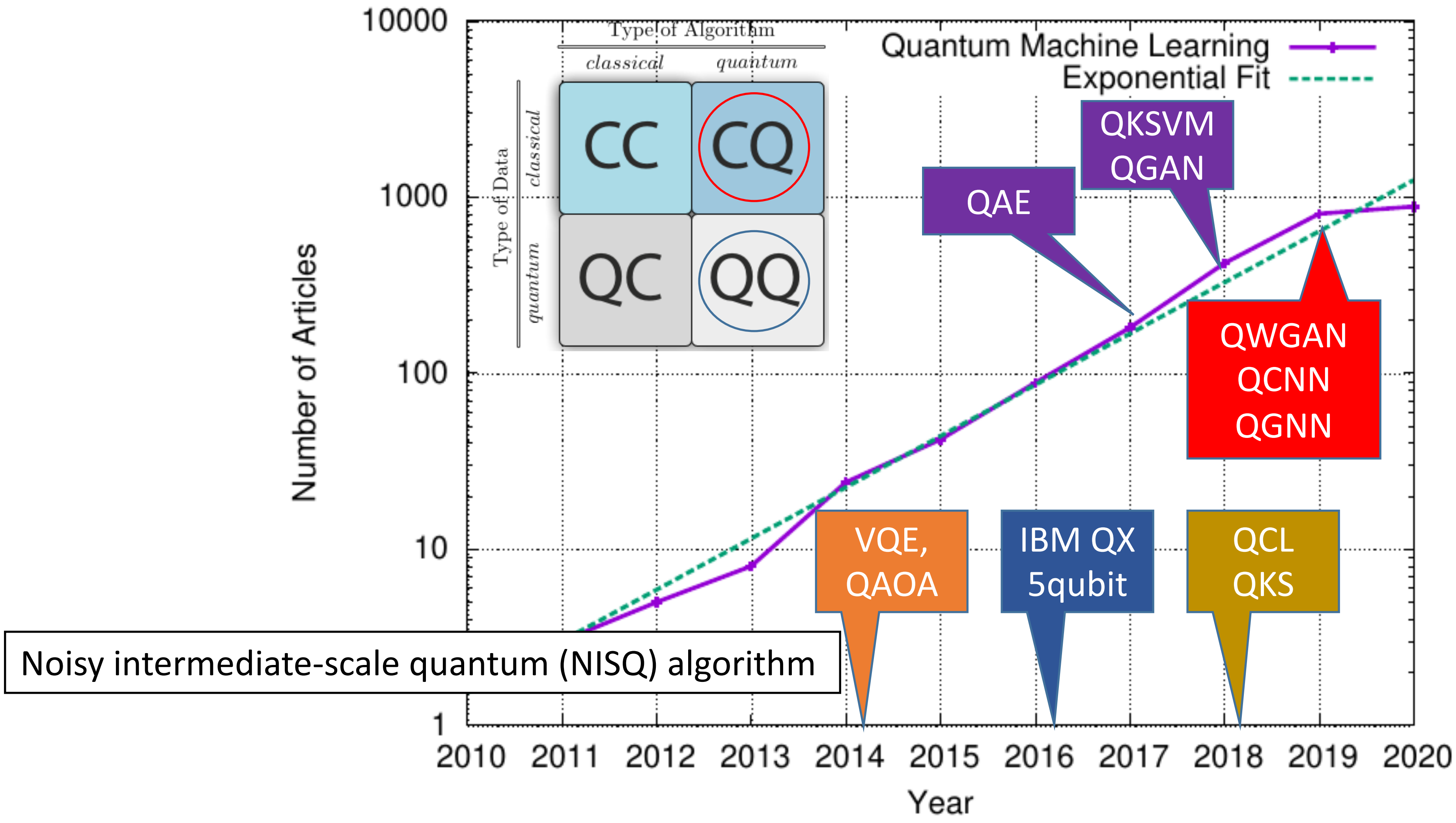
Quantum Supremacy

- Some reports claiming to have achieved ***quantum supremacy***:
  - Arute, F., Arya, K., Babbush, R. *et al.* Quantum supremacy using a programmable superconducting processor. *Nature* **574,** 505–510 (2019). https://doi.org/10.1038/s41586-019-1666-5
    - 53-qubit QPU: 200 sec. for 10,000-year job required for classical computer
  - Zhong HS, Wang H, Deng YH, Chen MC, Peng LC, Luo YH, Qin J, Wu D, Ding X, Hu Y, Hu P. Quantum computational advantage using photons. Science. 2020 Dec 18;370(6523):1460-3.
    - Boson sampling: 10^{14} faster than classic computer
- Quantum advantage is still argued for general applications





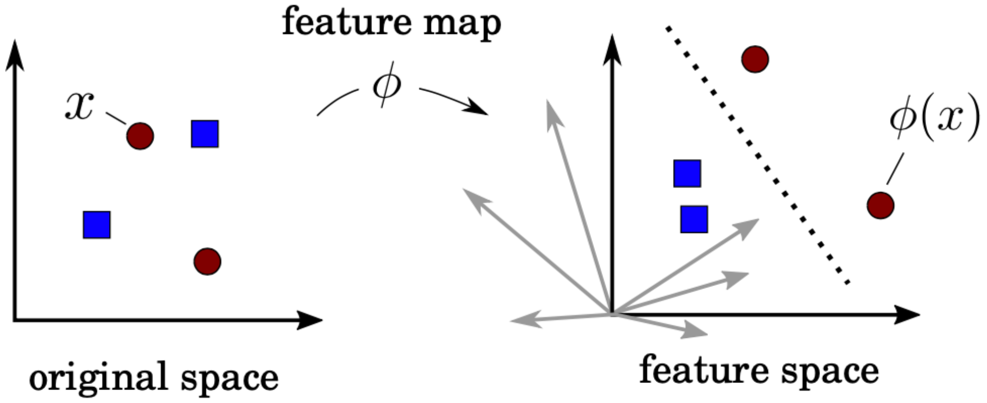
Quantum Machine Learning (QML)



VQE: Variational Quantum Eigensolver, QAOA: Quantum Approximate Optimization Algorithm
QAE: Quantum AutoEncoder, QKSVM: Quantum Kernel Support Vector Machine, Q(W)GAN: Quantum (Wasserstein) Generative Adversarial Network,
QCNN: Quantum Convolutional Neural Network, QGNN: Quantum Graph Neural Net, QX: Quantum Experience, QCL: Quantum Circuit Learning, QKS: Quantum Kitchen Sink

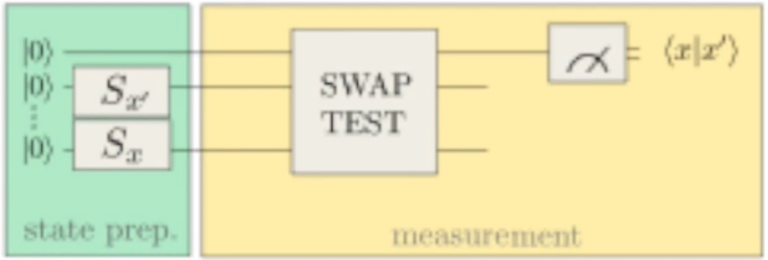
Quantum as Kernel

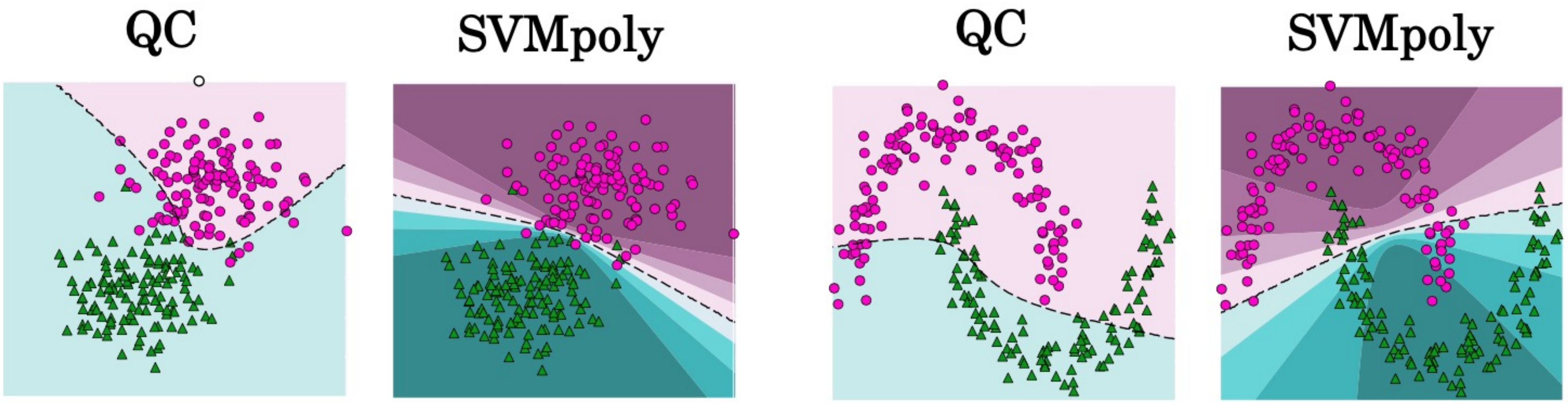
- Quantum operation is interpreted as Hirbert-space kernel operation [Schuld/Havlicek2018]



$$\kappa(x, x') = |\langle \phi(x) | \phi(x') \rangle|^2$$

Inner-product Kernel trick
= Overlap wavefunctions

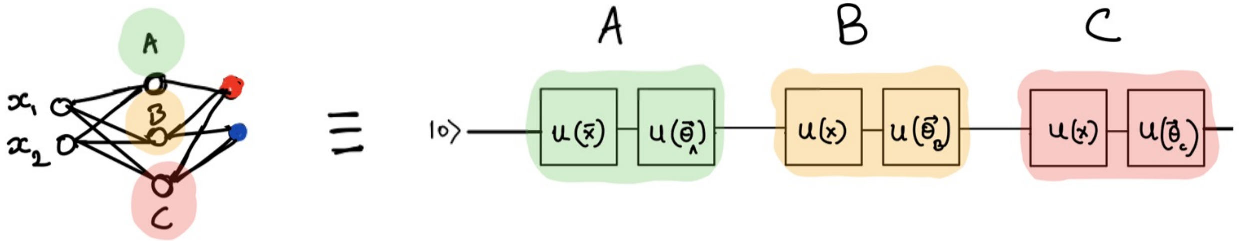




Universal Approximation Theorem/Property (UAT/UAP)

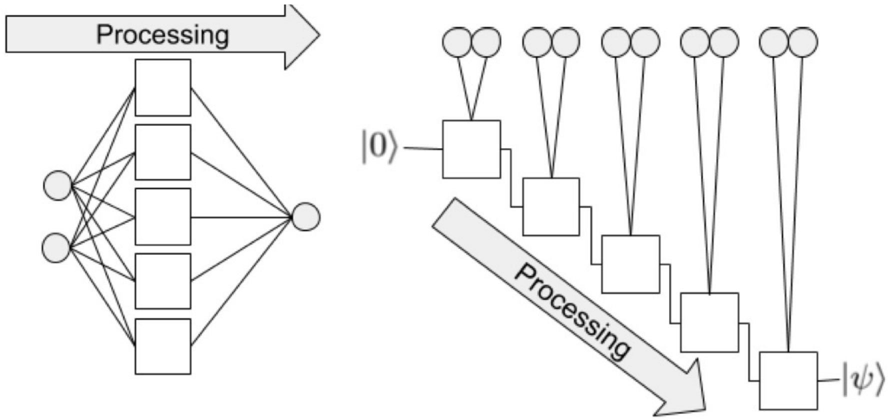
- Single hidden neural networks can approximate arbitrary bounded continuous functions [Cybenko 1989]
- Deep hidden neural networks can asymptoticly approximate arbitrary functions [Zhou 2017]
- UAP still holds for quantum processing [Perez 2019]
  - *Data re-uploading* trick

Wider/deeper neurons can reduce approximation error

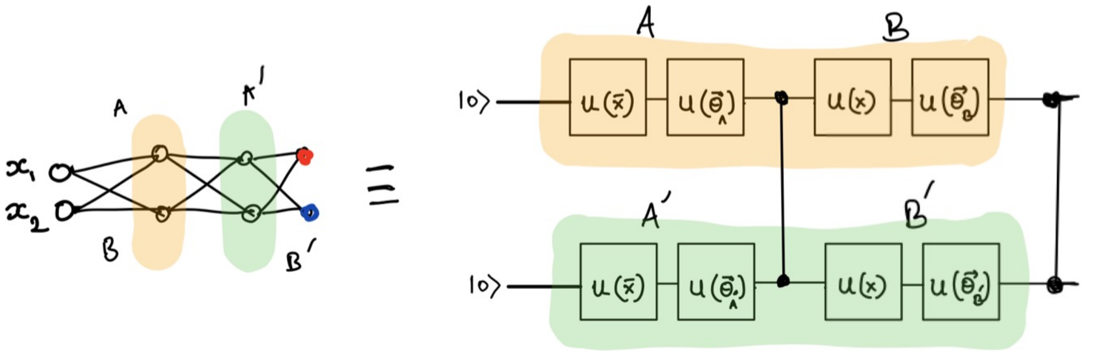


$$h(\vec{x}) = \sum_{i=1}^{N} \alpha_i \; \varphi\left(\vec{w}_i \cdot \vec{x} + b_i\right)$$

$$\mathcal{U}(\vec{x}) = U_N(\vec{x}) U_{N-1}(\vec{x}) \cdots U_1(\vec{x}) = \prod_{i=1}^{N} e^{i\vec{\omega}(\vec{\phi}_i(\vec{x}))\cdot\vec{\sigma}}$$



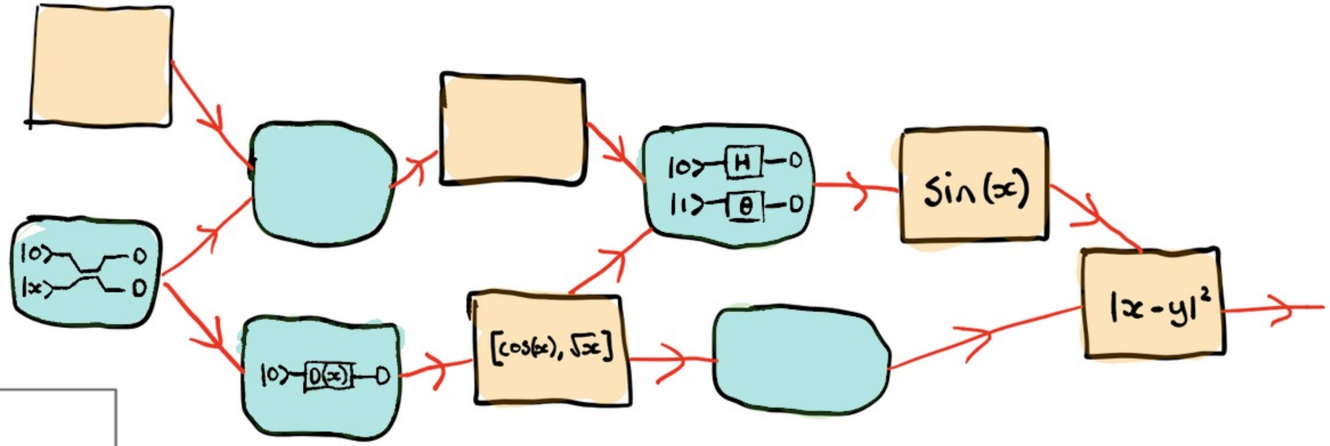
(a) Neural network (b) Quantum classifier

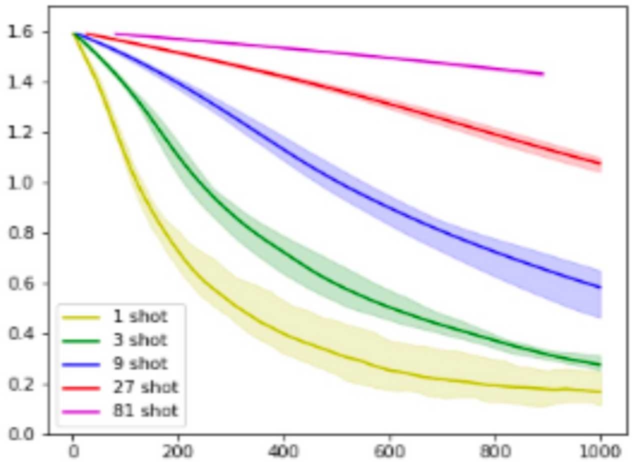


Wider neuron → More quantum layers
Deeper layer → More qubits

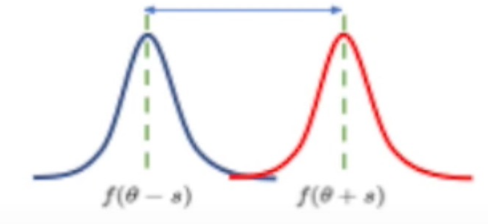
Differential Programming

- Quantum operation is differentiable:
  - *Parameter shift rule* [Mitarai/Schuld 2018] (exact gradient)
- Backpropagation through hybrid classical/quantum chips





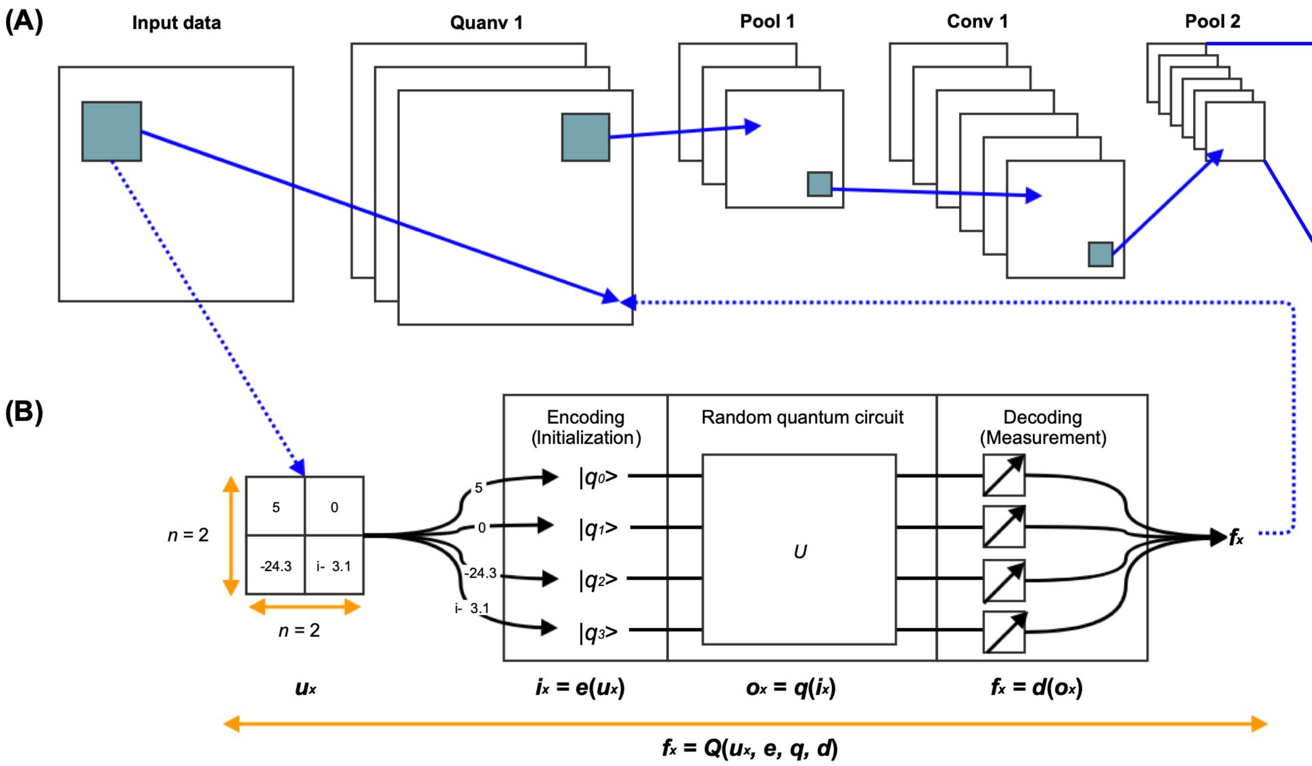
$$\boxed{\partial_\mu f(\mu) = c\big(f(\mu + s) - f(\mu - s)\big)}$$



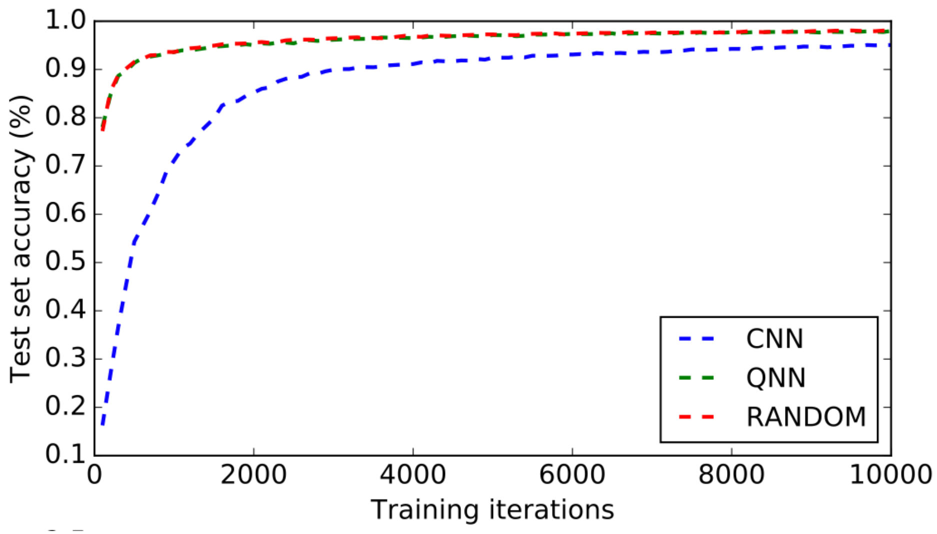
$$\partial_\mu f(\mu) \approx \frac{f(\mu + \frac{1}{2}\Delta\mu) - f(\mu - \frac{1}{2}\Delta\mu)}{\Delta\mu}$$

Quantum Convolution

- Quanvolutional Neural Network [Henderson2019]

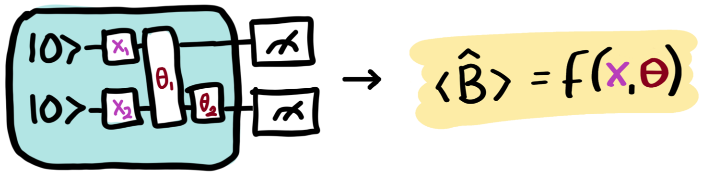


4-qubit QPU



Trainable Quantum Circuits as Parameterized DNN

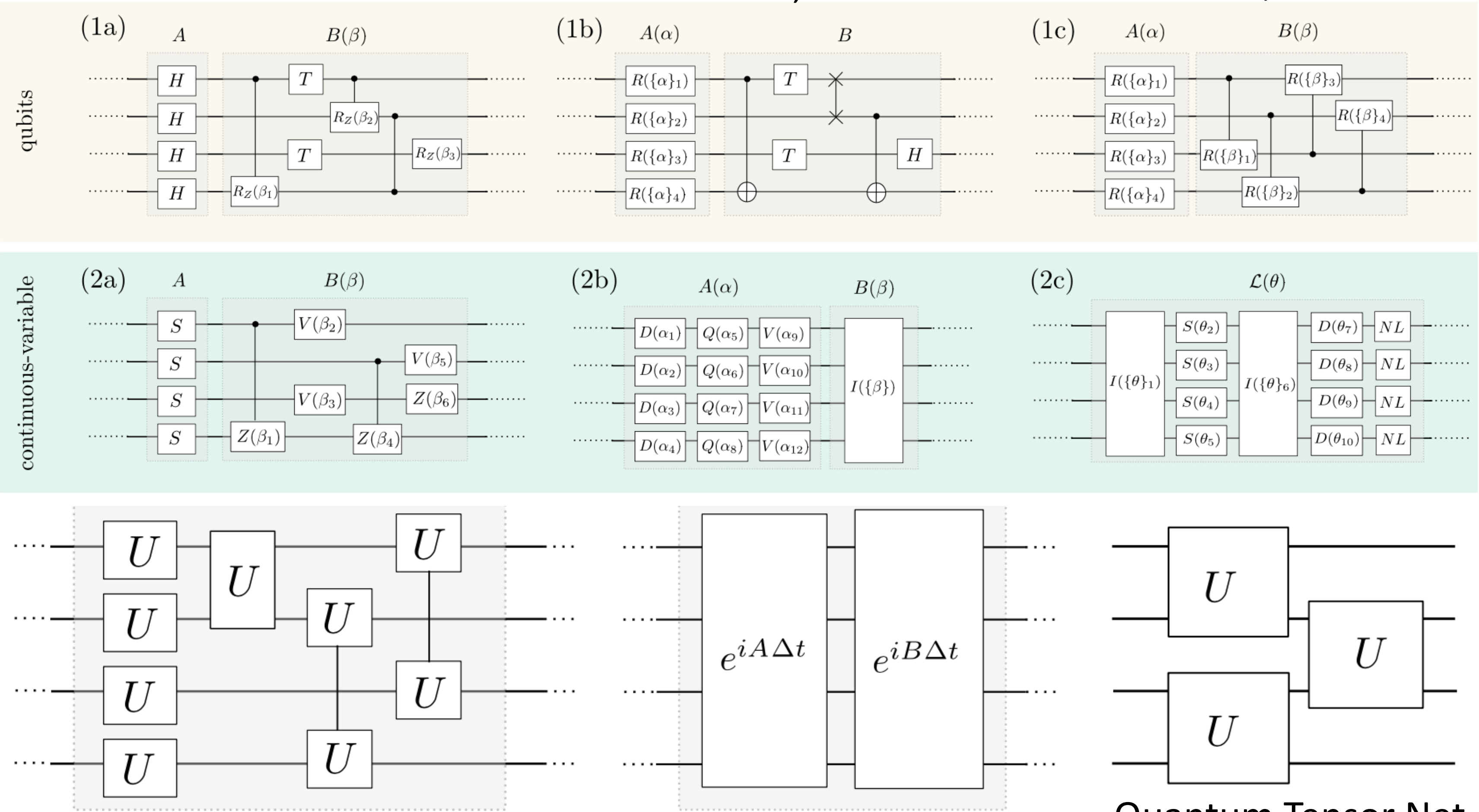
- Parametric quantum ansatz optimization



A-fix, B-variant

A-variant, B-fix

A-variant, B-variant



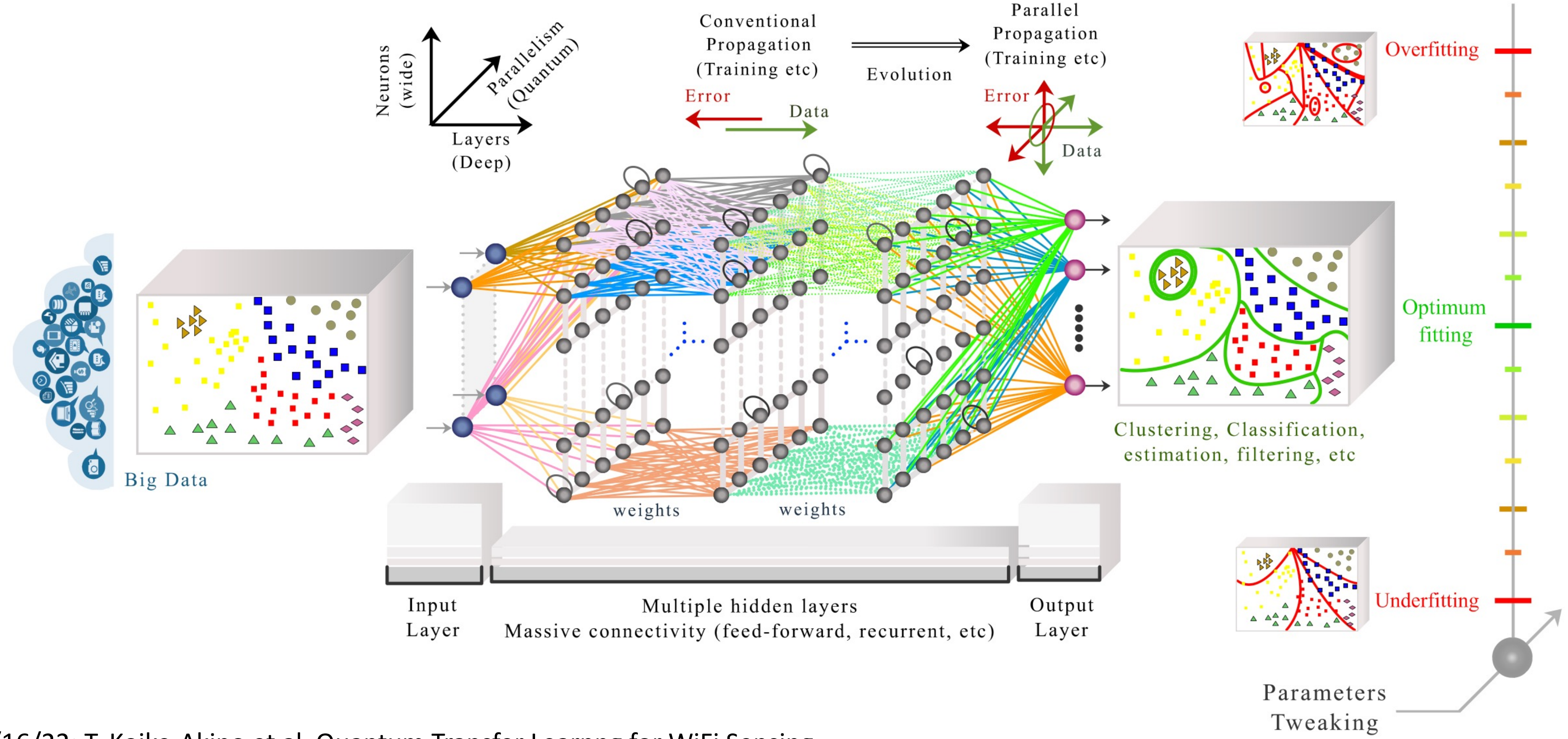
Layered gate

Alternating operation

Quantum Tensor Net

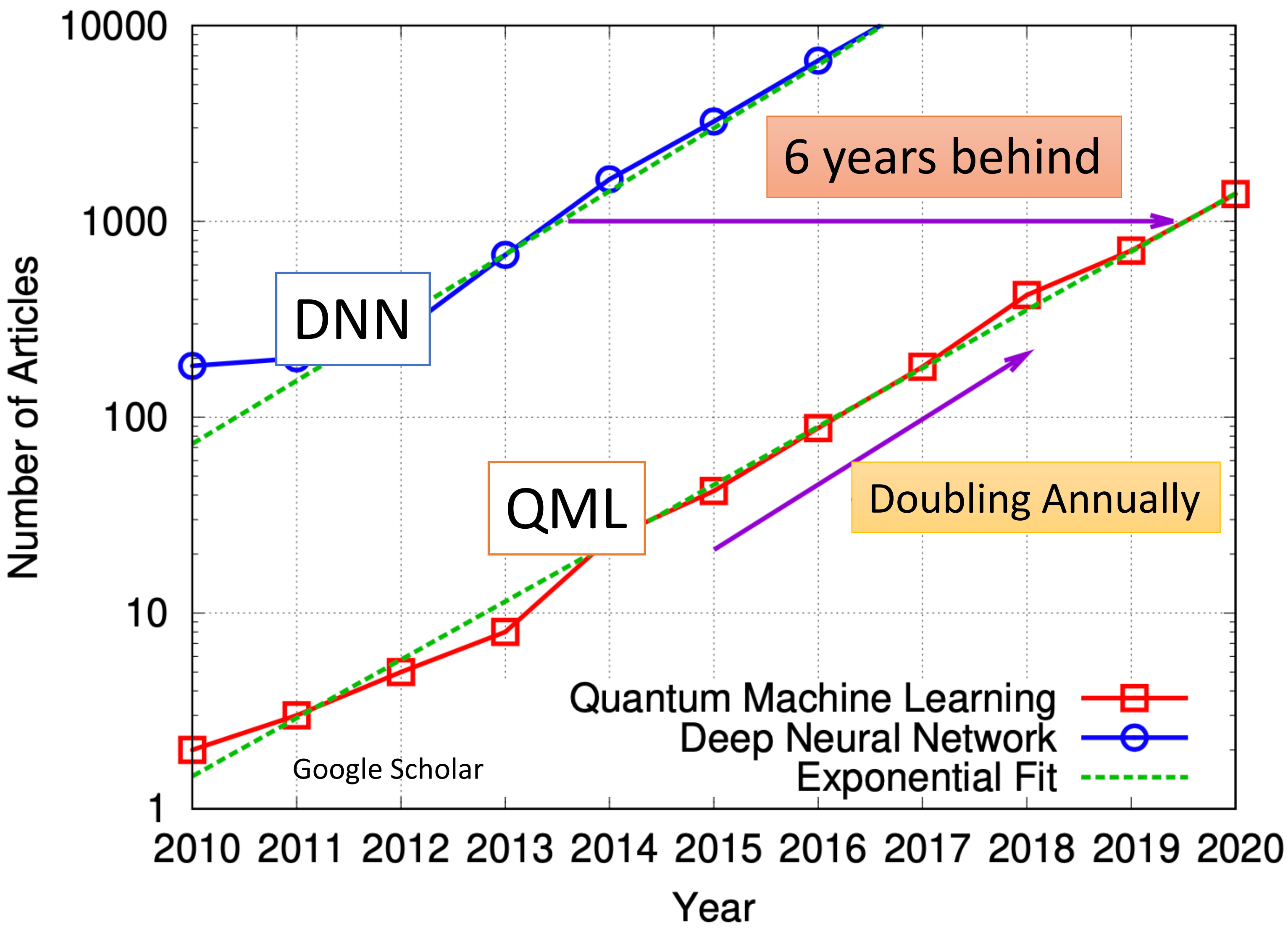
Quantum Neural Network (QNN)

- QML is a key major driver for **6G applications** [Nawaz et al. Access 2019]
- (Hyped) expectation of QNN advantage:
  - Fewer trainable parameters to support exponentially large quantum states in parallel
  - Parallel ensemble to prevent overfitting and underfitting
  - Low-power processing



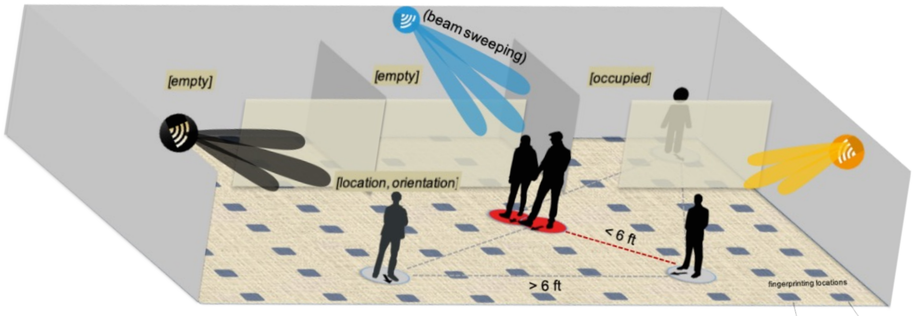
Quantum Machine Learning (QML)

- Number of articles on QML is doubling annually, just **6 years** behind of DNN

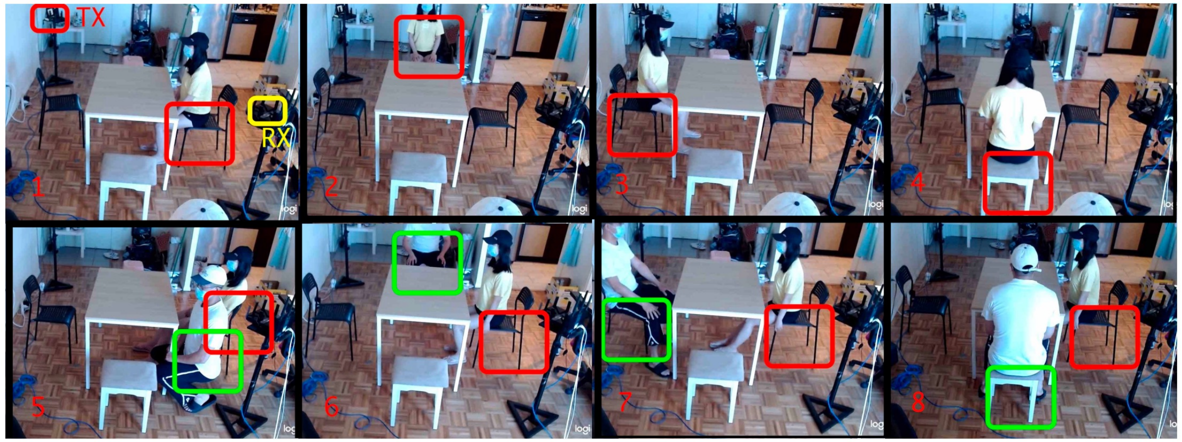


QML Meets WiFi Sensing

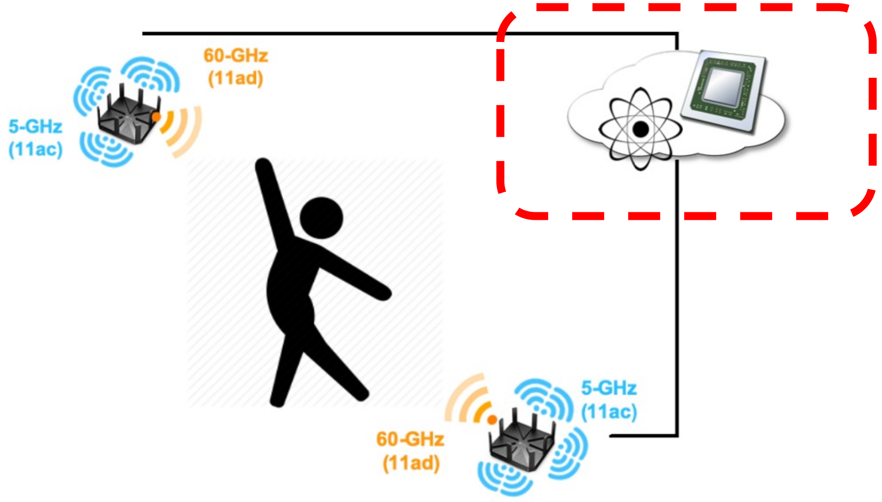
- Indoor Localization: [Koike-Akino, et al., "Fingerprinting-Based Indoor Localization with Commercial MMWave WiFi: A Deep Learning Approach", Access 2020]
- Human Monitoring: [Yu, et al., "Human Pose and Seat Occupancy Classification with Commercial MMWave WiFi", GLOBECOM 2020]



Indoor localization



Occupancy Sensing

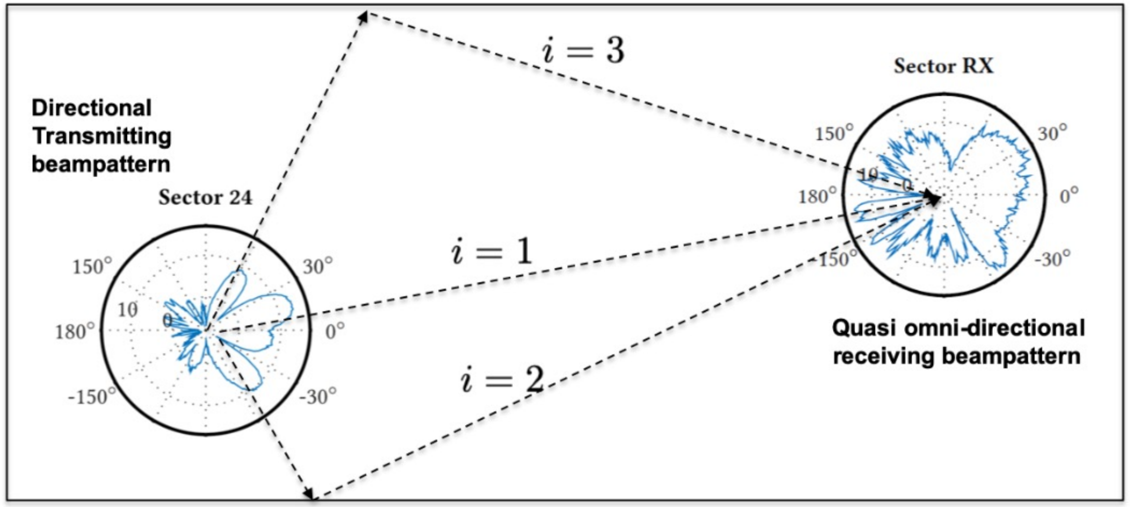


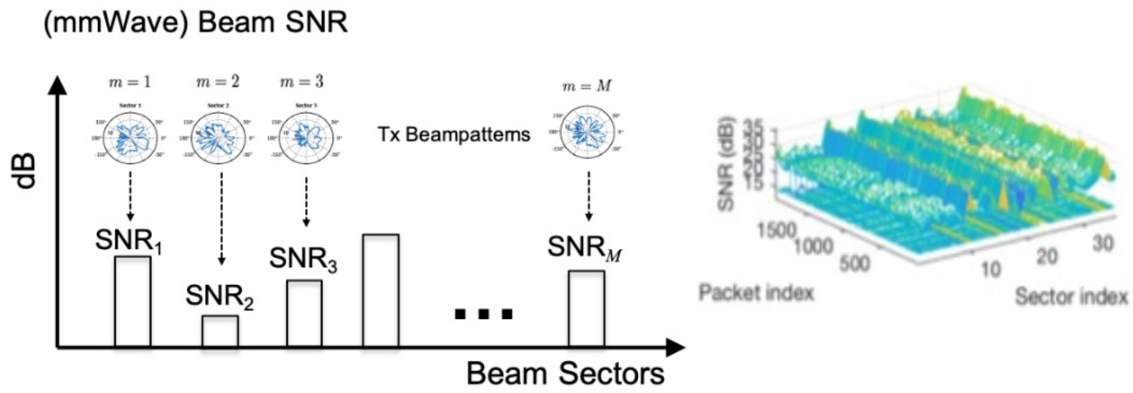
(a) Wi-Fi pose recognition empowered by QML

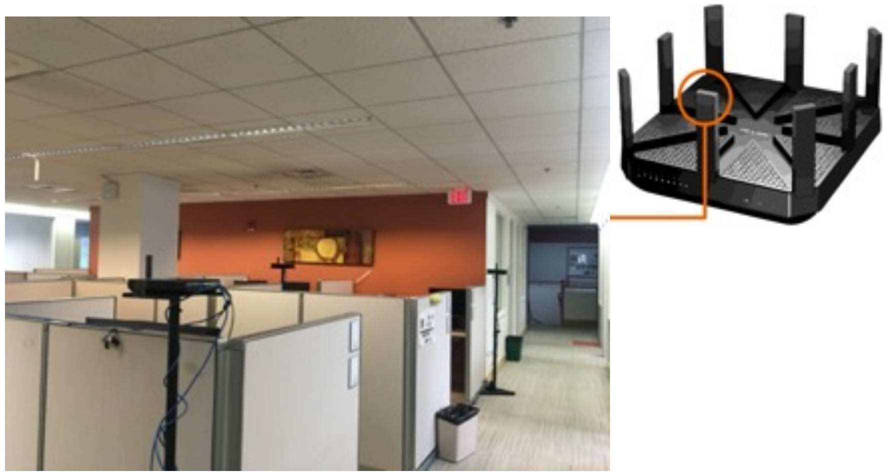
(b) Pose snapshots

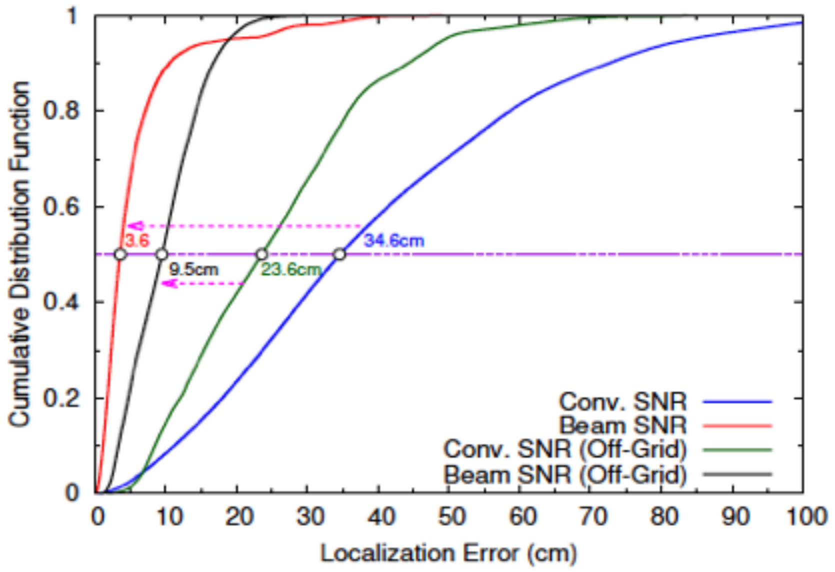
Beam SNR Measurement for mmWave Sensing

- IEEE 802.11ad protocol implements beam search/tracking
- Beam search can be used for indoor sensing; beam SNR



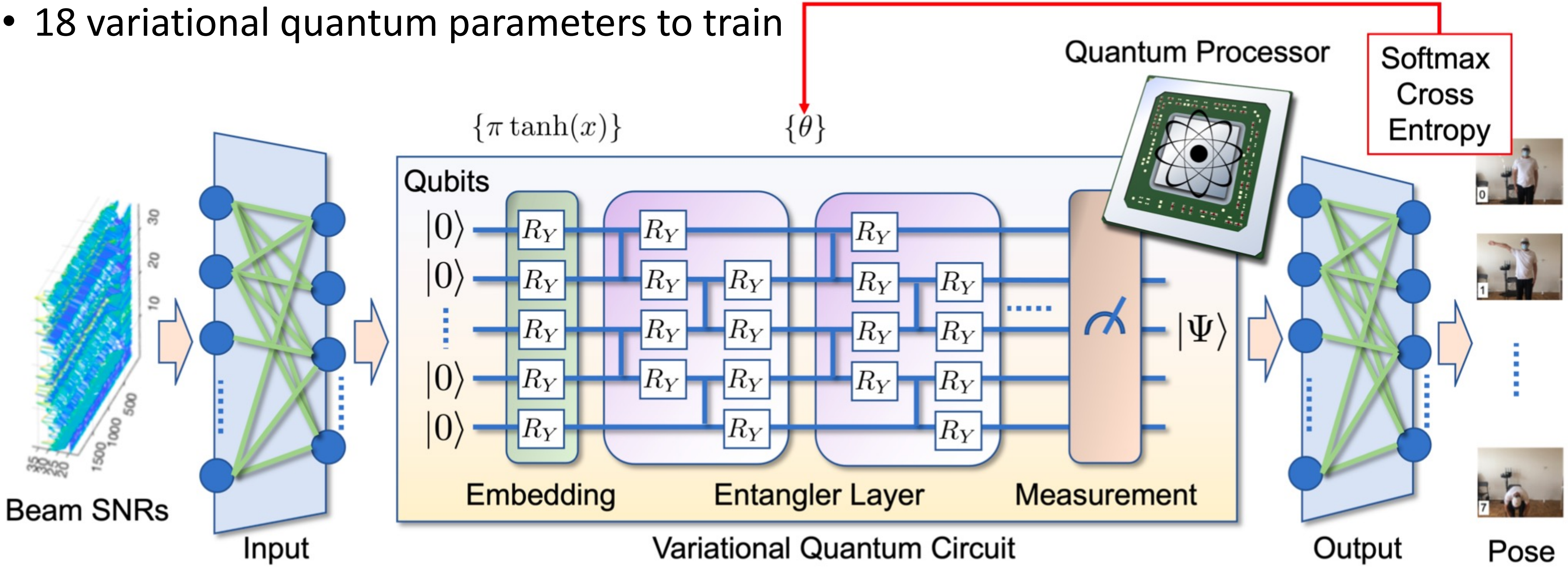






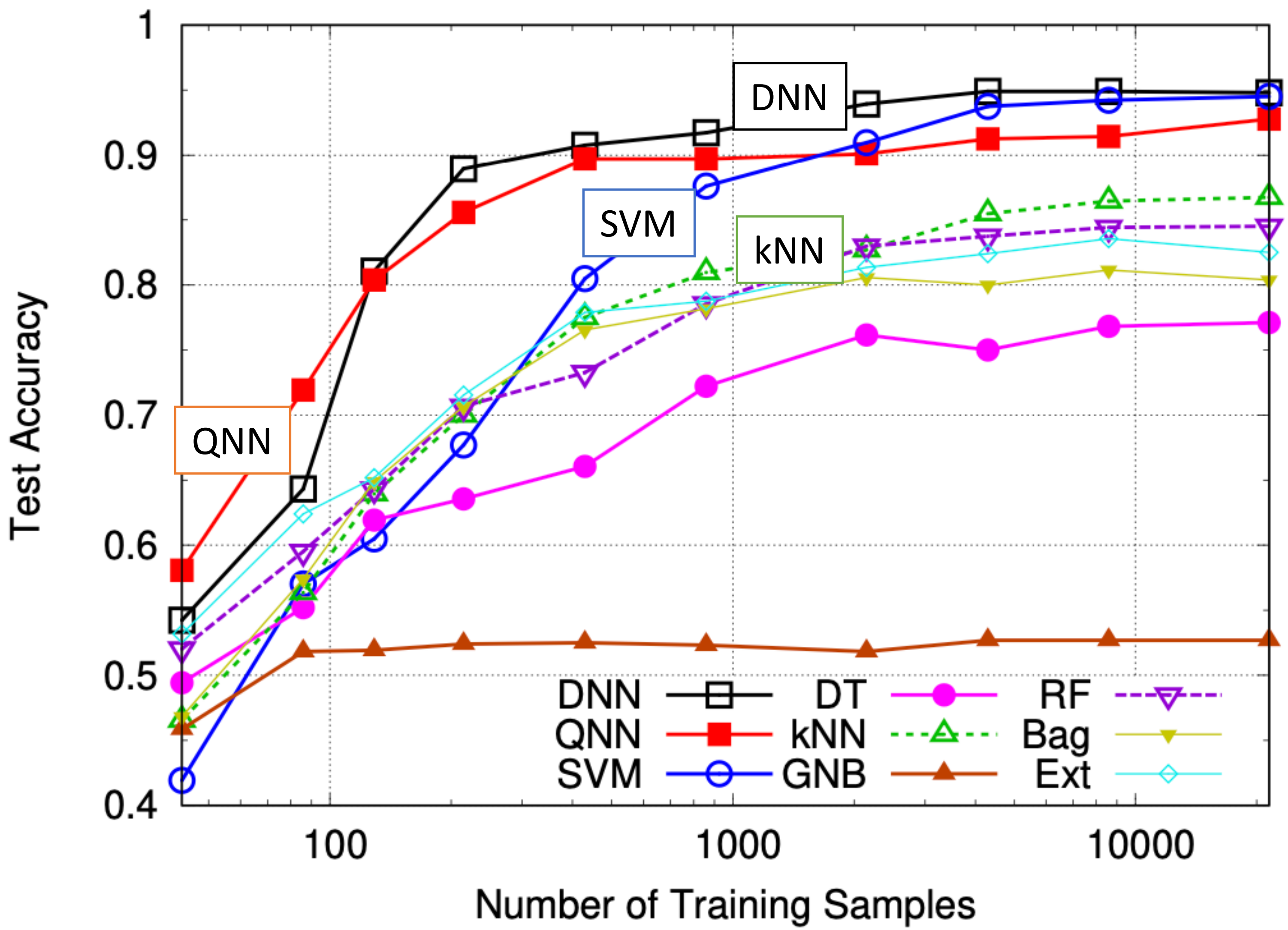
Quantum Neural Network

- Simplified two-design (STD) ansatz: https://arxiv.org/abs/2001.00550
  - Staggered Pauli-Y rotations with controlled Z gates
  - Holding statistical properties identical to ensemble random unitaries with respect to the Haar measure up to the first 2 moments: SO(2^N) -> 2N
- 18 variational quantum parameters to train



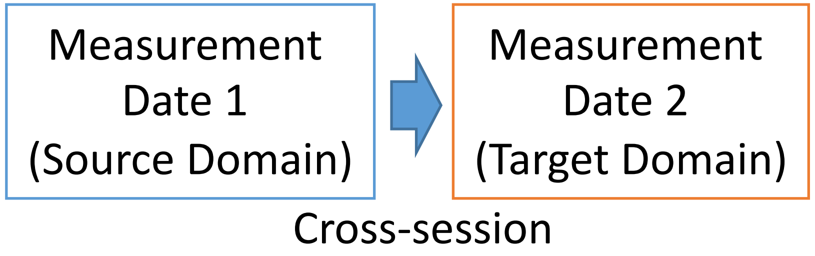
Comparison of Machine Learning Algorithms

- Various ML methods



DNN: 4-hidden; 100 nodes; Mish; 35k params
QNN: 2-layer STD ansatz; 18 params

(b) Pose snapshots



Domain Shift Issues: Within-Session vs. Cross-Session

- Source domain to target domain

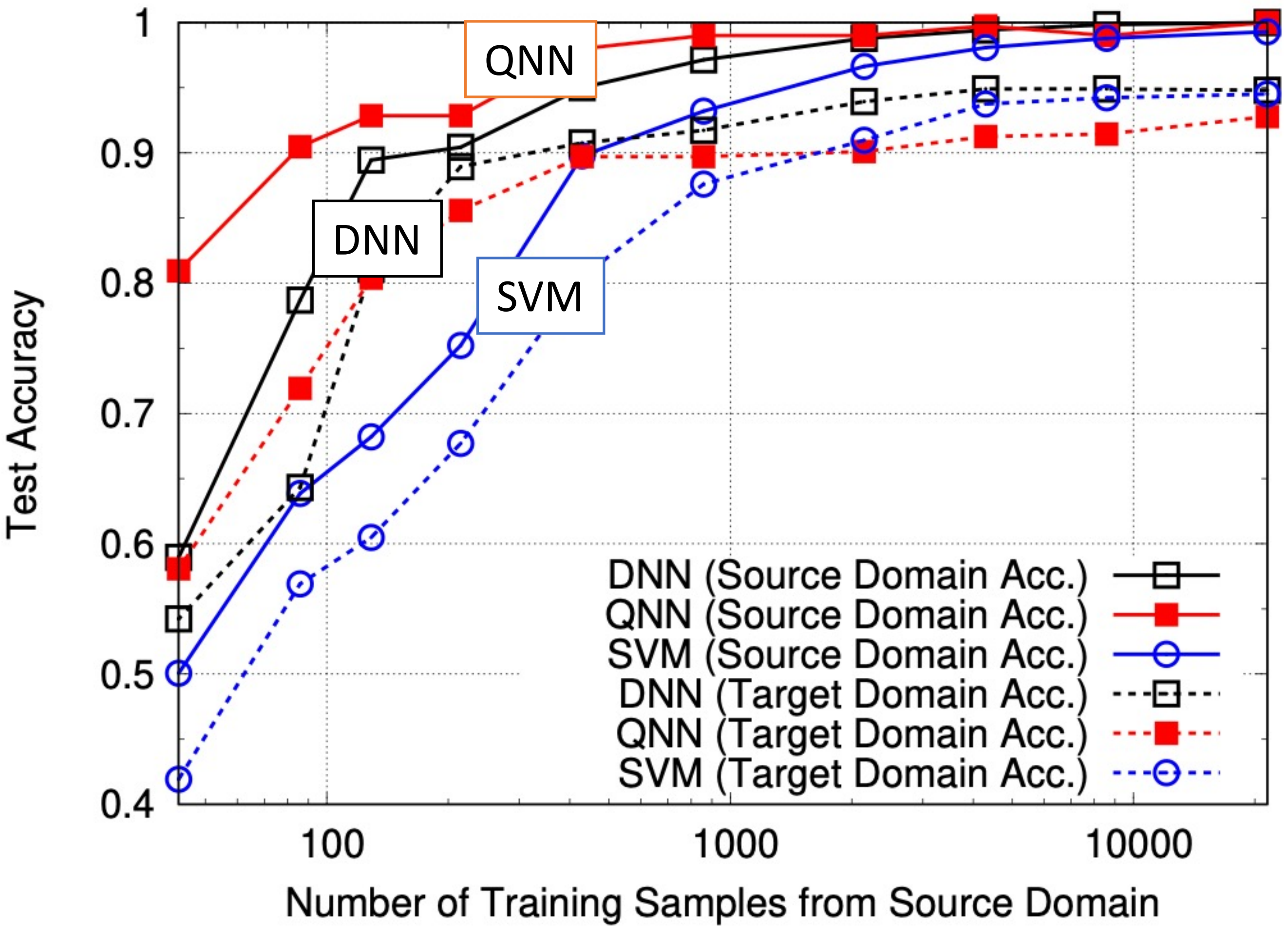
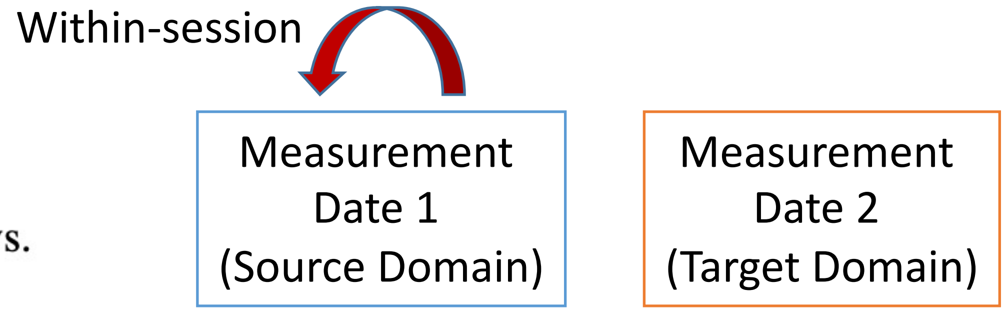


Fig. 7. Domain shift analysis with test accuracy (source/target domains) vs. training samples (source domain).

(b) Pose snapshots



Quantum Transfer Learning

- QNN performs comparable to DNN (35k params) while just 18 params

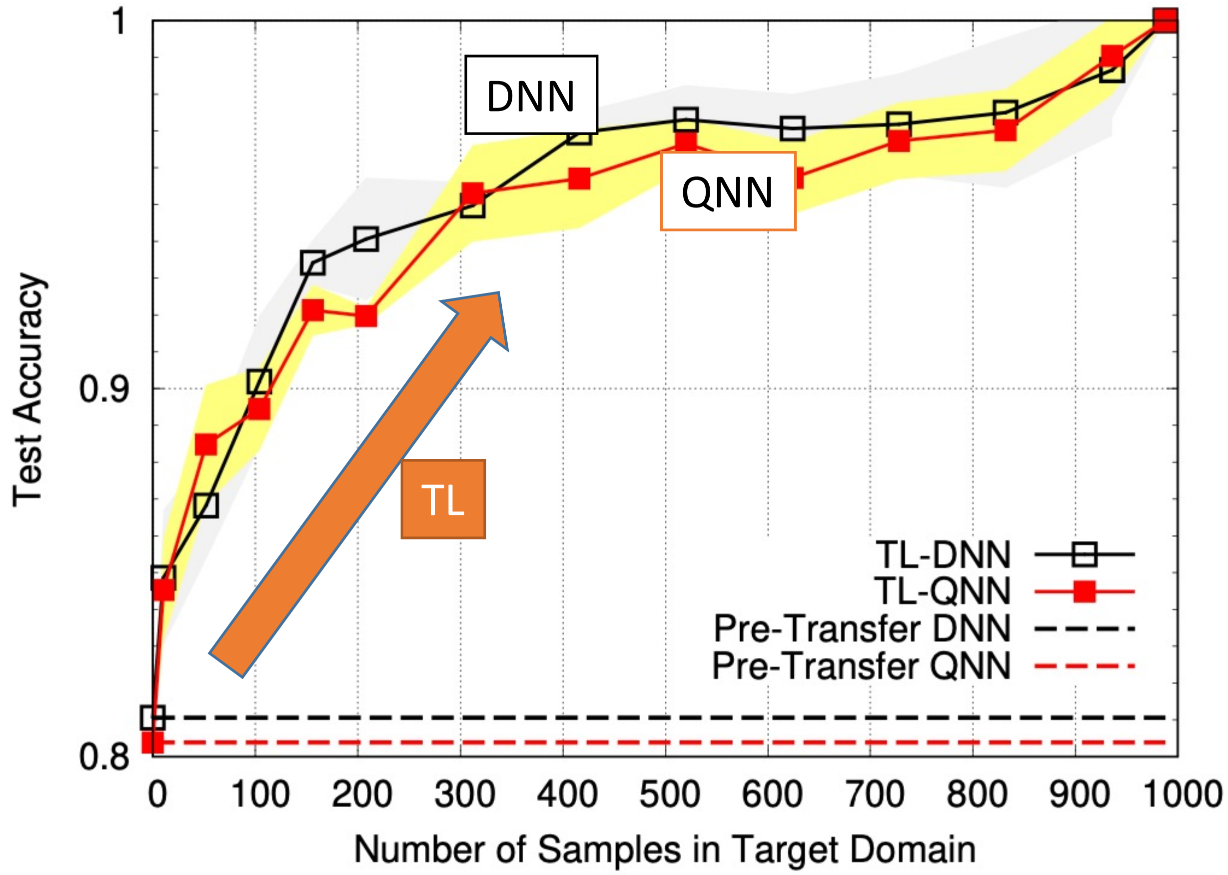
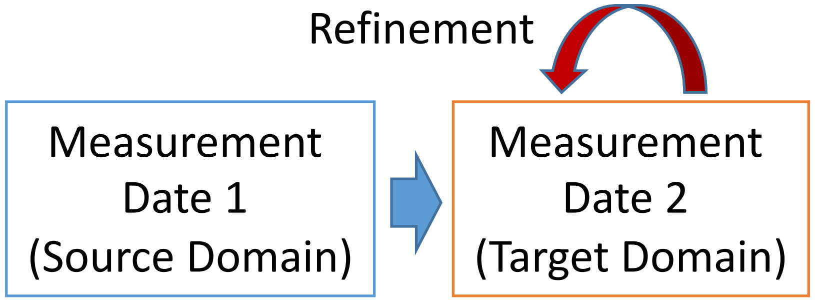
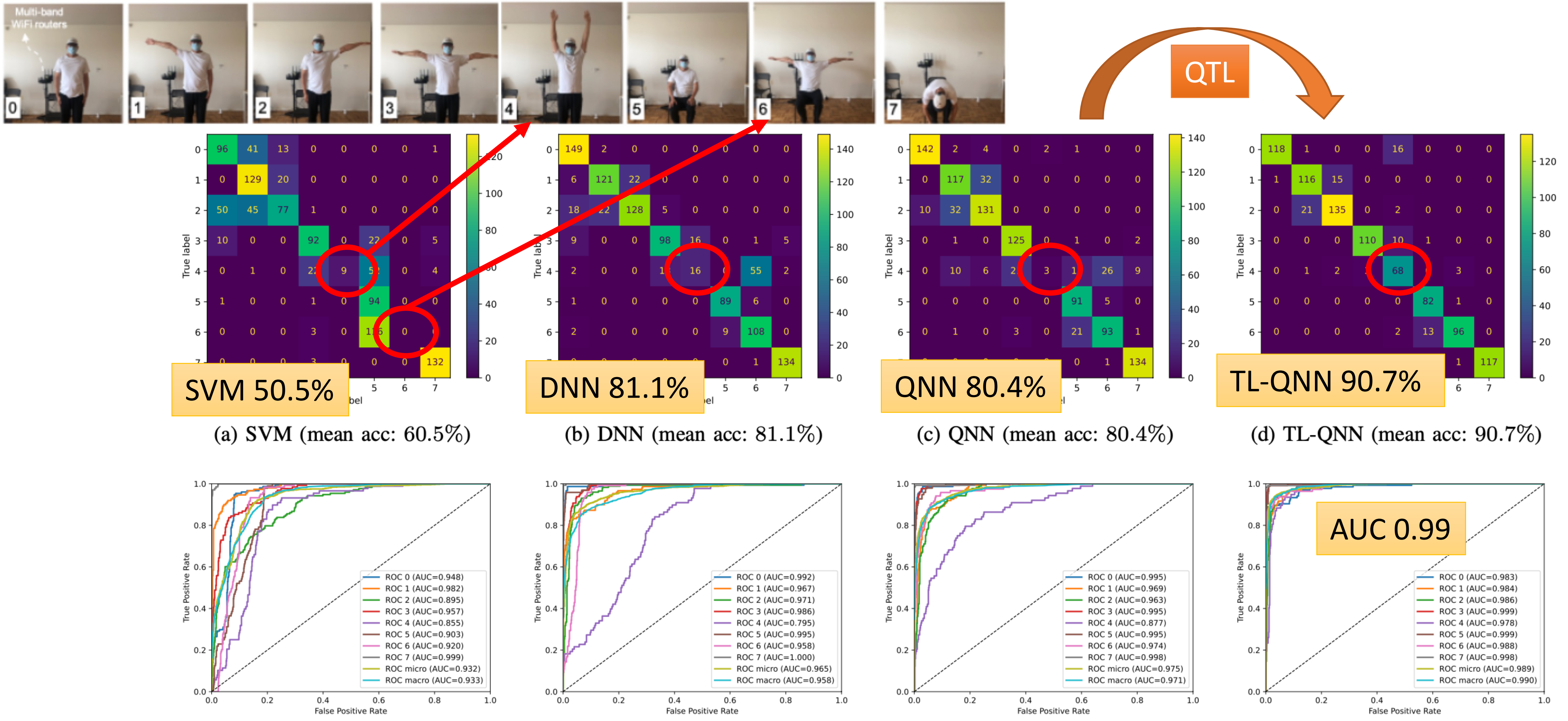


Fig. 8. Test accuracy vs. transfer samples (labeled data in target domain). Pre-transfer model uses 129 training samples (0.3% data labeled in source domain).

(b) Pose snapshots



Transfer Learning Gain



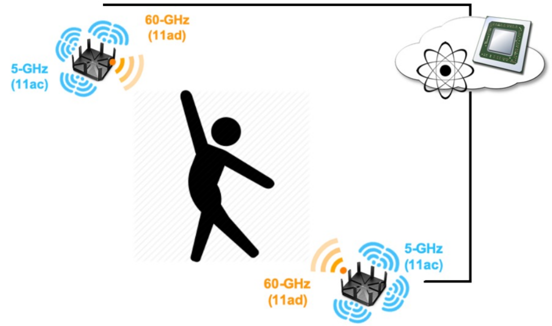
(a) SVM (mean acc: 60.5%) (b) DNN (mean acc: 81.1%) (c) QNN (mean acc: 80.4%) (d) TL-QNN (mean acc: 90.7%)

(e) SVM (AUC: 0.933) (f) DNN (AUC: 0.958) (g) QNN (AUC: 0.971) (h) TL-QNN (AUC: 0.990)

Fig. 6. Confusion matrices (top row) and ROC curves (bottom row) for 8-pose recognition with 129 training samples (0.3% data labeled in the source domain). TL-QNN uses 104 transfer samples (10% data labeled in the target domain).

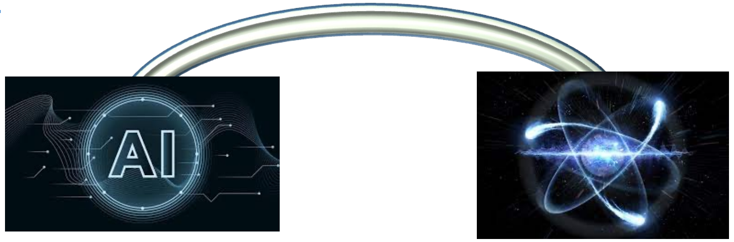
Conclusions

- We showed recent AI trends overview: ML for everything in community
- We overviewed recent advancement on **QML**
- We proposed QML transfer learning in **integrated sensing & communications (ISAC)**
  - Demonstrated the first proof-of-concept study for future quantum-era
  - Experimented the feasibility of COTS WiFi sensing systems for indoor monitoring
  - Showed significant gain to deal with a measurement domain shift issue
  - Achieved state-of-the-art DNN performance with few-parameter QML
  - Validated nearly 100% accuracy for pose estimation
- There are many fascinating topics and high potentials for future work
- Questions?
  - Please contact me: koike@merl.com



(a) Wi-Fi pose recognition empowered by QML

AI x Quantum



(b) Pose snapshots

MITSUBISHI ELECTRIC

Changes for the Better