

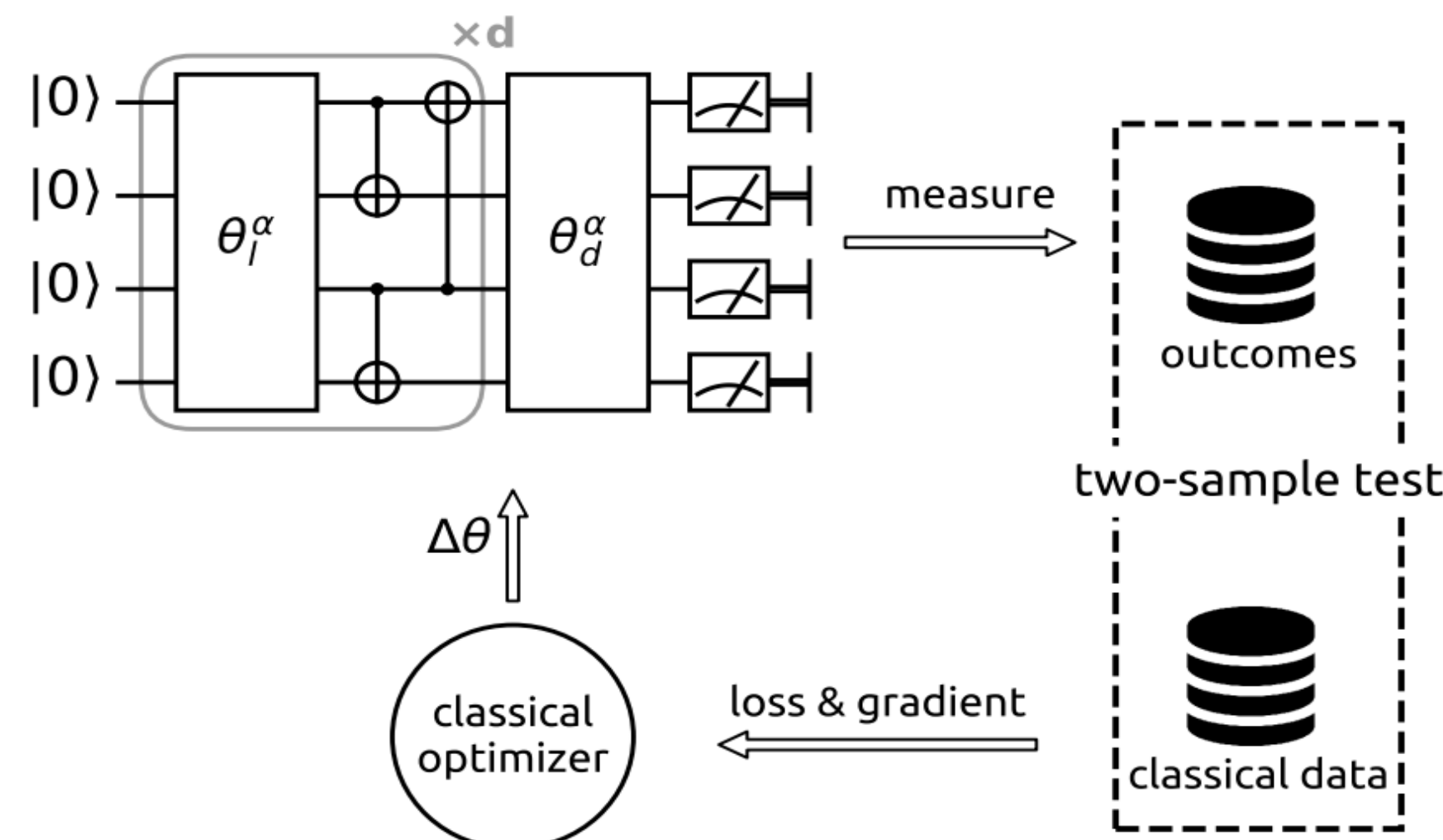
Quantum Augmentation for Imbalanced Cancer Images Using Amazon Braket

2026 Quantum Machine Learning REU Program

Sophia Knutson¹, Nandika Goyal², Gennaro De Luca³, Glen Uehara², Andreas Spanias²
¹SCS, CMU, ²School of ECEE, ASU, ³SCAI, ASU

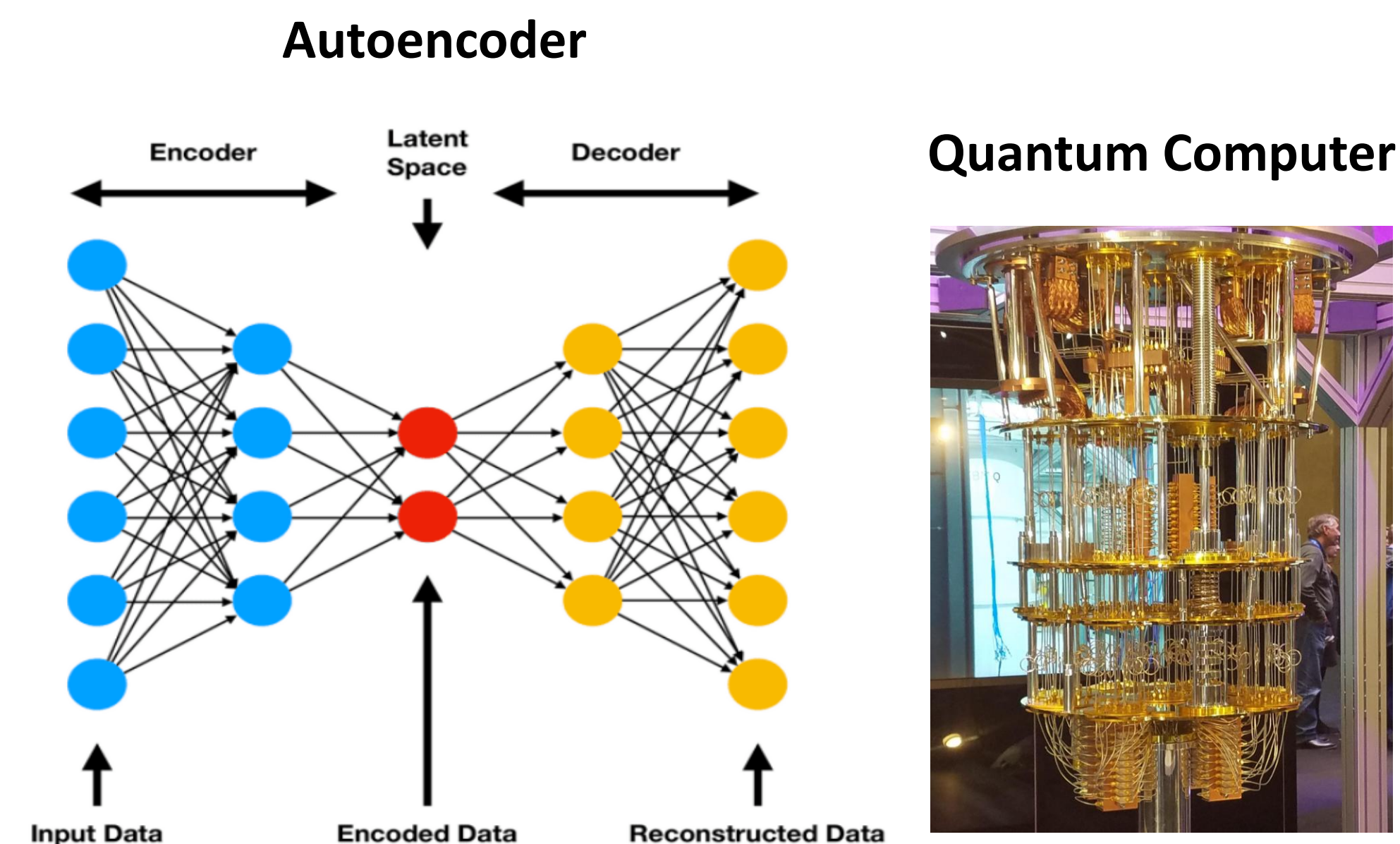
RESEARCH BACKGROUND/DESCRIPTION

- Machine learning classifiers are increasingly used for cancer detection tasks [1]
- Many cancer image datasets contain few malignant examples compared to benign ones - training on such datasets can lead to poor model performance
- Quantum Circuit Born Machines (QCBMs) outperform many classical models, specifically when training data is scarce [2]
- We can use QCBMs to generate synthetic data to augment the datasets to train more accurate classifiers



METHODS & MATERIALS

- Use BUSI dataset
- Train baseline classifier (ResNet-18) and record F1 and AUC-ROC numbers
- Loss function for variational autoencoders:
 - $\text{loss} = \text{MSE}(x, x_{\text{recon}}) + (1 - \text{SSIM}(x, x_{\text{recon}}))$
- Build and train QCBMs on Braket SV1 using MMD as loss function [3]
- Generate and decode synthetic latent vectors using QCBMs and decoders
- Train new classifier on augmented dataset (original BUSI + synthetic images)

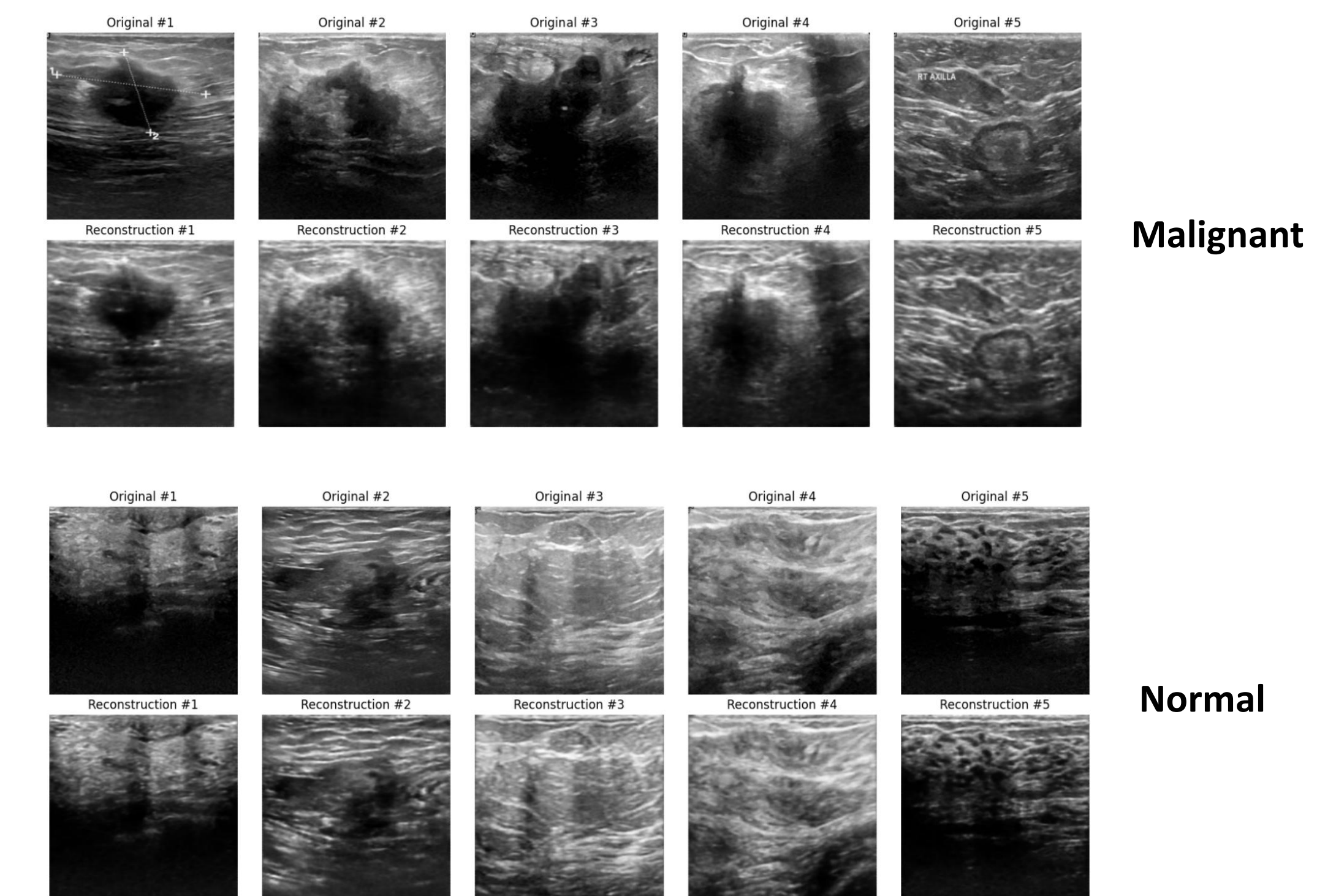


REFERENCES

- [1] N. Goyal, G. Uehara, and A. Spanias, "Quantum-Enhanced Cancer Detection for Histopathologic Images," in Proc. 2025 IEEE Int. Conf. Image Process. (ICIP), Anchorage, AK, USA, Sep. 2025, pp. 1342–1347, doi: 10.1109/ICIP55913.2025.11084595.
- [2] M. Hibat-Allah, M. Mauri, J. Carrasquilla, and A. Perdomo-Ortiz, "A framework for demonstrating practical quantum advantage: comparing quantum against classical generative models," Communications Physics, vol. 7, no. 1, p. 68, 2024.
- [3] J.-G. Liu and L. Wang, "Differentiable learning of quantum circuit born machines," Physical Review A, vol. 98, no. 6, p. 062324, 2018

PRELIMINARY RESULTS

- Benchmark classifier trained on non-augmented data: Accuracy – 0.8974, F1 Score – 0.8000, AUC-ROC – 0.9555
- Autoencoder training produced unsatisfactory results when trained on dataset with mixed classes
- Best reconstruction image quality was achieved when training separate autoencoders for separate image classes (normal, benign, malignant)



RESEARCH OBJECTIVES/PLAN

- Train two variational autoencoders (one for normal and one for malignant cancer images) to compress images into 16-dimensional latent vectors
- Train two QCBMs on normal and malignant latent vectors to generate new synthetic latent vectors
- Decode generated latent vectors to create synthetic full resolution images
- Train new classifier on augmented data and compare to baseline, GAN augmentation, and SMOTE oversampling
- Use AWS Braket for Simulations

CONCLUSIONS + FUTURE WORK

- Training separate autoencoders for separate image classes produced significant reconstructed image quality improvement
- Future work: Build and train QCBMs, generate and decode synthetic latent vectors, train final classifier, and compare to baseline model