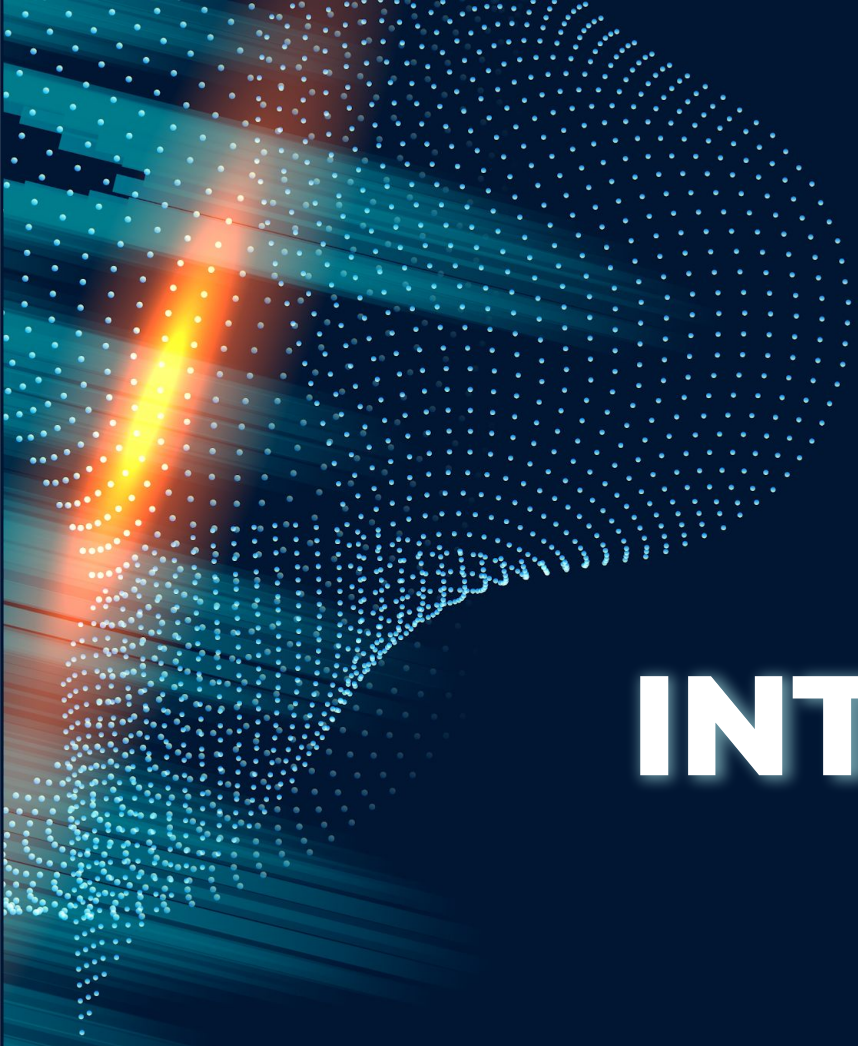
Quantum Adaptive Excitation Network with Variational Quantum Circuits for Channel Attention

Authors: Yu-Chao Hsu, Kuan-Cheng Chen, Tai-Yue Li, Nan-Yow Chen

Presentors: Chirag Raisingh, Evan Nawfal, Robert Walsh, Quinn Senych

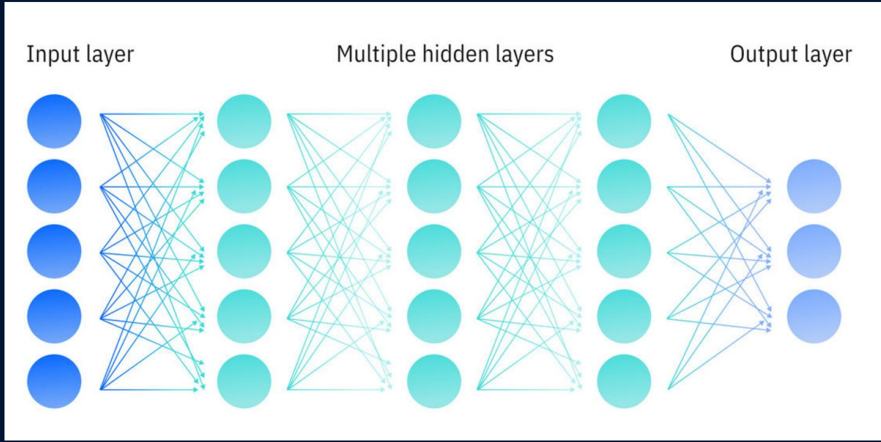
CONTENTS



INTRODUCTION

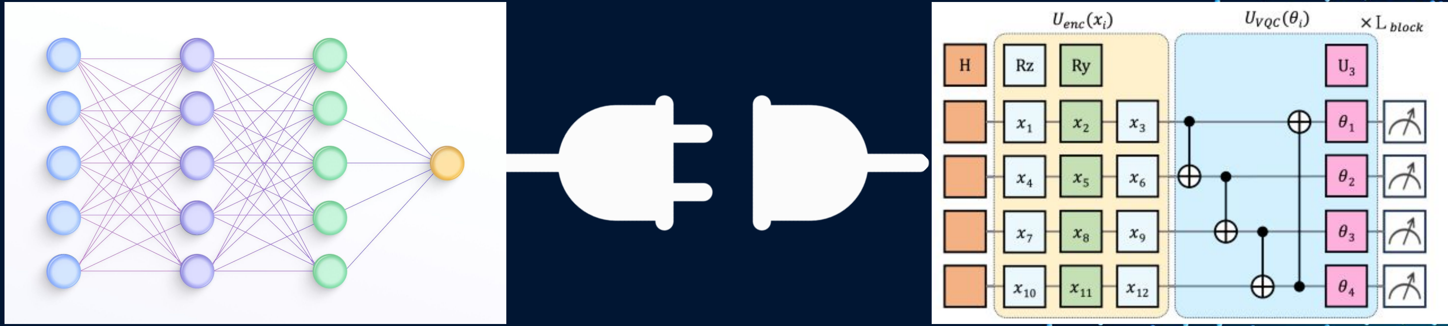
Problem: ML Models Keep Getting Bigger

- Modern Deep Learning models require massive computing power
- Networks grow deeper/wider to extract complex patterns
- Interest in new ways to increase model efficiency



Goal of the Paper

- Enhance deep learning models using quantum circuits
- Replace part of a Convolutional Neural Network (CNN) called a Squeeze and Excitation block with a trainable quantum circuit
- Evaluate accuracy and training time



01 BACKGROUND THEORY

01

CONVOLUTIONAL NEURAL NETWORKS

What are CNNs and how do they work?

02

EXCITATION BLOCKS

Why do we need excitation blocks? What is a SAE and VQC?

03

PROJECT DESIGN

What does the paper actually implement? What is our project goal?

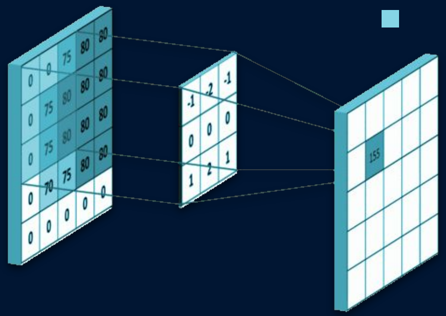
04

ADDITIONAL INFO

What datasets are used for training? How do we visualize our data?

CONVOLUTIONAL NEURAL NETWORKS (CNN)

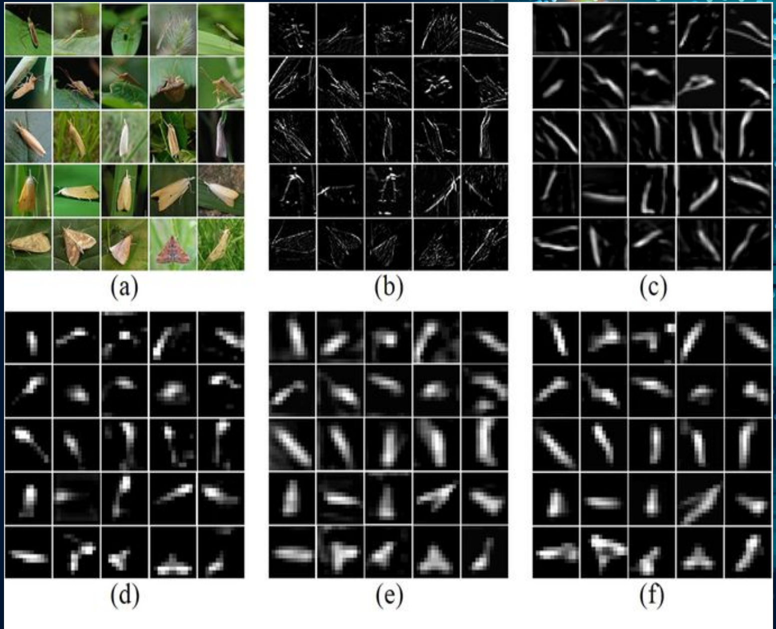
- Excels at image, audio, and signal detection
- Made up of 3 different kinds of layers:
    - Convolutional Layer
        - Kernel moves through input matrix and outputs dot product in the output matrix
    - Pooling Layer
        - Downsamples input matrix to increase efficiency
    - Fully Connected Layer
        - Connects all neurons to previous ones so all weights are taken into account



WHY CHANNEL ATTENTION MATTERS

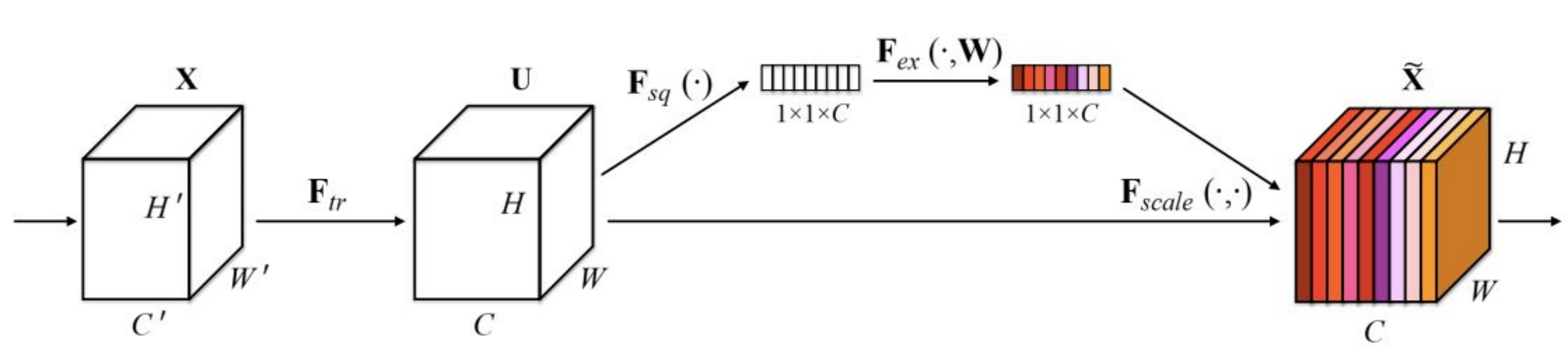
- CNN feature maps have many channels, each capturing different patterns (edges, textures, colors, semantics)

- Standard CNN's treat these channels with equal weighting, but certain channels are more relevant than others



SQUEEZE & EXCITE NETWORK (SENet)

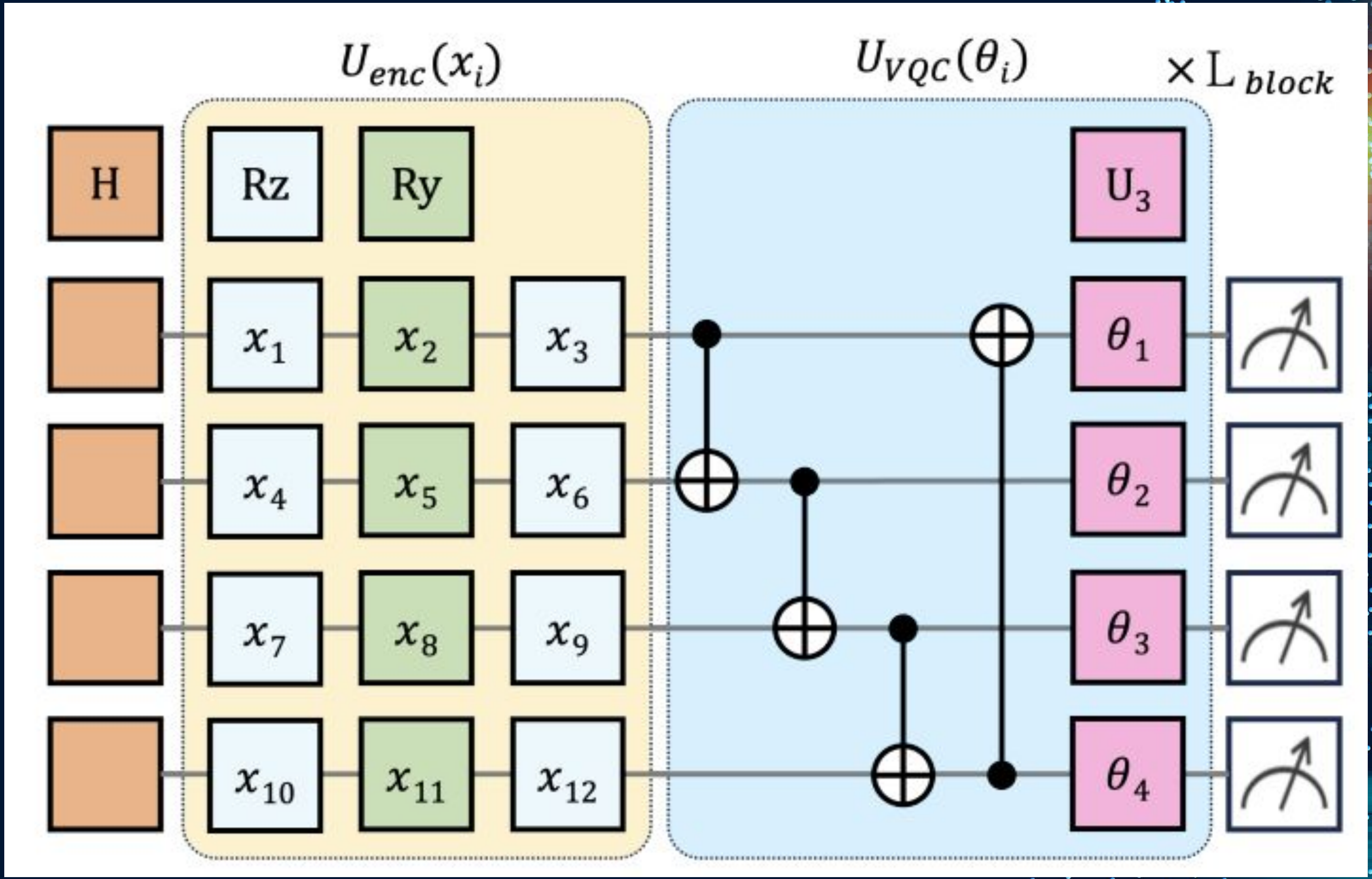
- Boosts CNN accuracy by reweighting feature channels
- Squeeze: global average pooling summarizes each channel
- Excitation: small Multilayer Perceptron learns channel importance scores
    - Specifically 2 Fully Connected Layers with Sigmoid and ReLU attention features.
- Reweights channels to amplify useful features



$\mathbf{X}$
H'
W'
C'
$\mathbf{F}_{tr}$
$\mathbf{U}$
H
W
C
$\mathbf{F}_{sq}(\cdot)$
$1\times1\times C$
$\mathbf{F}_{ex}(\cdot,\mathbf{W})$
$1\times1\times C$
$\mathbf{F}_{scale}(\cdot,\cdot)$
$\widetilde{\mathbf{X}}$
H
W
C

VARIATIONAL QUANTUM CIRCUITS (VQC)

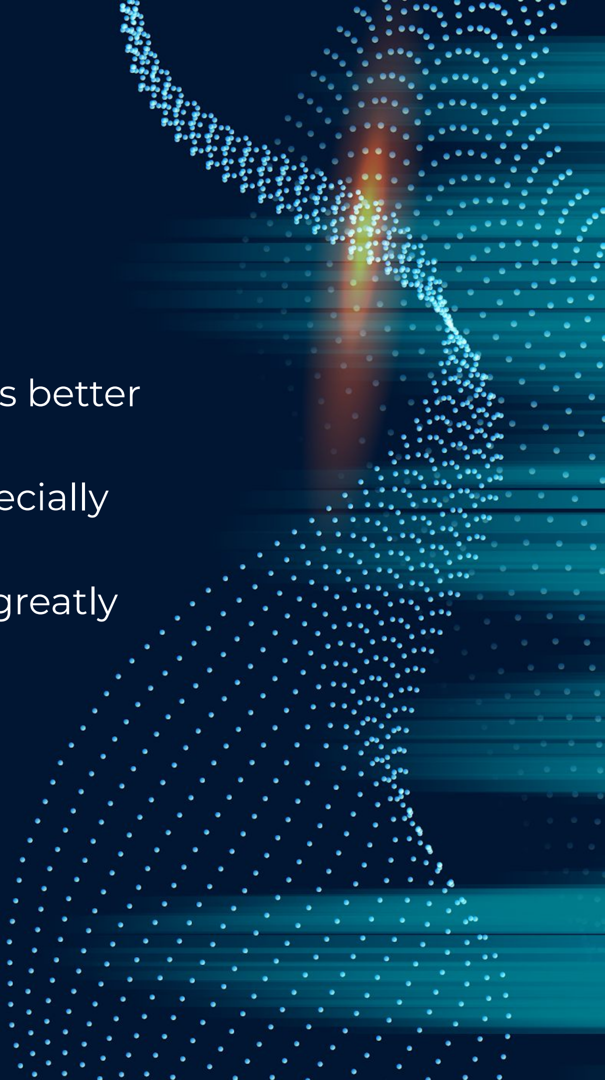
- VQCs are circuits that depend on free parameters
- Follows 3 main steps:
    - Preparation of Initial State
    - Quantum Circuit, U(θ), with some ansatz value of θ.
    - Measurement of an observable at the end
- In our paper, the input for the VQC were the angle encoded qubits that came from the classical input data.



$U_{enc}(x_i)$
$U_{VQC}(\theta_i)$
$\times L_{block}$
H
Rz
Ry
U_3
x_1
x_2
x_3
x_4
x_5
x_6
x_7
x_8
x_9
x_{10}
x_{11}
x_{12}
θ_1
θ_2
θ_3
θ_4

WHAT THE PROPOSED MODEL DOES (HIGH LEVEL)

- The paper suggests that:
  - The quantum block captures channel relationships better than the classical block
  - Higher classification accuracy result achieved, especially on RGB images
  - More quantum layers improves accuracy without greatly increasing model size



PROJECT GOALS

- Reimplemented QAE-Net using Python + PennyLane
- Train our model using MNIST, F-MNIST, CIFAR-10 datasets using PyTorch
- Compared results with the original paper - Find flaws/differences
- Evaluate whether quantum excitation meaningfully improves performance

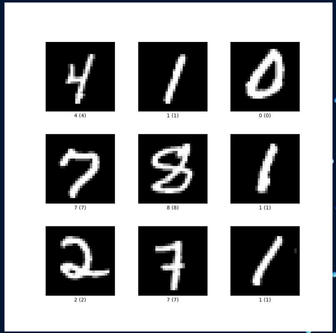
PyTorch + Pennylane

ADDITIONAL INFO: TRAINING DATASETS

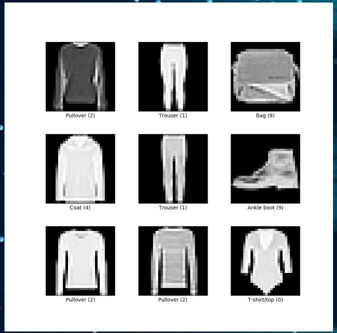
- CIFAR-10
  - Colour
  - Images of animals, vehicles
- MNIST
  - Black & White
  - Images of numbers
- F-MNIST
  - Black & White
  - Images of Clothing

CIFAR-10

MNIST

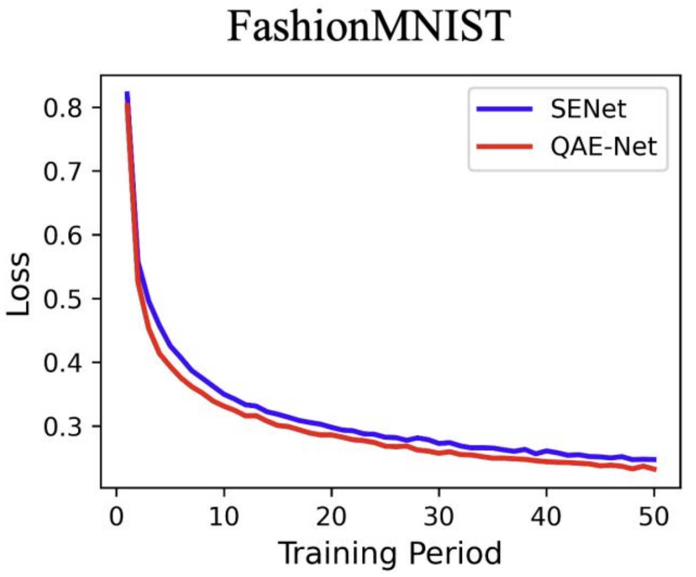


F-MNIST



ADDITIONAL INFO: READING TRAINING GRAPHS

- Loss
  - How accurate are predictions
- Epoch
  - Weights adjusted after each batch
  - A run through the whole dataset is an Epoch



02

CODE
DEMONSTRATIO
N

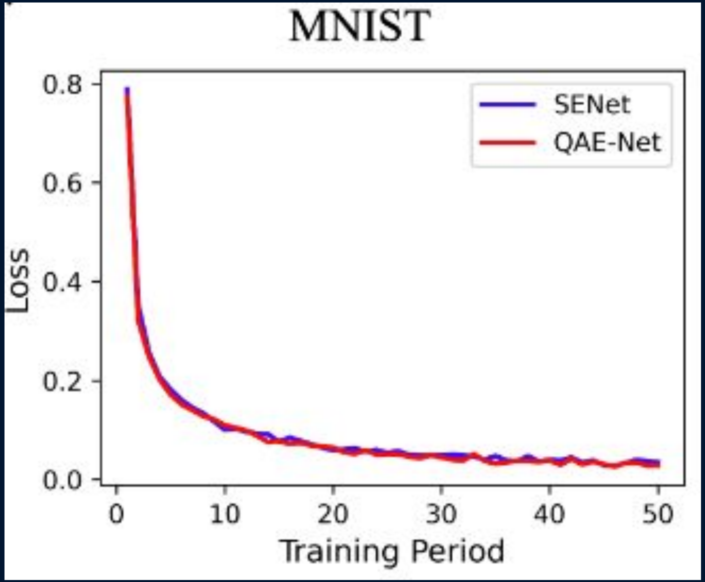
03 RESULTS COMPARISON

MULTI-CLASS CLASSIFICATION RESULTS

Dataset	Method	Ch.	Epoch	Paper Acc. (%)	Our Acc. (%)
MNIST	SENet	1	50	97.9	99.09
	QAE-Net	1	50	**98.0**	**99.21**
F-MNIST	SENet	1	50	91.0	89.85
	QAE-Net	1	50	**91.3**	**90.33**
CIFAR-10	SENet	3	200	76.72	**66.32**
	QAE-Net	3	200	**89.08**	66.22

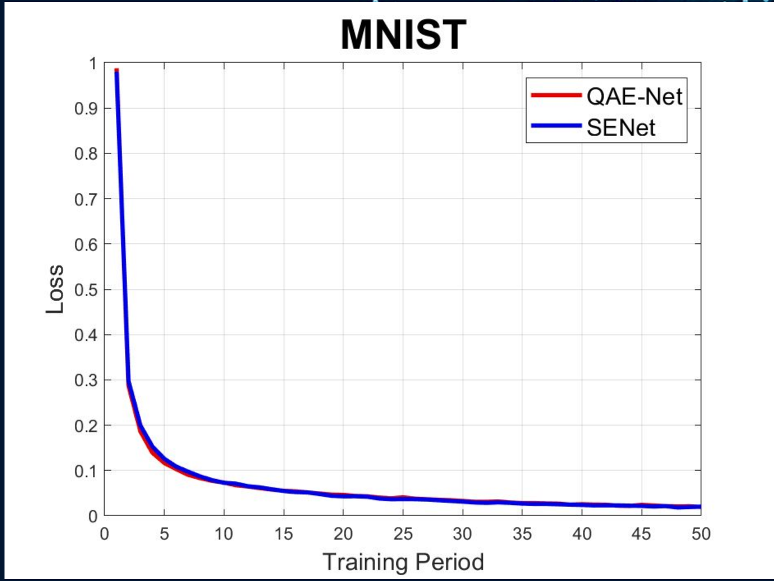
MNIST

Paper



SENet Accuracy: 97.9%
QAE-Net Accuracy: **98.0%**

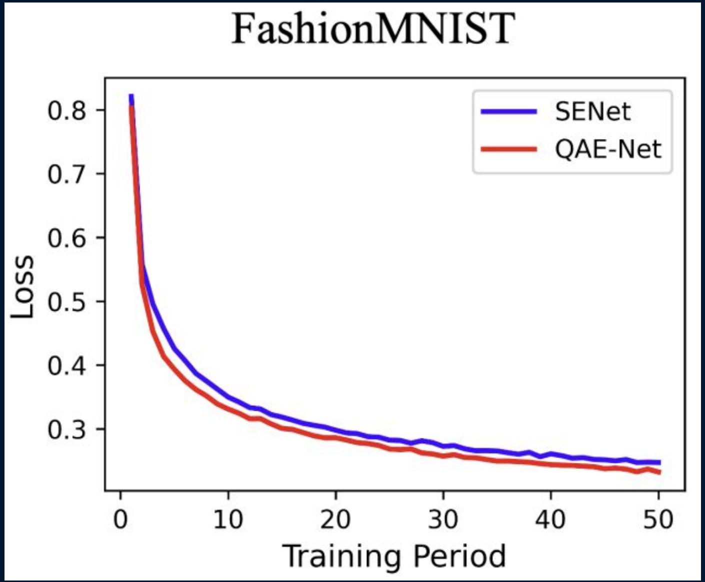
Ours



SENet Accuracy: 99.09%
QAE-Net Accuracy: **99.21%**

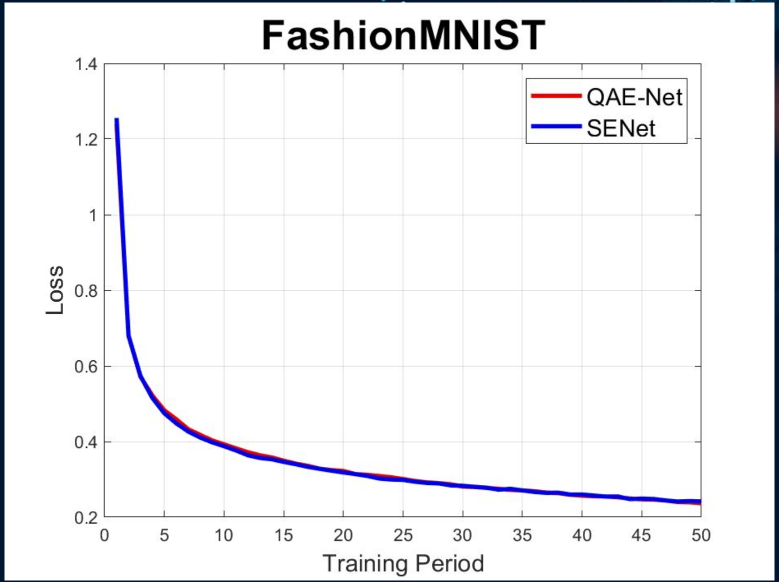
FashionMNIST

Paper



SENet Accuracy: 91.0%
QAE-Net Accuracy: **91.3%**

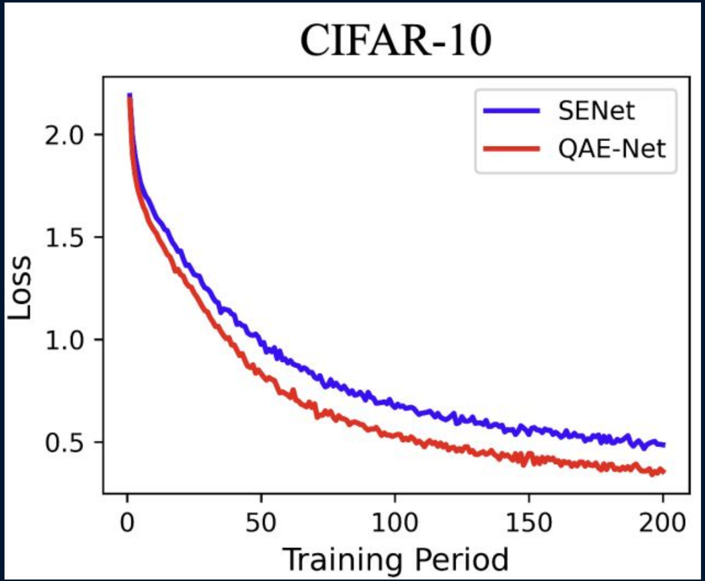
Ours



SENet Accuracy: 89.85%
QAE-Net Accuracy: **90.33%**

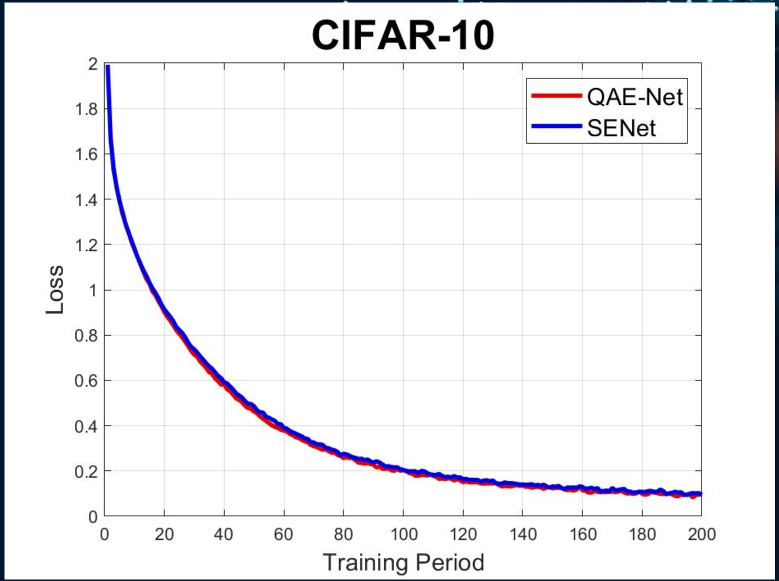
CIFAR-10

Paper



SENet Accuracy: 76.72%
QAE-Net Accuracy: **89.08%**

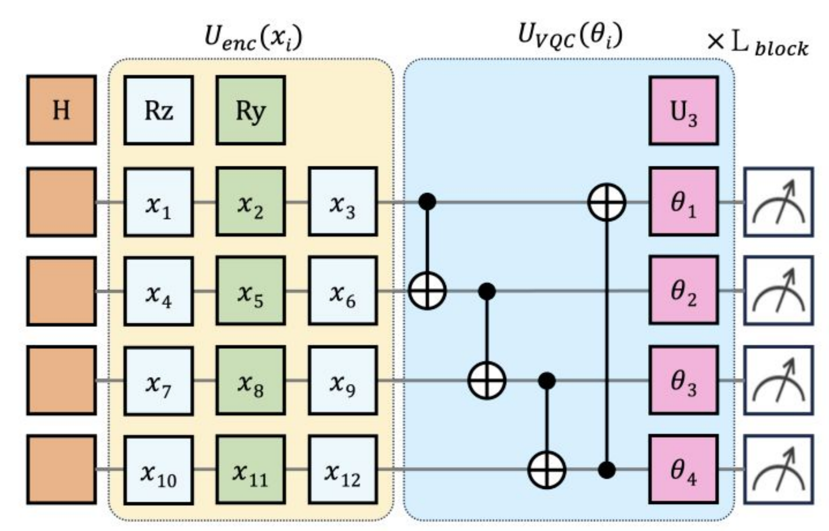
Ours



SENet Accuracy: **66.32%**
QAE-Net Accuracy: 66.22%

VARIATIONAL LAYERS ON CIFAR-10

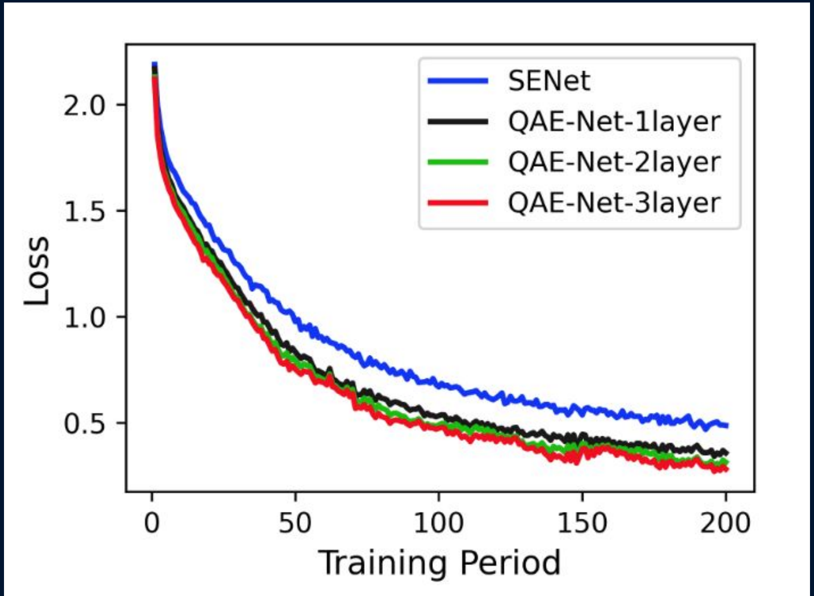
- Test conditions
  - Learning rate: 0.001
  - 200 Epochs
  - 4 Qubits for QAE-Net



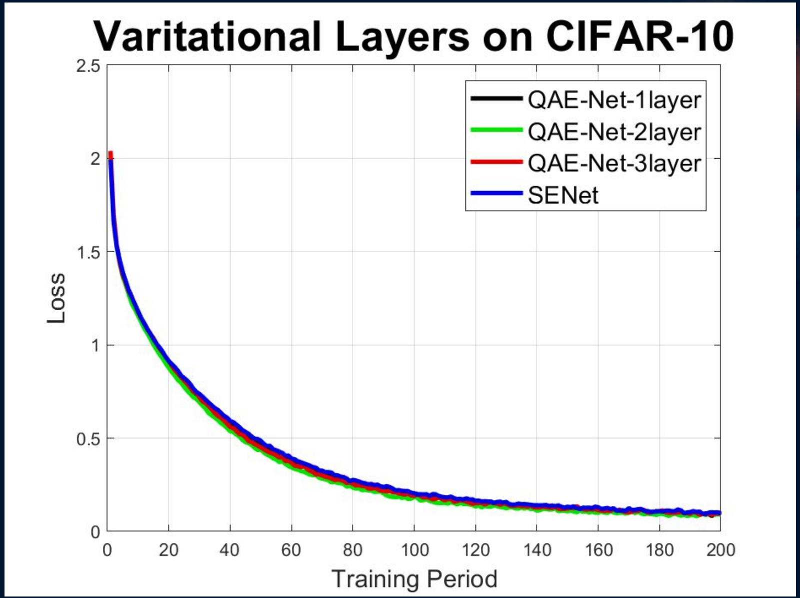
Model	SENet	QAE-Net (1 Layer)	QAE-Net (2 Layer)	QAE-Net (3 Layer)
Our Acc. %	66.32	66.22	66.76	64.30
Paper Acc. %	76.72	89.08	90.10	92.30

VARIATIONAL LAYERS ON CIFAR-10

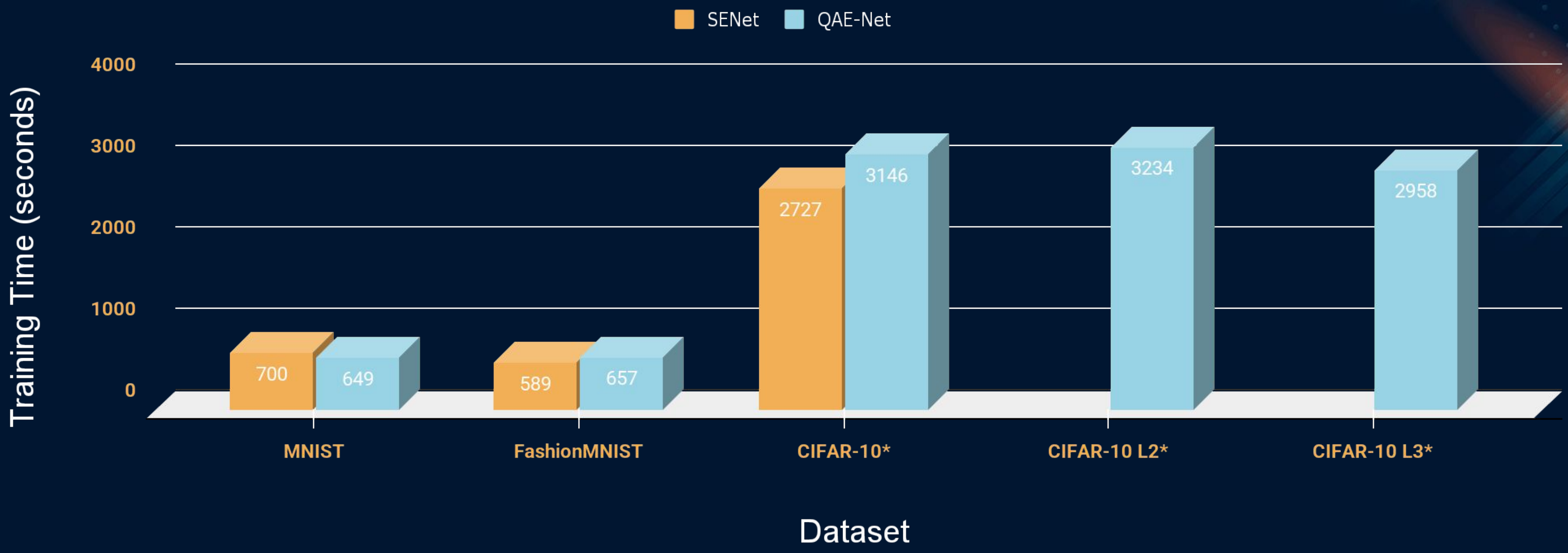
Paper



Ours



TRAINING TIME



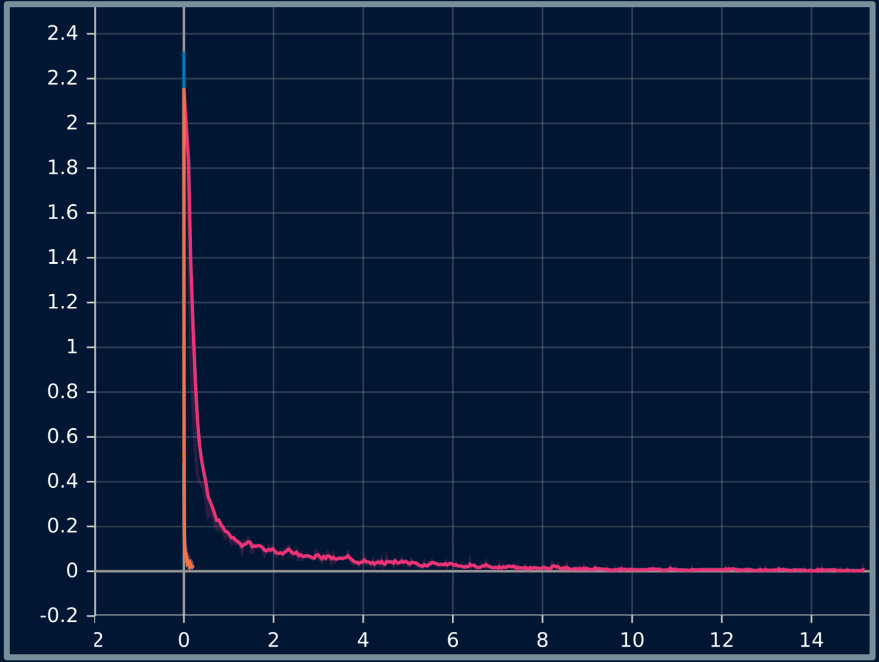
*Note that CIFAR Runs for 200 Epochs, instead of 50

04

REPRODUCIBILITY CHALLENGES

TRAINING SPEED

- Speed with the paper specifications was around ~15 hours for MNIST
  - ~50 hours for each CIFAR-10 model.
- We made some adjustments



HOW DID WE INCREASE TRAINING SPEED?

- Paper specifies "parameter-shift" differentiation method
    - This is how gradients are calculated on actual quantum hardware
- Batched parameters
- Paper specifies "lightning.gpu" as the device
    - We used "default.qubit"

AMBIGUITY IN CLASSICAL MODEL

- Many implied dimensions when configuring the Neural Network
    - Ex: linear layer dimensions in Excitation block
- Relationship between layers and their functions not explicitly defined
    - Activation
    - Pooling Layers

Model	Dataset	Paper Params	Our Params
SENet	(F)-MNIST	39,602	39,602
	CiFAR-10	142,634	142,602
QAE-Net	(F)-MNIST	39,366	39,386
	CiFAR-10	142,570	142,386

05 | FINAL COMMENTS

RECALL: WHAT PROPOSED MODEL DOES (HIGH LEVEL)

The paper suggests that:

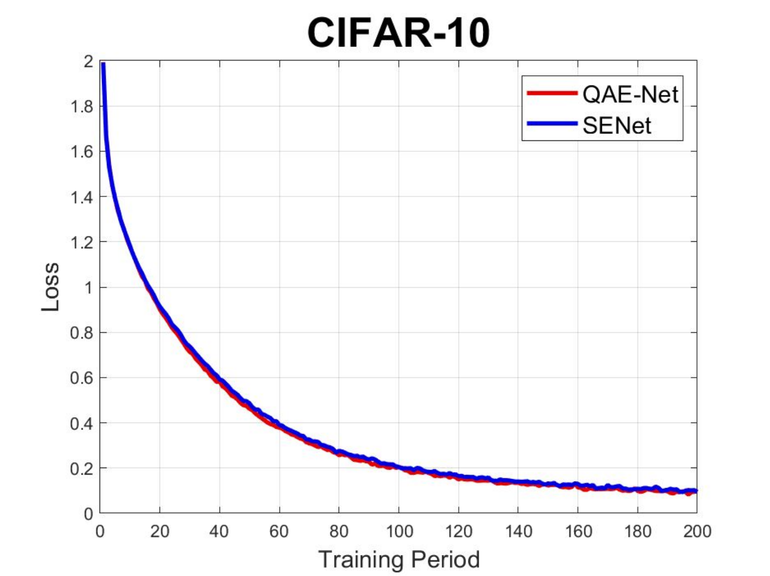
a. The quantum block captures channel relationships better than the classical block
b. Higher classification accuracy result achieved, especially on RGB images
c. More quantum layers improves accuracy without greatly increasing model size

We found that:

a. No evidence the quantum block captures channel relationships any better
b. Higher classification accuracy result only achieved for black & white images, small gains
c. More quantum layers worsened accuracy (2 layers was best)

FUTURE STUDIES

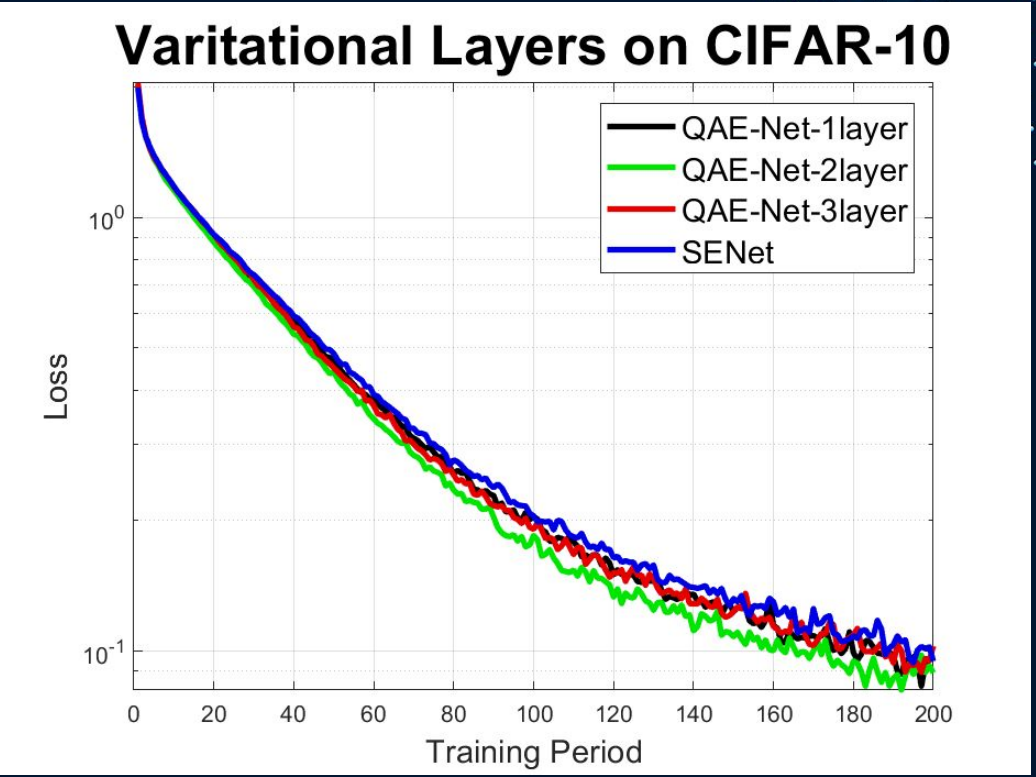
- QAE-Net Performs best on colour images
  - Try training on a different dataset
- Does Increased number of Epochs allow SEN-Net to compete with QAE in all cases?
- For most accurate results (0.1%), it would be ideal to train each model multiple times and take the average
- Are more quantum layers better?

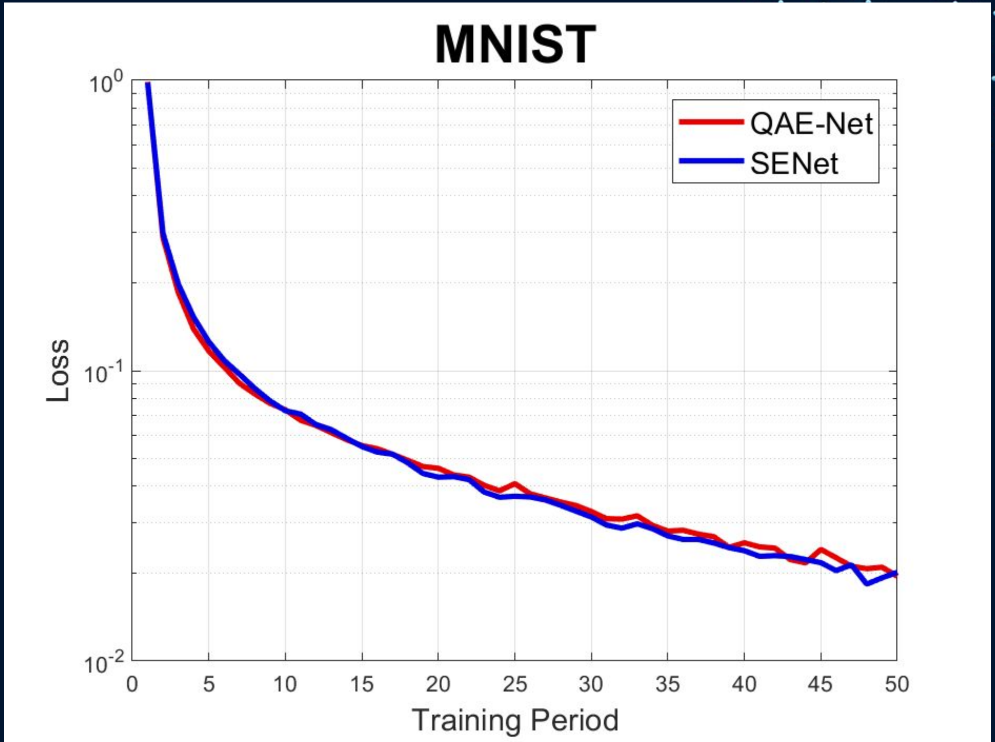


2.1 | RETURN: CODE DEMONSTRATIO N

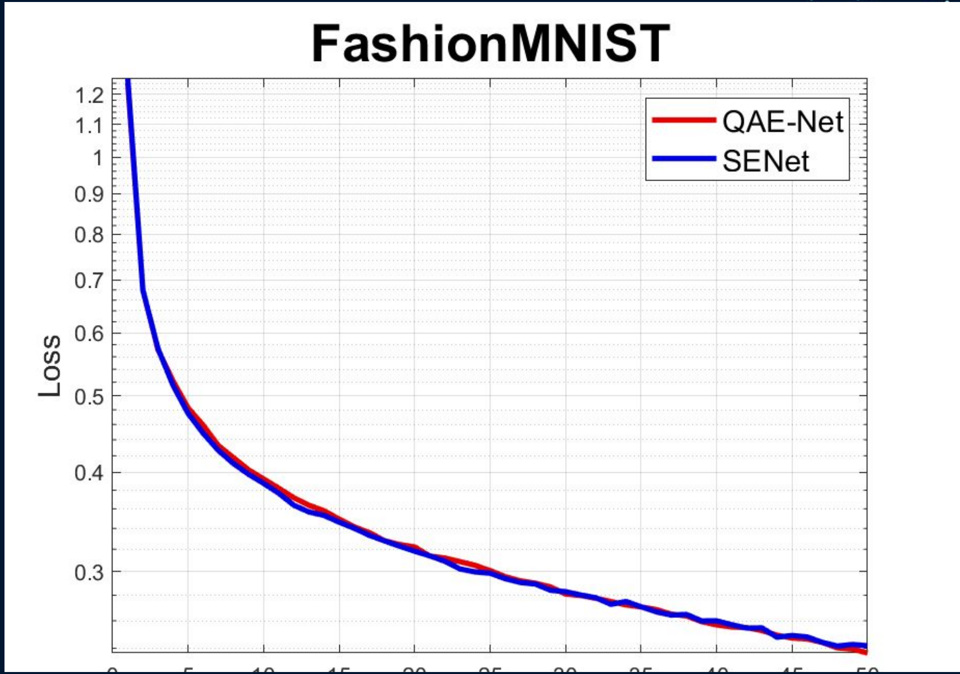
THANK YOU

Appendix A:
Log Plots

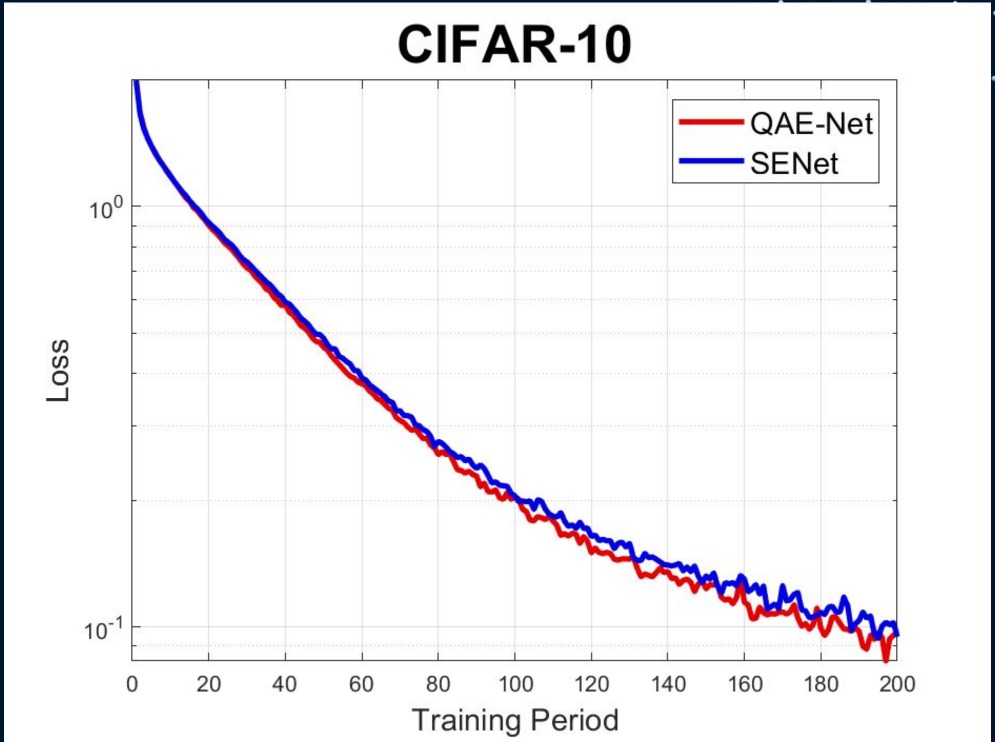

Varitational Layers on CIFAR-10
QAE-Net-1layer
QAE-Net-2layer
QAE-Net-3layer
SENet
10^0
10^{-1}
Loss
0
20
40
60
80
100
120
140
160
180
200
Training Period



MNIST
QAE-Net
SENet
Loss
Training Period
10^0
10^{-1}
10^{-2}
0
5
10
15
20
25
30
35
40
45
50



FashionMNIST
Loss
1.2
1.1
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
QAE-Net
SENet


CIFAR-10
QAE-Net
SENet
Loss
10^0
10^{-1}
0
20
40
60
80
100
120
140
160
180
200
Training Period