Pion/Kaon Identification at STCF DTOF Based on CNN/QCNN

Zhipeng Yao, Teng Li, Xingtao Huang

Shandong University

2023.8.13

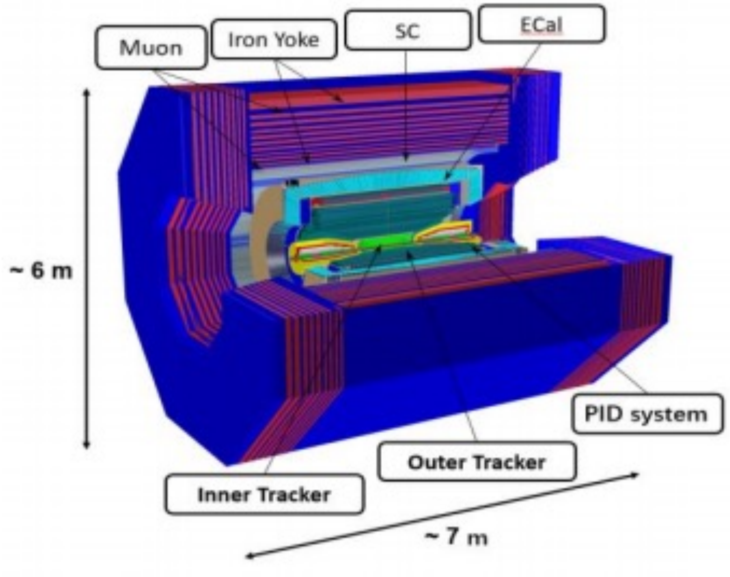
Outline

- Super Tau Charm Facility
- DIRC-like time-of-flight detector
- PID Based on Convolutional Neural Network
- PID Based on Quantum Convolutional Neural Network
- Summary

Super tau-Charm Facility

The Super Tau Charm Facility (STCF) proposed in China is a new-generation facility of electron positron collider

- the peak luminosity above 0.5×10^{35} cm^{-2}s^{-1}
- center-of-mass energies covering 2-7 GeV
- potential for further upgrading to improve the peak luminosity and realize beam polarization in the future



Broad Physics at tau-charm Energy Region

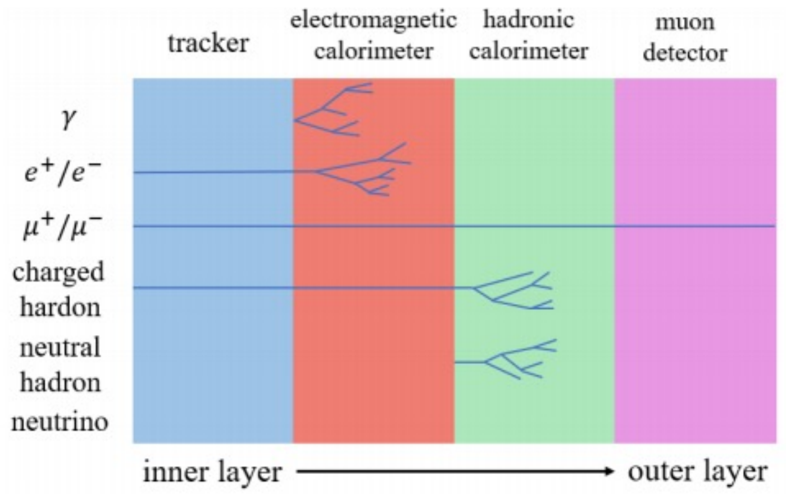
- Rich of physics with c quark and τ leptons
- Important playground for study of QCD, hadron physics
- Search for new physics beyond the Standard Model

From the interaction point outward, the STCF detector consists of a tracking system(**ITK** and **MDC**), a particle identification (**PID**) system, an electromagnetic calorimeter (**EMC**), a superconducting solenoid (**SCS**) and a muon detector (**MUD**).

Particle identification System

The PID is one of the most fundamental tools in various physics studies. The PID for the full momentum range is essential for charm physics studies and fragmentation function studies.

- The identification of hadrons in the low momentum range is achieved through measurements of the specific energy loss rate (**dE/dx**) by the **MDC**.
- The identification of leptons and neutral particles is provided by the **EMC** and the **MUD**.
- To enhance PID and charged hadrons in the high momentum range, the PID system of the STCF is designed



The PID system uses two different Cherenkov detector technologies:

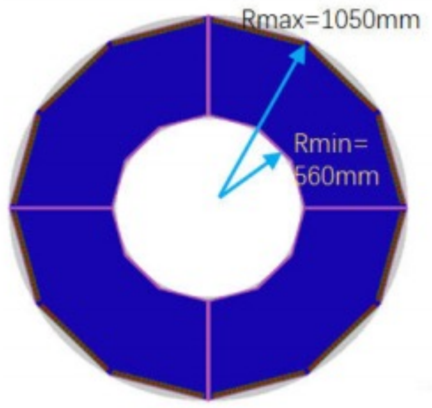
- a Ringing Imaging Cherenkov detector (**RICH**) in the barrel
- a time-of-flight detector based on the detection of the internal total-reflected Cherenkov light (**DTOF**) in the endcap

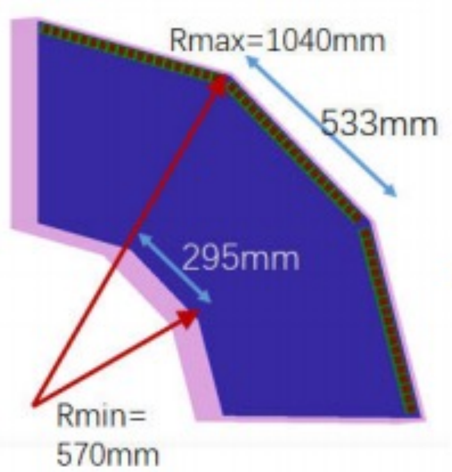
to achieve a 3σ separation between kaons and pions with a momentum up to 4 GeV/c.

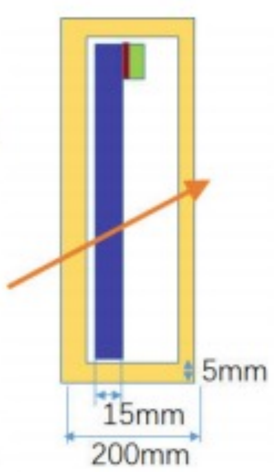
DTOF

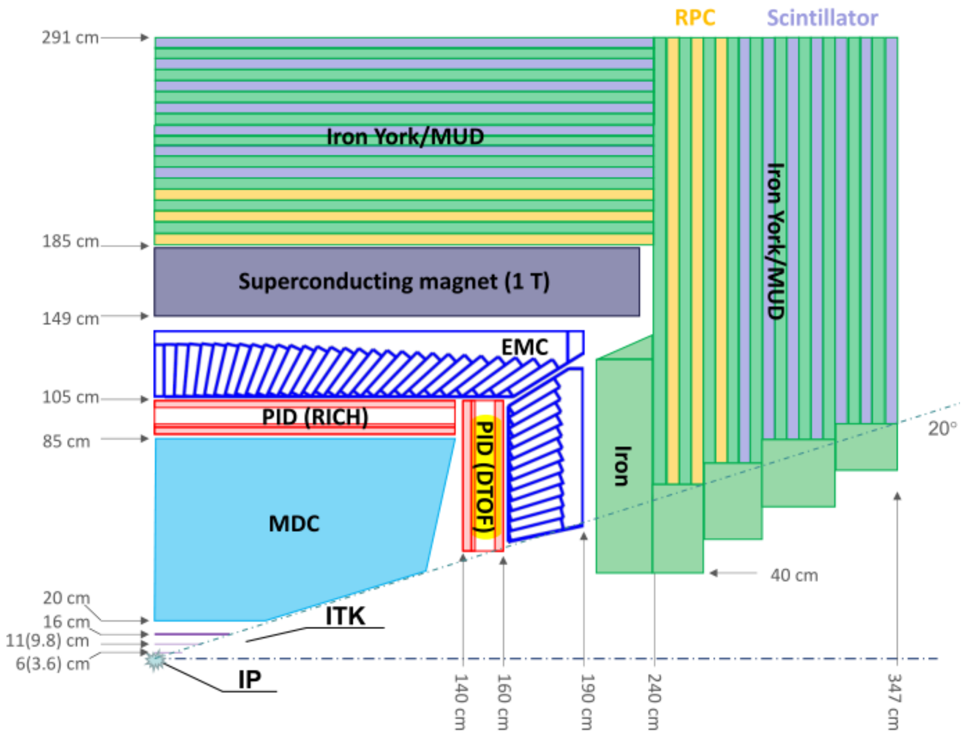
The DTOF consists of two identical endcap discs positioned at $\sim \pm 1400$ mm away from the collision point along the beam direction. Each disc is made up of several quadrantal sectors, with an inner radius of $\sim$ 560 mm and an outer radius of $\sim$ 1050 mm.

- covering in polar angles of $\sim 22^\circ - 36^\circ$
- synthetic fused silica radiator
- Photoelectric Detection: Multi-Anode PMT









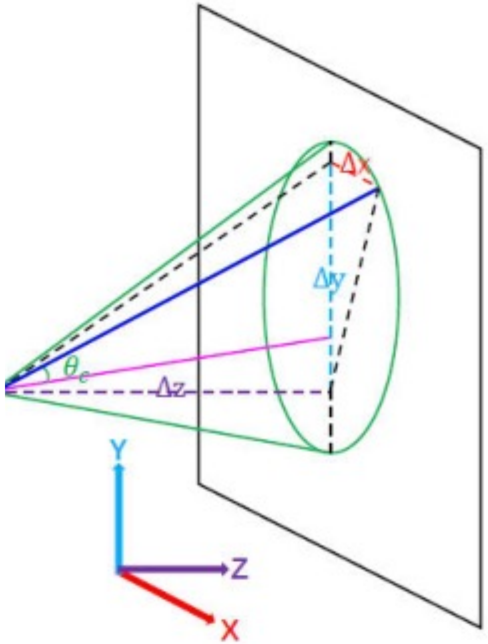
Schematic layout of the STCF detector concept

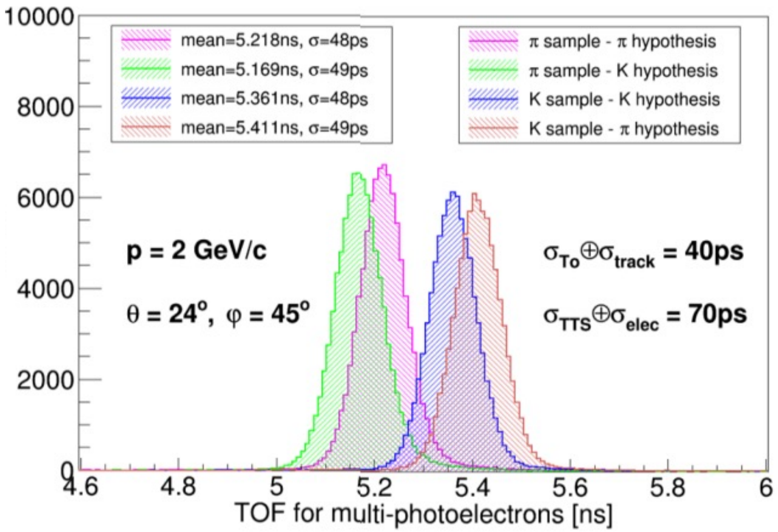
The likelihood method for PID

Building likelihood probability density function based on reconstructed TOF distribution

- $cos(\overline{\theta_c}) = \frac{1}{n_p \beta} = \frac{\vec{v_t} \cdot \vec{v_p}}{|\vec{v_t}| \cdot |\vec{v_p}|}$
  - $\vec{v_t} = (a, b, c), \ \ |\vec{v_t}| = 1$
  - $\vec{v_p} = (\Delta x, \Delta y, \Delta z)$
- $LOP = H \cdot \sqrt{\frac{\Delta x^2 + \Delta y^2 + \Delta z^2}{\Delta y^2}}$
- $TOF = T - TOP - T_0 = T - \frac{LOP \overline{n_g}}{c} - T_0$

$$\mathcal{L}_h = p_h(N_{p.e.}) \prod_{i=0}^{N_{p.e.}} f_h(TOF_i)$$



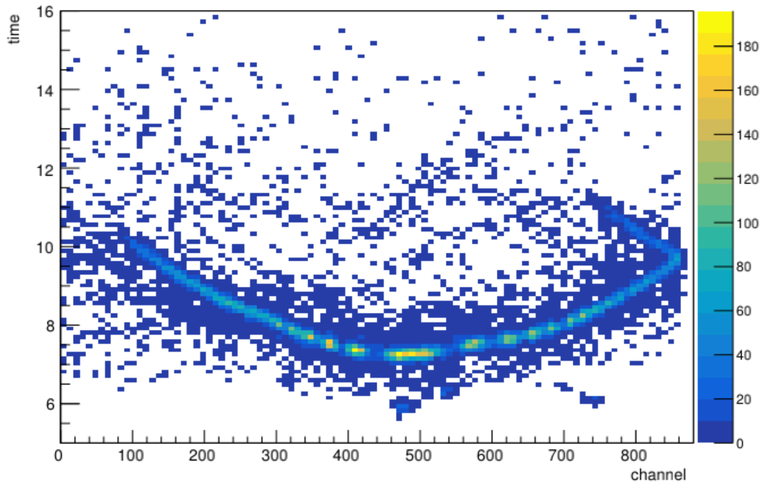


Qi, B et al., DIRC-like time-of-flight detector for the experiment at the Super Tau-Charm Facility. *Journal of Instrumentation*, *16*(08), P08021.

Convolutional Neural Network

The likelihood method

- Utilizes the timing information of different particle hypotheses
- But ignores spatial information (topology of photons)



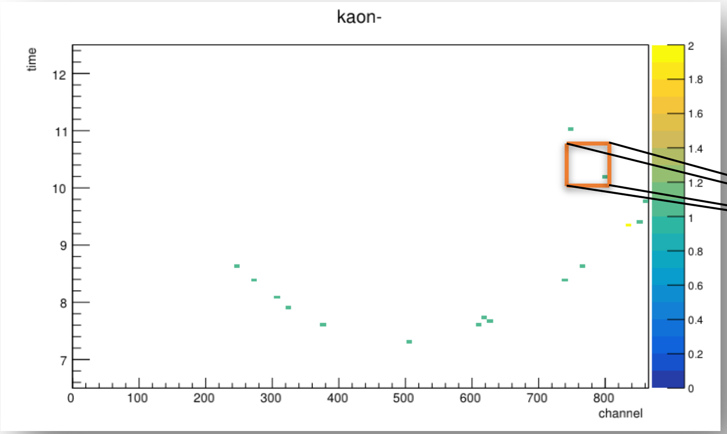
The pixel map of photons:

- X-label: the hit position of Cherenkov collected photon by PMT
- Y-label: the arrival time of Cherenkov collected photon by PMT
- Value: the number of photons within this bin

The image-like data represents the topologies of Cherenkov photons generated by different particles

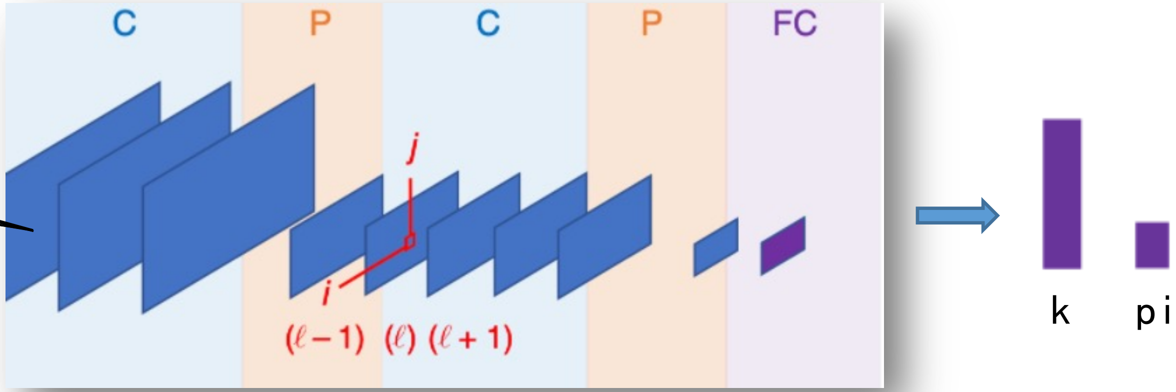
To exploit the PID performance of DTOF, we developed a convolutional neural network (**CNN**) for pions/kaons identification.

The Structure of CNN



single event

convolution



CNN consists of interlaced convolutional layers and pooling layers, and ends with a fully connected layer.

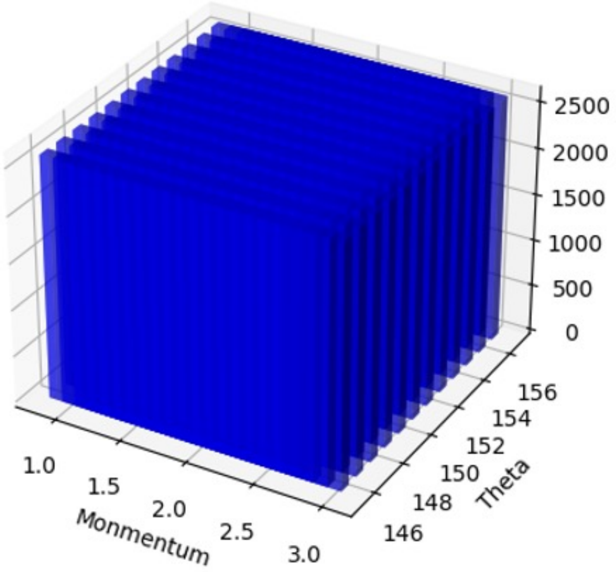
- The primary purpose of the convolution layer is to extract new hidden features using convolution kernels
- The pool layer is used to reduce the dimension of data, reducing the resources required for learning and avoiding overfitting
- The full connection layer adopts softmax full connection, and the activation value obtained is the picture feature extracted by convolutional neural network.

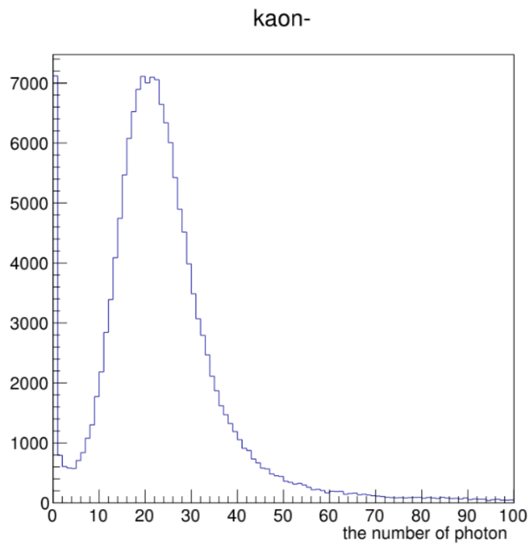
data sample

MC sample is produced with the Offline Software of Super Tau-Charm Facility (OSCAR)

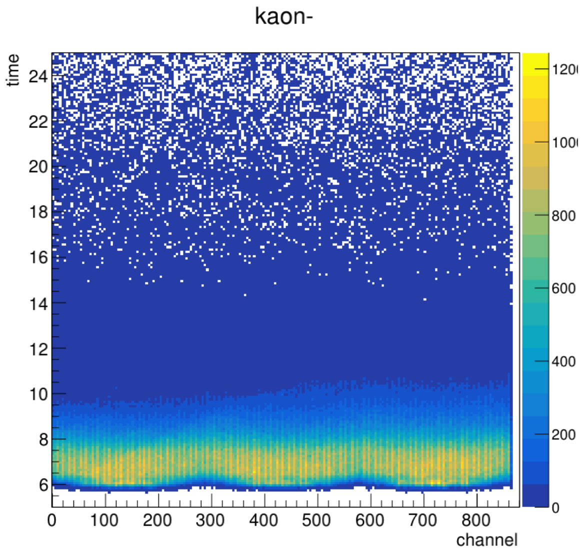
- pi+ : pi- : k+ : k- = 1 : 1 :1: 1
- 0.6 Gev < p < 2.4 Gev
- 23° < theta < 35°

- $0 \le$ channel ≤ 868
- $5.5 \le$ time ≤ 15.5 ns
- Bin number: channel * time = 217 * 200





the number of photon



2D chnanel-time

The Performance of CNN

The structure and parameters of CNN:

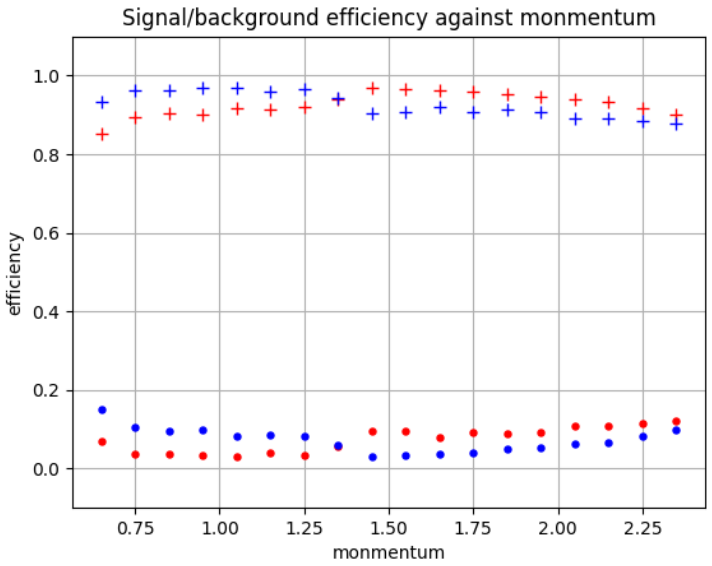
- Conv2D (16, (3, 3), activation='relu'), MaxPooling2D ((2, 2))
- Conv2D (32, (3, 3), activation='relu'), MaxPooling2D ((2, 2))
- Conv2D (32, (3, 3), activation='relu'), MaxPooling2D ((2, 2))
- Flatten(), Input((p, theta)), Concatenate()([model.output, input_features])
- Dense(256, activation='relu'), Dropout(0.2), Dense(2)
- learning_rate = 0.00001, batch_size = 128

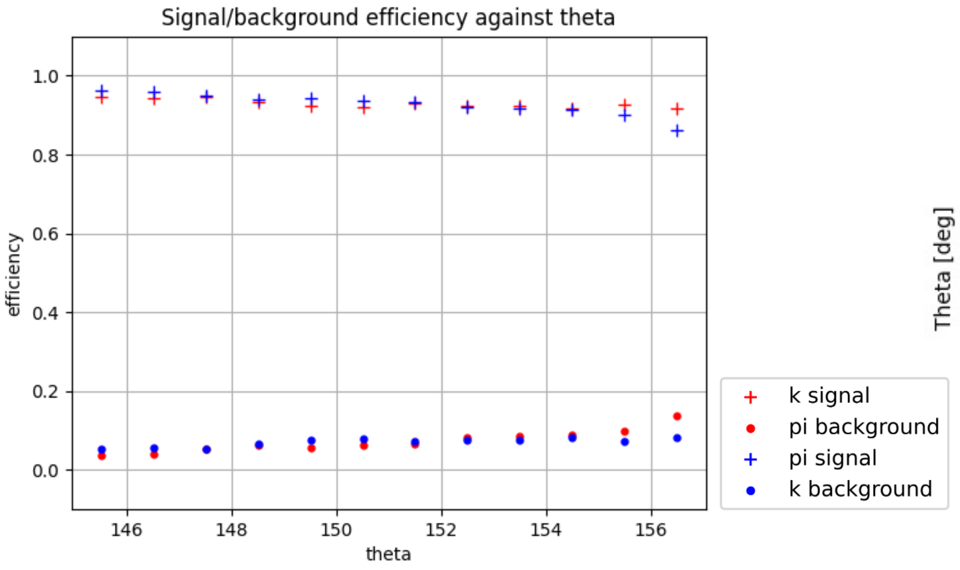
Data set:

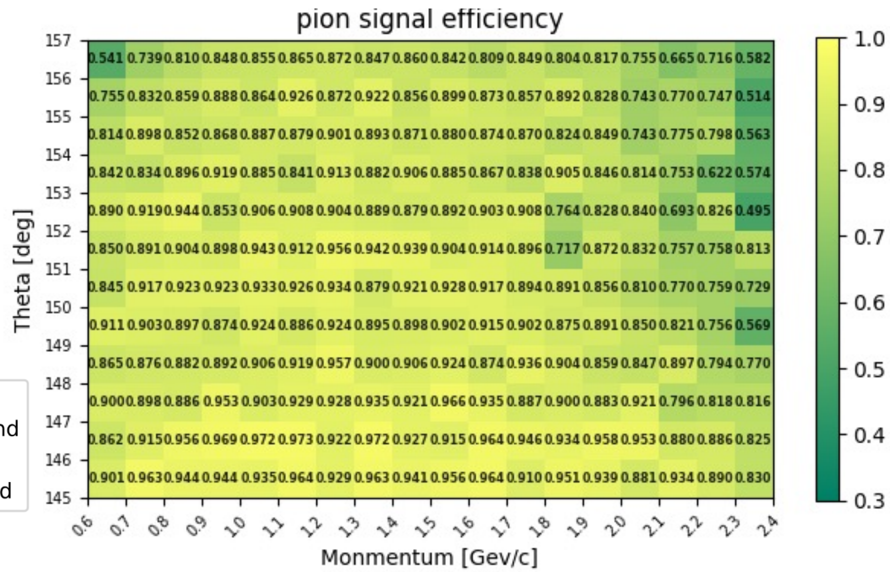
- training set: 200k
- validation set :100k
- test set : 100k

Test set accuracy : **92.86%**

Background efficiency not exceeding 3%







The preliminary results show the CNN model has a promising performance against the pion/kaon identification.

Quantum machine learning

Question: can we do better with the help of quantum machine learning?

- Quantum machine learning: under the domain of quantum computing/algorithm
  - Provide alternatives/enhancement for traditional machine learning algorithms
- Potential quantum advantage for ML problems
  - It utilizes high-dimensional Hilbert space through superposition and entanglement to explore more useful information.
- Basic idea: use a quantum device to extract features from the origin image-like data, before feeding data into the CNN
- Based on the classical CNN, a quantum convolution neural network (**QCNN**) is developed as a proof-of-concept work exploring possible quantum advantages provided by quantum machine learning methods.

Quantum Convolutional Neural Network

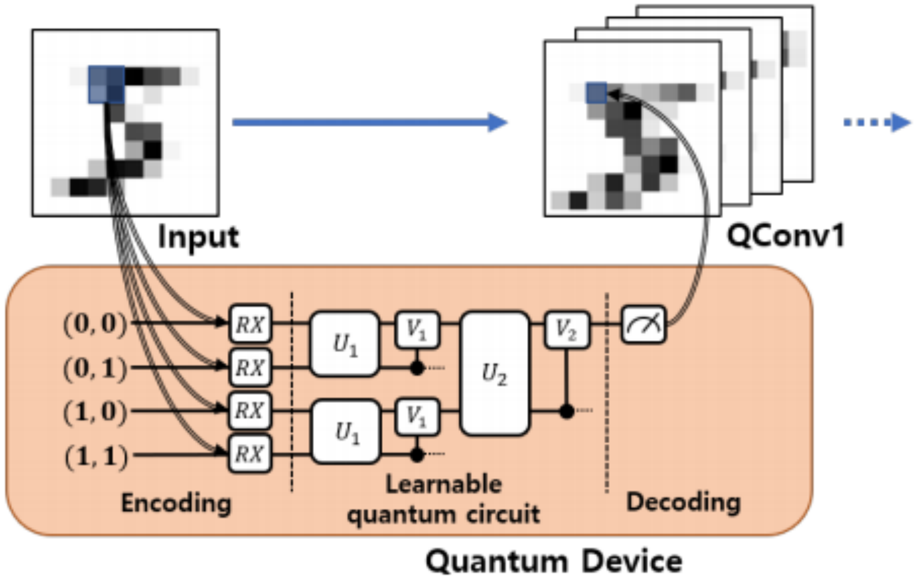
Leveraging the capabilities of the **TensorFlow Quantum** and **Cirq** Simulator platforms, we have developed a trainable quantum convolution layer that can replace the traditional convolution layer in CNN.

1. Data Encoding Circuit

Because existing quantum computers are still limited to small quantum systems, the quantum convolution layer does not apply the entire image map to a quantum system at once, but processes it as much as the filter size at a time.

A small region of the input image, in our work a 2×2 square, is embedded into a quantum circuit. This is achieved with RX rotation gate applied to the qubits initialized in the |0> state.

$$RX(\theta) = \exp\left(-i\frac{\theta}{2}X\right) = \begin{pmatrix} \cos\frac{\theta}{2} & -i\sin\frac{\theta}{2} \\ -i\sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{pmatrix}$$



http://doi.org/10.1109/ICTC49870.2020.9289439

Quantum Convolutional Neural Network

2. Quantum convolutional Kernel

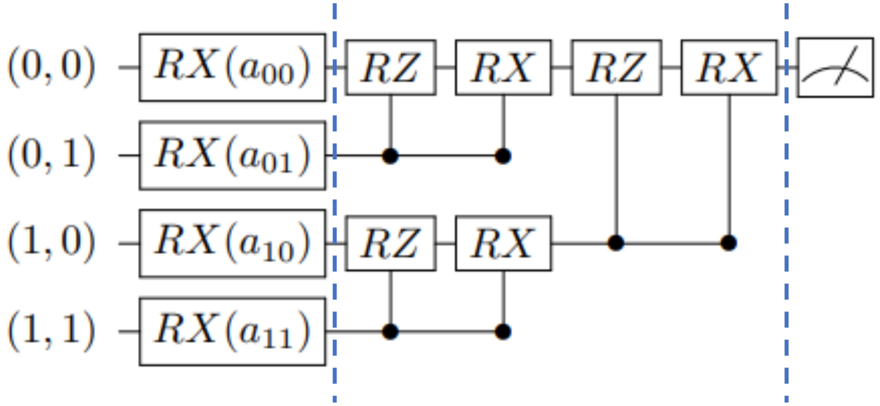
Utilize a parameterized variational quantum circuit to take spatially-local subsections of images from a dataset as input. In our work, we use some entanglement gates with a parameterized phase.

CXPowGate

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & gc & -igs \\ 0 & 0 & -igs & gc \end{bmatrix}$$

CZPowGate

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & e^{i\pi t} \end{bmatrix}$$



In the standard language of CNN, this would correspond to a convolution with a 2×2 *kernel* and a *stride* equal to 1.

3. Decoding

The decoding process gets new classical data by measurement quantum states.

According to the basic principle of quantum mechanics, the expectation value for measuring the observable is deterministic.

$$f(\boldsymbol{\theta}) = \left\langle \psi | W^{\dagger}(\boldsymbol{\theta}) \hat{O} W(\boldsymbol{\theta}) | \psi \right\rangle$$

The quantum convolution can be followed by further classical CNN layers

The Performance of QCNN

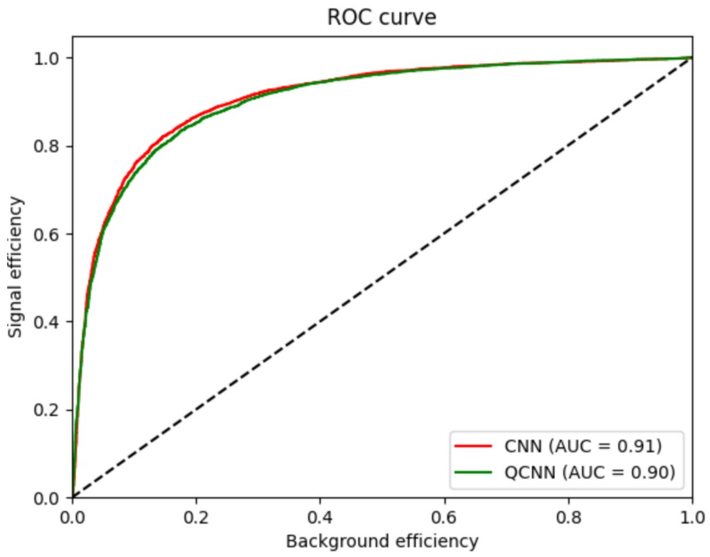
The 217 × 200 size MNIST dataset was downscaled to 32×32 size

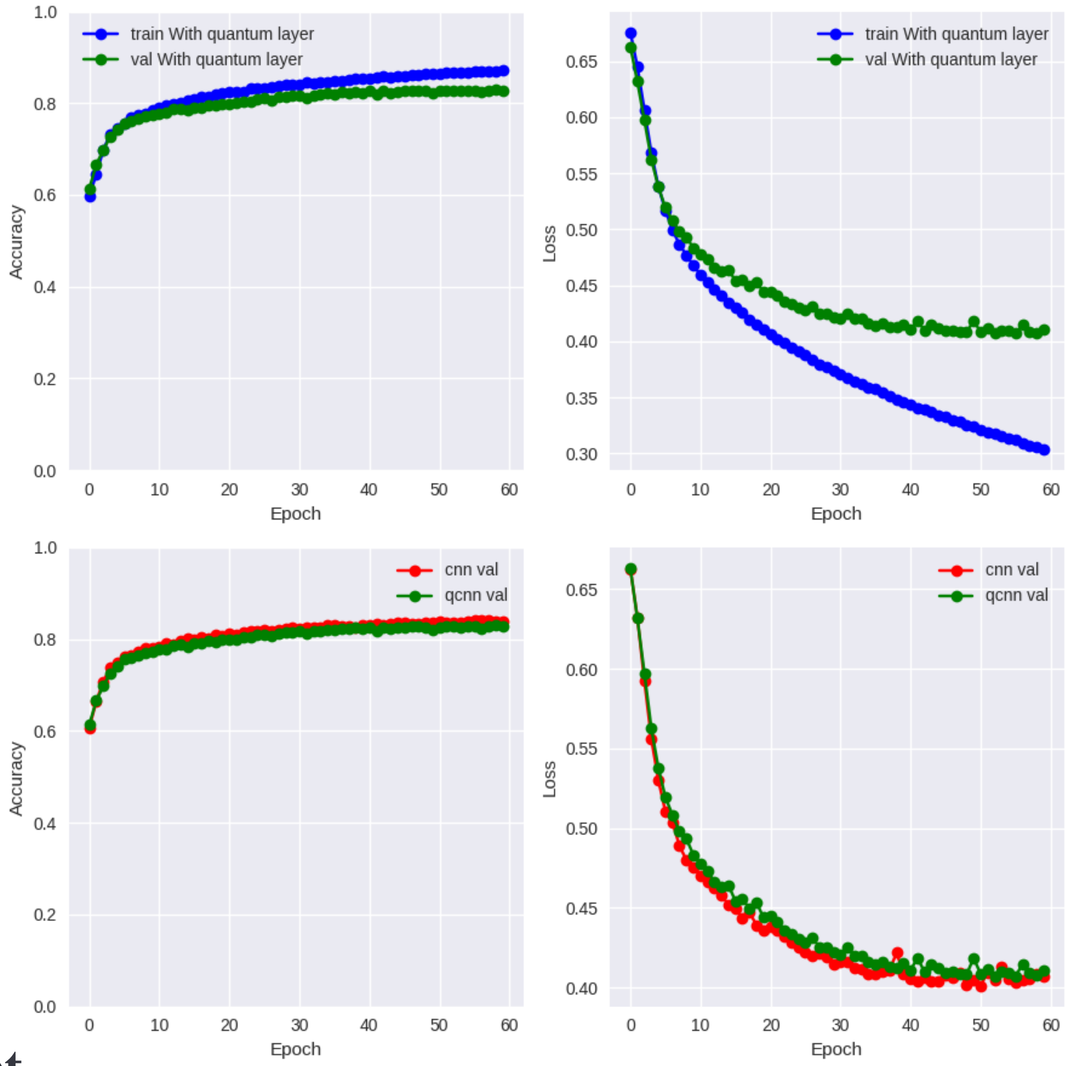
The structure and parameters of QCNN:

- **QCONV**(1, (2, 2), activation=‘relu’), MaxPooling2D ((2, 2))
- Conv2D (16, (2, 2), activation=‘relu’), MaxPooling2D ((2, 2))
- Flatten(), Dense(128, activation=‘relu’), Dense(2)
- learning_rate = 0.0001, batch_size = 16

Data set:

training set: 20k
validation set : 10k
test set : 10k

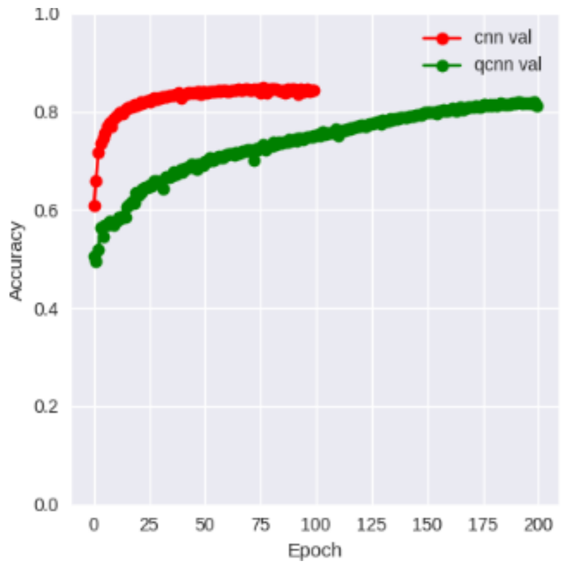


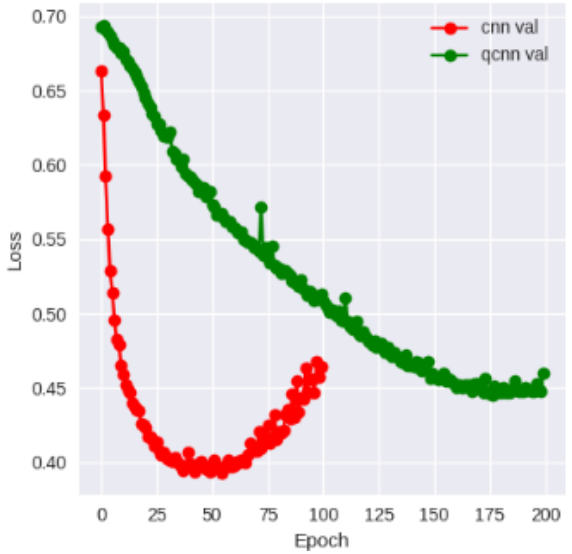


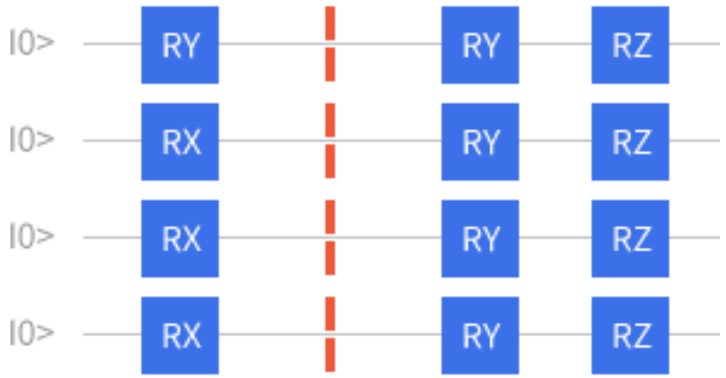
QCNN achieved similar performance to CNN in same dataset.

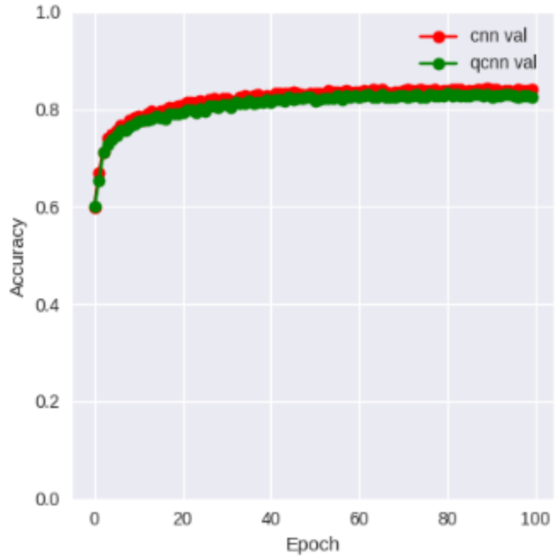
The Performance of QCNN

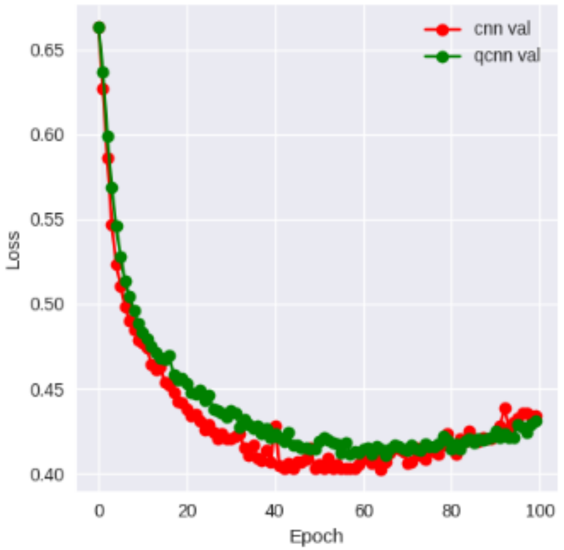
Design of Variational Circuits





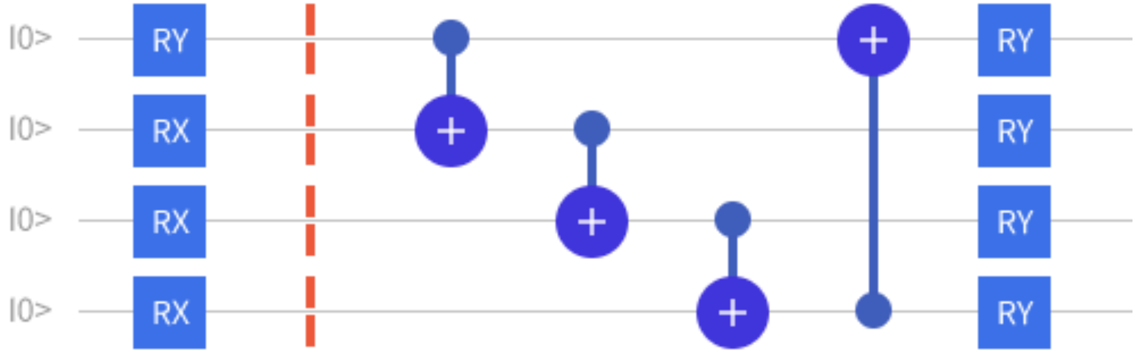


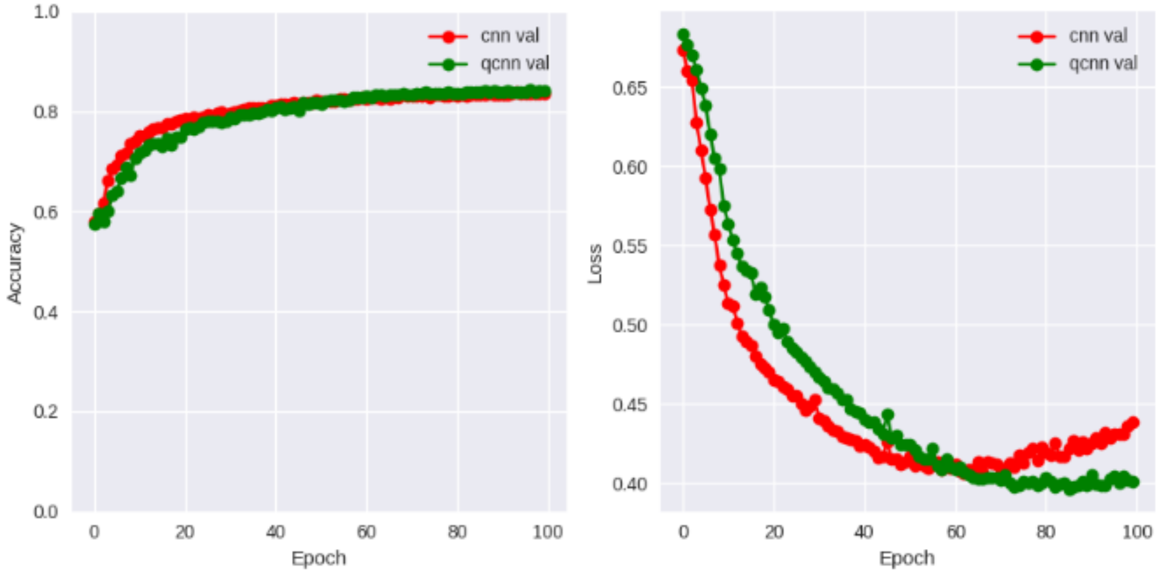


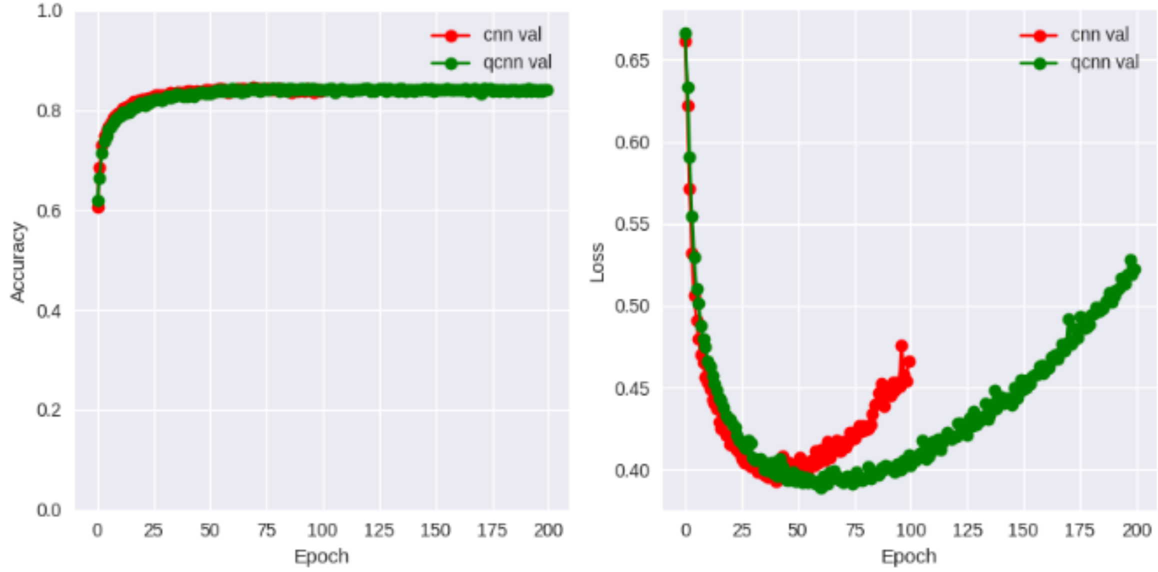


More **parameterized** quantum gates lead to performance improvement

The Performance of QCNN







Entanglement operations also contribute to the performance improvement

Summary

- Targeting at the Pion/Kaon identification problem at STCF, a CNN is developed taking the photon hit positions and photon arrival times as inputs
- The preliminary results show that the CNN model has a promising performance for the Pion/Kaon ID problem

- To explore better performance, a quantum CNN that uses a trainable quantum convolutional kernel are developed.
- The quantum version of CNN acheives similar performance comparing to classical CNN on small datasets.
- Further studies are still in progress, as a proof-of-concept of using QCNN to process HEP experiment data.

Thank you！