SHANDONG UNIVERSITY

27th International Conference on Computing in High-Energy & Nuclear Physics

Particle Identification at STCF DTOF Detector Based on Classical/Quantum Convolutional Neural Network

Zhipeng Yao, Teng Li, Xingtao Huang

Shandong University Institute of Frontier and Interdisciplinary Science

2024.10.21 Krakow Poland

Outline

I. Background

II. π/K Identification Based on Classical Convolutional Neural Networks

III. Quantum Computing and Quantum Machine Learning

IV. Feasibility Study of Quantum Convolutional Neural Networks for PID

V. Summary

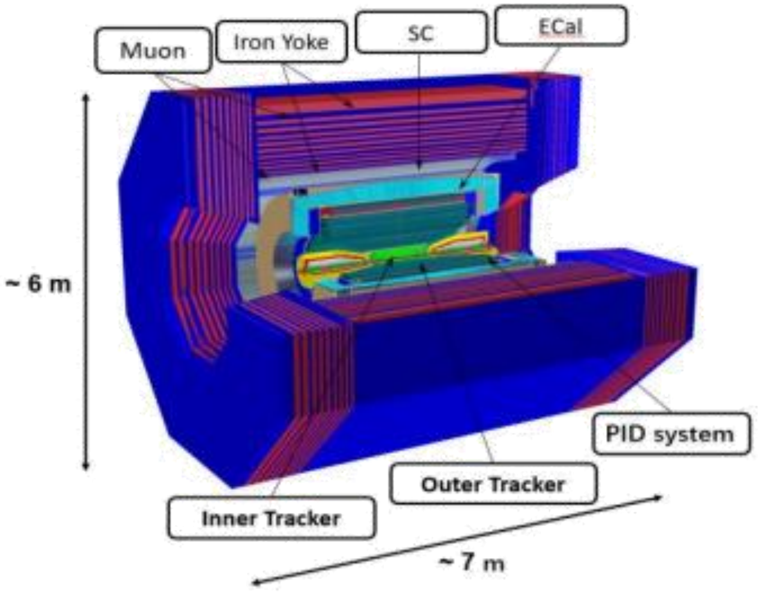
1 Super Tau-Charm Facility

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

- Center-of-Mass Energy 2-7 GeV
- peak luminosity 0.5×10^{35} $cm^{-2}s^{-1}$
- higher-luminosity upgrades and beam polarization in the future

rich physics research

- charm quarks and τ leptons
- Non-perturbative strong interactions and hadronic structure
- Search for new physics



From the interaction point outward:

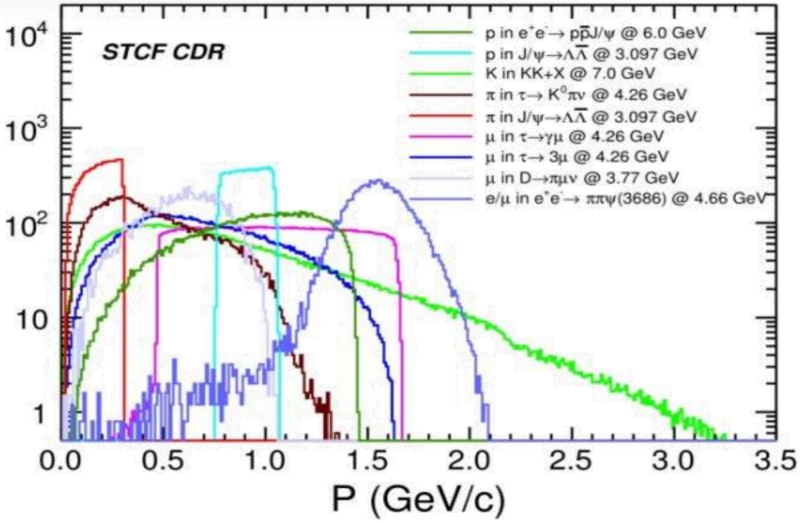
- Tracking system (ITK and MDC)
- Particle identification system (PID)
- Electromagnetic calorimeter (EMC)
- Superconducting solenoid (SCS)
- Muon detector (MUD)

ITk • $< 0.25\% X_0$ / layer • $\sigma_{xy} < 100$ μm	Cylindrical μRWELL CMOS MAPS
MDC • $\sigma_{xy} < 130$ μm • $\sigma_p/p \sim 0.5\%$ @ 1 GeV • dE/dx~6%	Cylindrical Drift chamber
PID • π/K (and K/p) 3-4σ separation up to 2GeV/c	RICH with MPGD DIRC-like TOF
EMC E range: 0.025-3.5GeV σ_E (%) @ 1 GeV Barrel: 2 .5 Endcap: 4 Pos. Res. : 5 mm	pCsI + APD
MUD • 0.4 - 2 GeV • π suppression >30	RPC + scintillator

1 PID Requirements

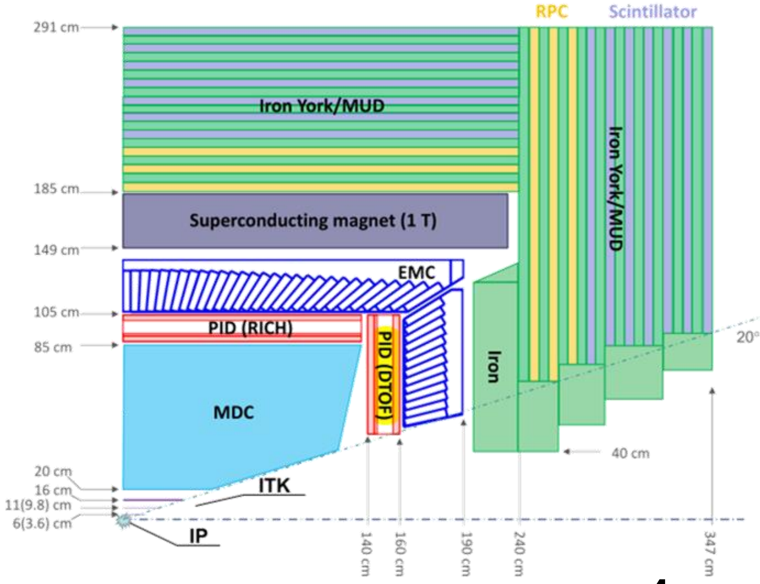
Particle identification (PID) is an important tool for conducting physics research in collider experiments.

- $p < 2$ GeV/c, π/K misidentification rate < 2%, identification efficiency > 97%
- $p > 0.7$GeV/c, μ identification efficiency > 95%; 0.5 < p < 0.7, > 70%
- Good neutral particle identification capability



The **PID system** uses two 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

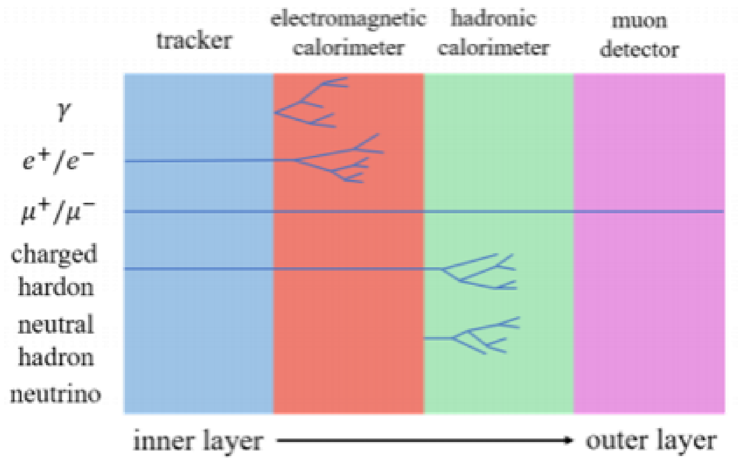


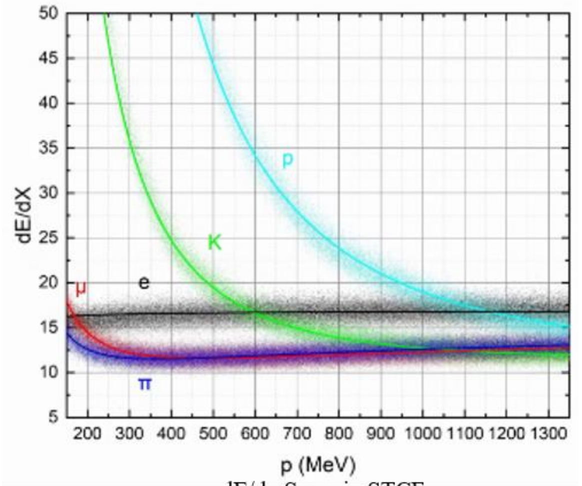
1 Machine learning-based PID Technology

Relying solely on information from a single sub detector makes it challenging to accurately distinguish particles

Machine learning: excellent performance in PID by extracting useful features in high-dimensional spaces

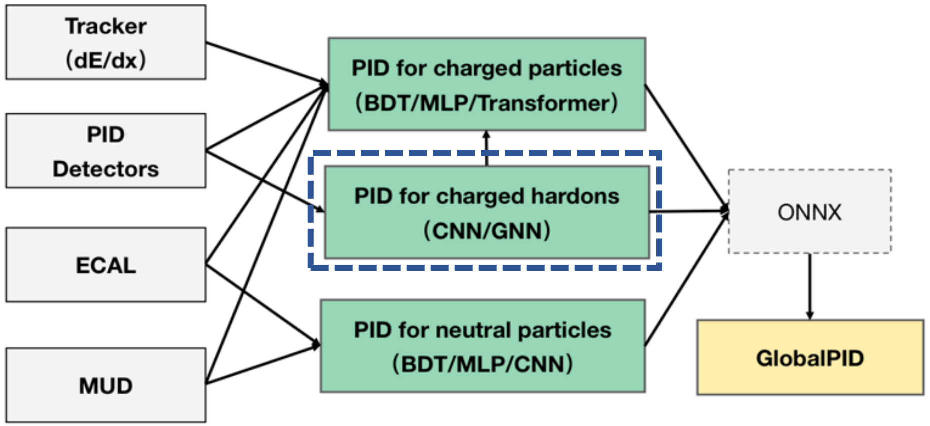
- Combining information from multiple sub detectors
- Fully utilize the original response of the detector
- Decision trees and neural networks have gradually become the mainstream methods of PID





dE/dx Sepa. in STCF

Various PID techniques based on ML algorithms for different types of detectors used in STCF

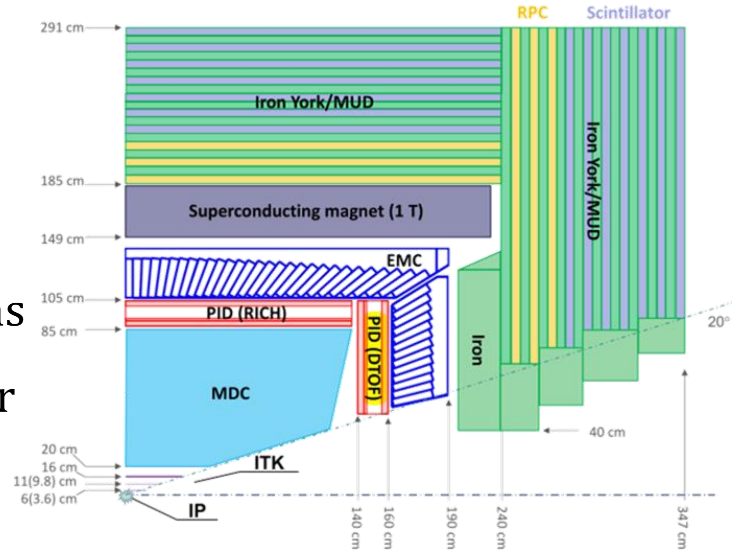


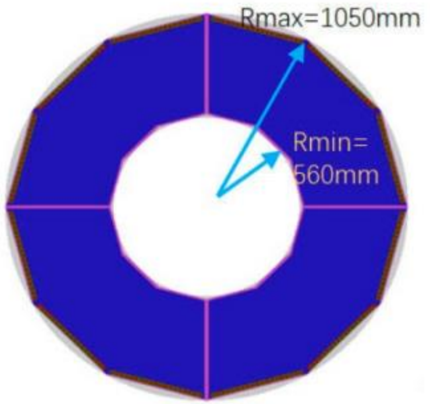
PID algorithms based on machine learning for STCF

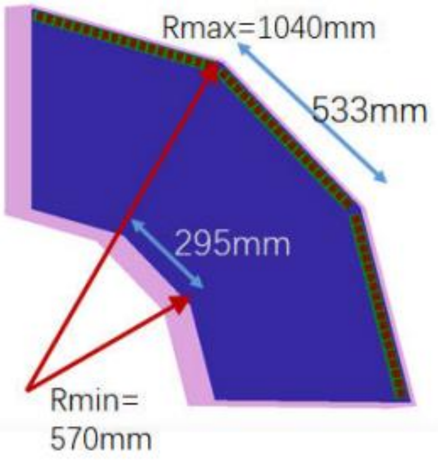
2 DTOF Detector

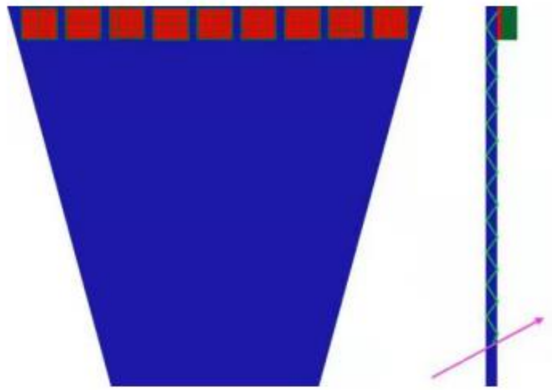
As the endcap PID detector of the STCF, the DTOF employs a technology based on the detection of internally total-reflected Cherenkov light

- DTOF consists of two identical discs, containing multiple sectors
- covering in polar angles of ~23-36°
- synthetic fused silica serves as the Cherenkov radiator to generate photons
- an array of MCP-PMTs are optically coupled to the radiator along the outer side to detects the Cherenkov photons



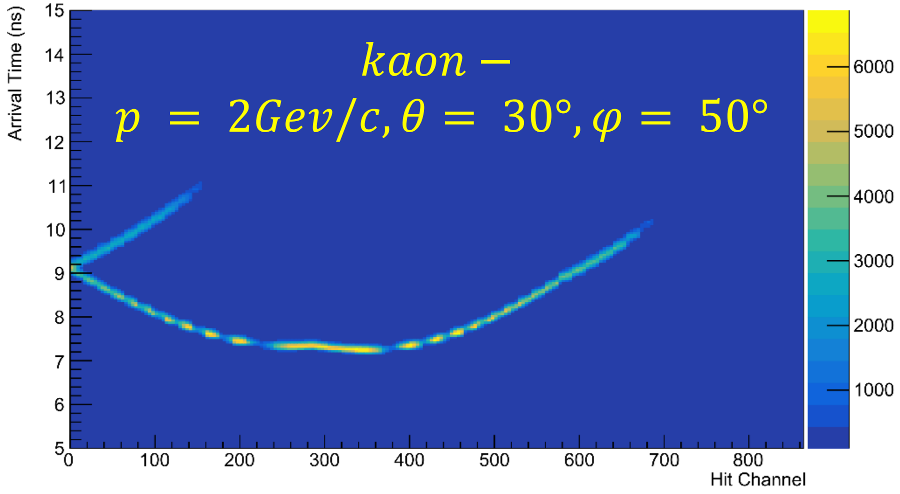


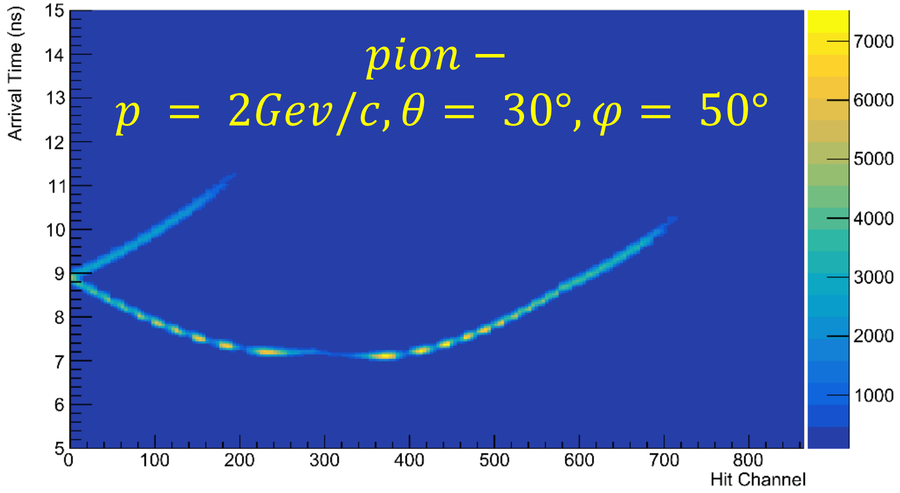




Two-dimensional pixel map

Combine the time-space information of Cherenkov photons hitting the PMT





Using the original response from the detector, construct a two-dimensional pixel map:

- X-label: Hit channels of Cherenkov photons received by the PMT
- Y-label: Arrival time of Cherenkov photons received by the PMT
- Value: Number of photons in the bin

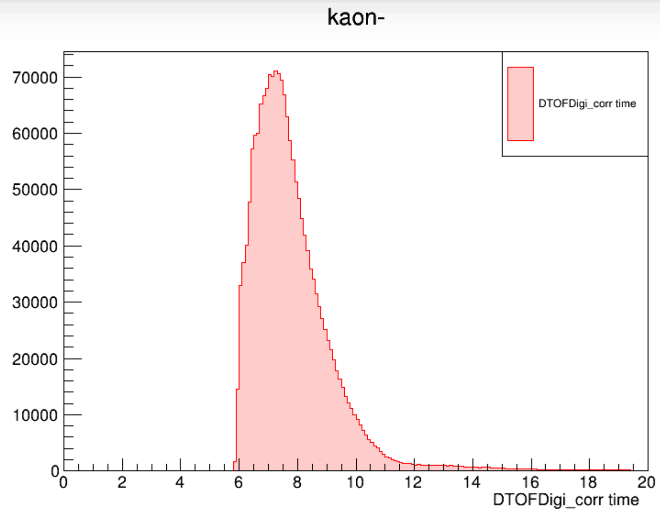
A larger pixel area indicates that at the current time, photons have a greater probability of hitting the corresponding channel

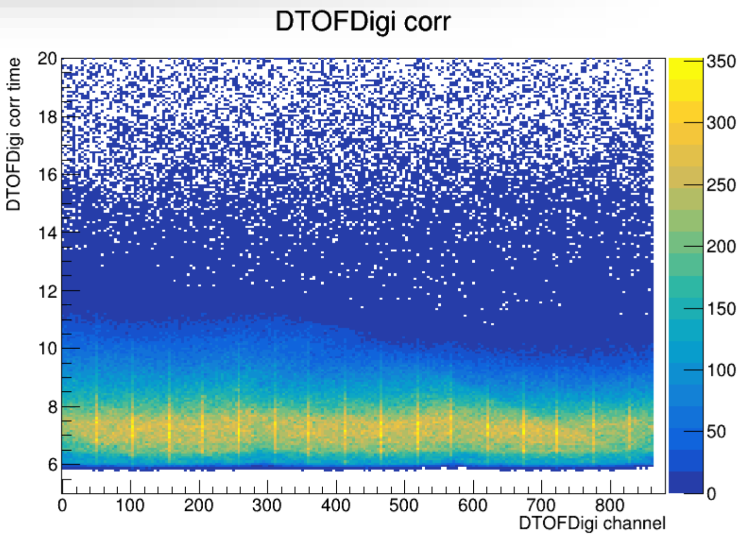
CNN Algorithm is a suitable choice

2 Data Sample

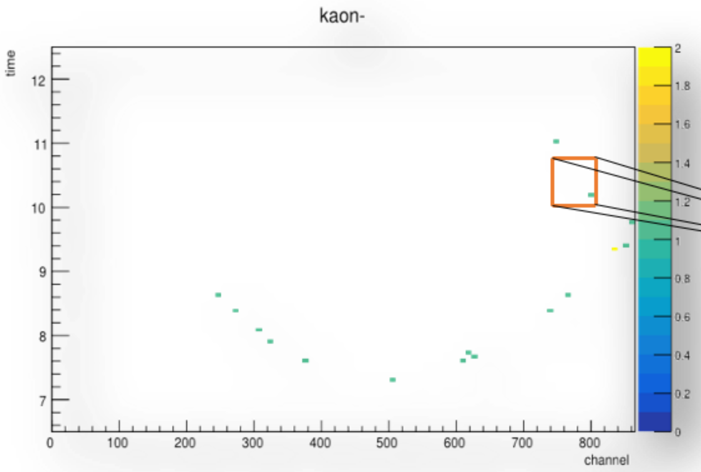
OSCAR simulates the digitized MC samples for Pion/Kaon

- $p \in$ 0.3-2.4 GeV/c, $\theta \in 23-36^{\circ}$, $\varphi \in$ 0-2π
- $0 \leq$ Channel ≤ 864
- $5.5 \leq$ Time ≤ 15.5 ns (Time resolution ~ 50 ps)
- bin number: Channel * Time =216* 200



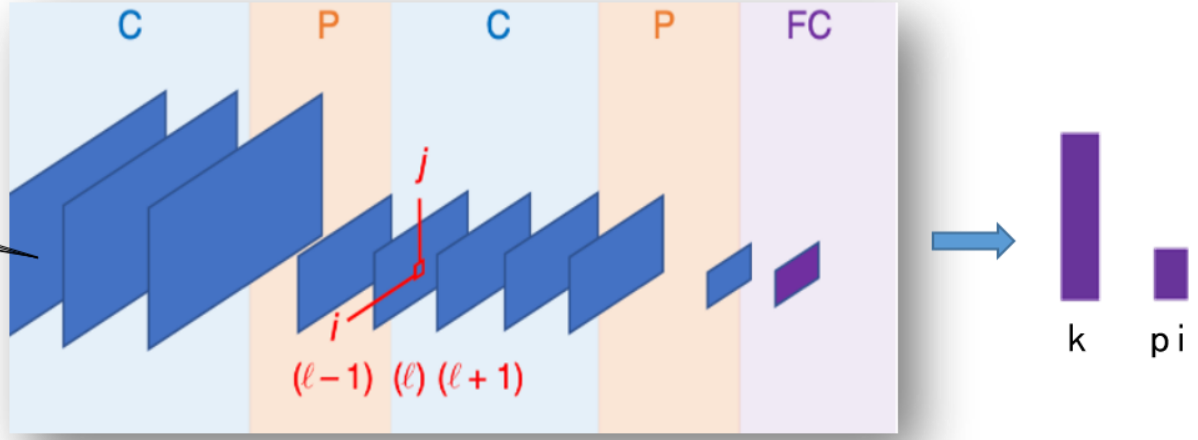


CNN Algorithm : process the 2D pixel map constructed for each event



single event

convolution



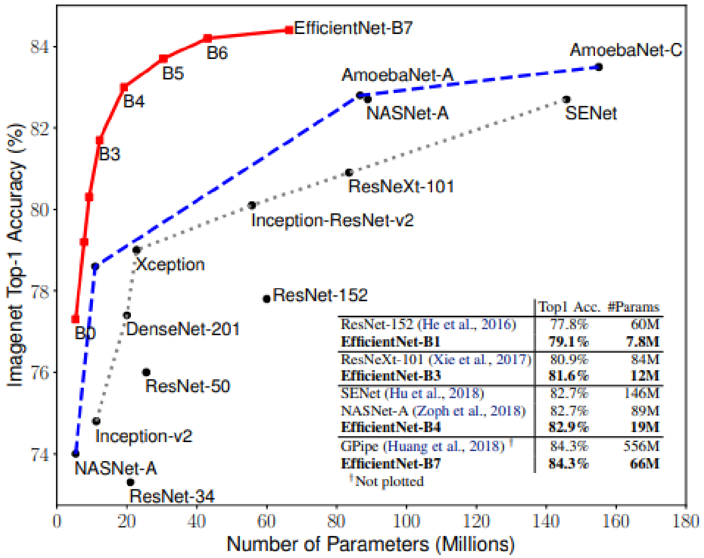
2 CNN Model: EfficientNetV2

EfficientNet

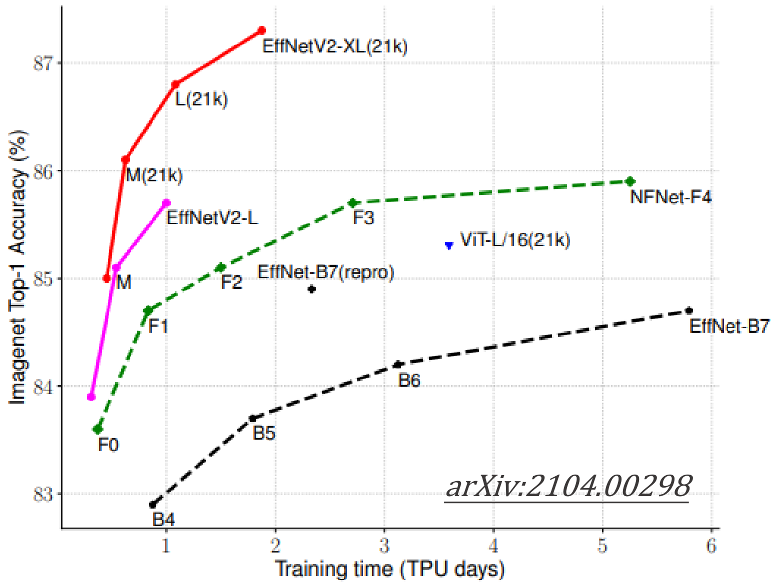
Version1： Utilizes a composite coefficient to uniformly scale the depth, width, and input image of the network

Version2： More lightweight, it reduces the number of parameters, thereby accelerating computation speed

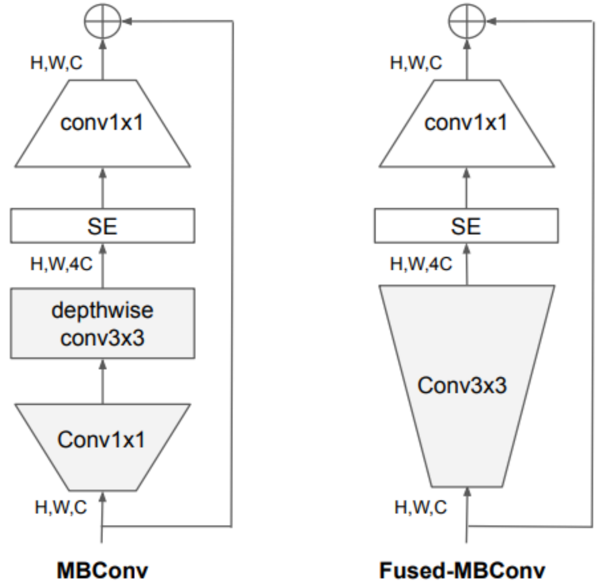
V1 Accuracy



V2 Improvement



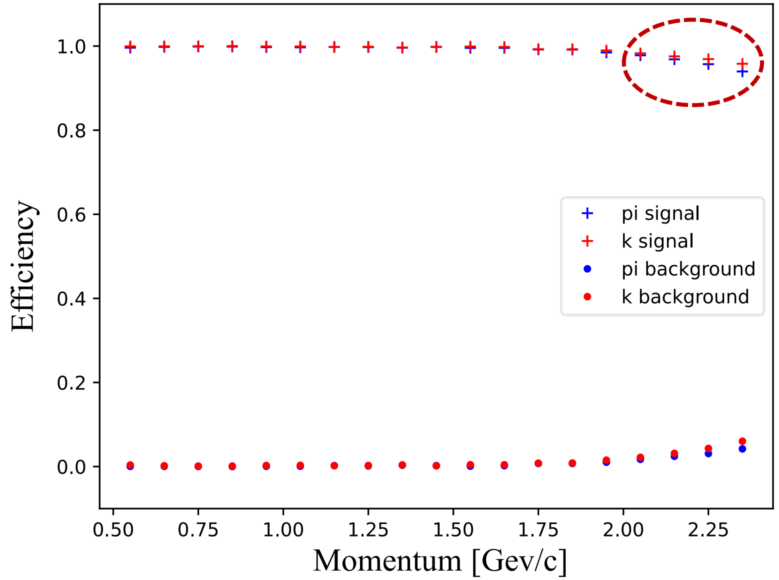
(Fused)MBConv

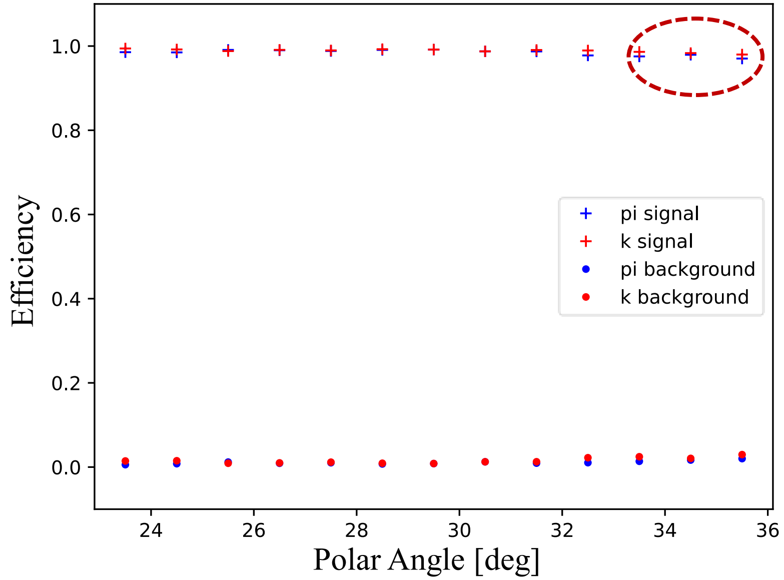


- Using EfficientNetV2-Small as the baseline model, adjusting and optimizing the network structure
- The model input consists of a 2D image constructed from the photon hit channels and arrival times
- Tracking momentum and position extrapolated to DTOF is added to a fully connected layer

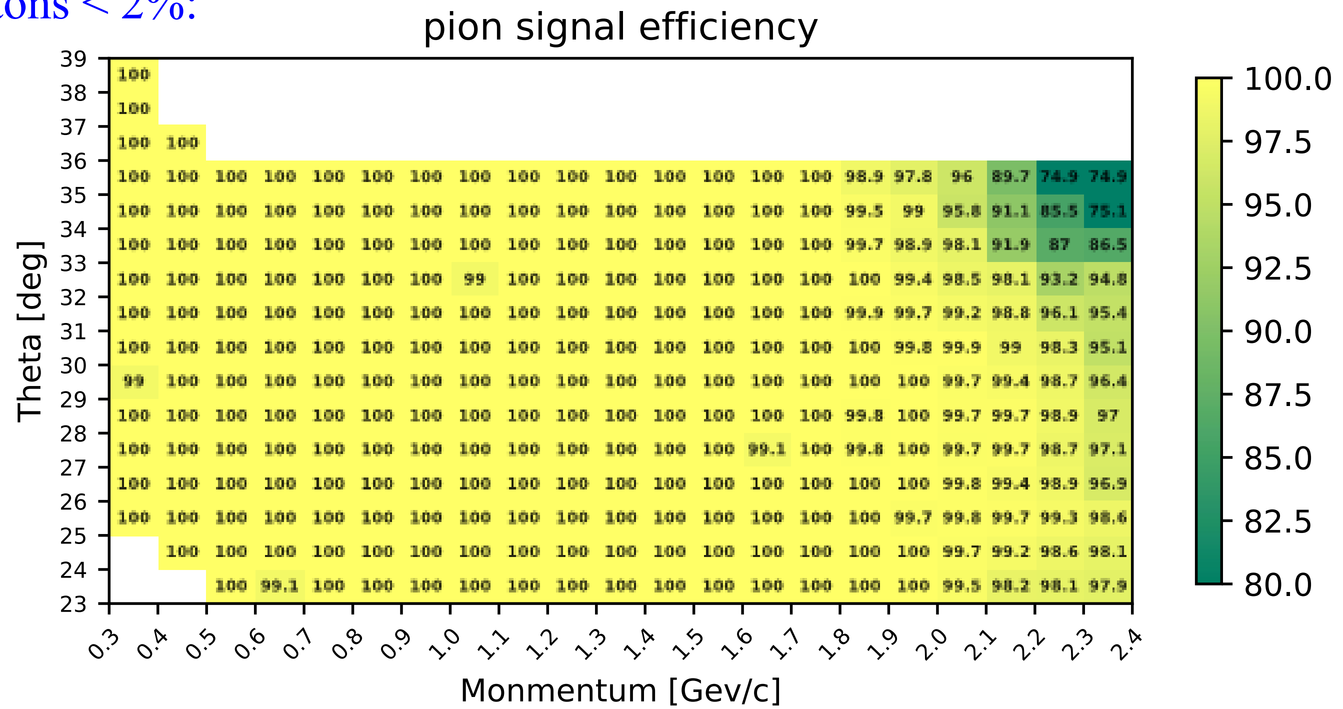
2 CNN PID Performance

The signal efficiency and background misidentification rate for pion/kaon across momentum and polar angle





adjust thresholds on the predicted probabilities to control the misid rate of kaons < 2%:

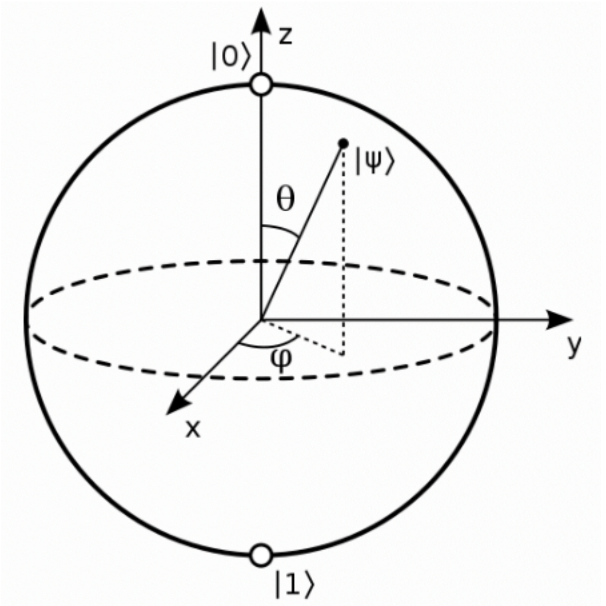


- The CNN model exhibits good PID capability in most of the phase space
- Meet the STCF requirements for π/K identification (p<2Gev/c, eff >97%)
- The performance in the high momentum and large angle shows a decline

3 Quantum Computing and Quantum Machine Learning

Future HEP experiments with higher energy requirements will pose significant challenges to traditional methods, and current ML algorithms will also consume more computational resources.

- **Quantum machine learning: under the domain of quantum computing/algorithm**
  - Provide alternatives/enhancement for traditional machine learning algorithms
- **Potential quantum advantage for ML problems**
  - simplify computational complexity, accelerating computation speed
  - leverag more information in high-dimensional Hilbert spaces through superposition and entanglement



3 QML for HEP

The application of QML algorithms in HEP

- **Quantum Generative Adversarial Networks (QGAN) for physical simulations**
- **Quantum Graph Neural Networks (QGNN) for track reconstruction**
- **Quantum Support Vector Machines (QSVM) and Variational Quantum Classifiers (VQC) for event classification**

Due to the limitations of NISQ, most current QML algorithms are hybrid quantum-classical algorithms.

Parameterized quantum circuits are used as structural layers in machine learning models.

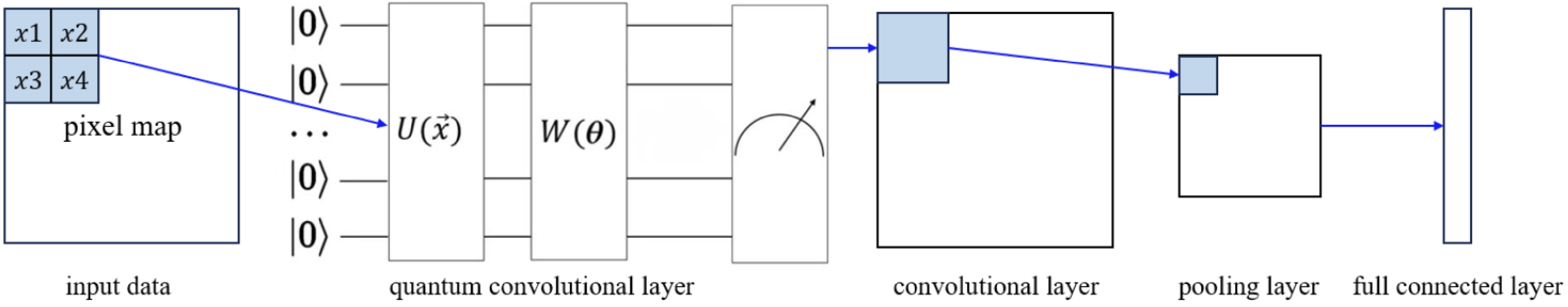
Based on classical CNNs, we have developed a Quantum Convolutional Neural Network (QCNN) to conduct feasibility studies on the π/K identification, exploring potential quantum advantages.

1. Data encoding circuit

Classical data cannot be processed directly on quantum system and must be mapped into quantum state space through data encoding circuits. The quantum circuit does not apply the entire image map to a quantum system at once, but processes it using RX rotation gate as much as the filter size at a time.

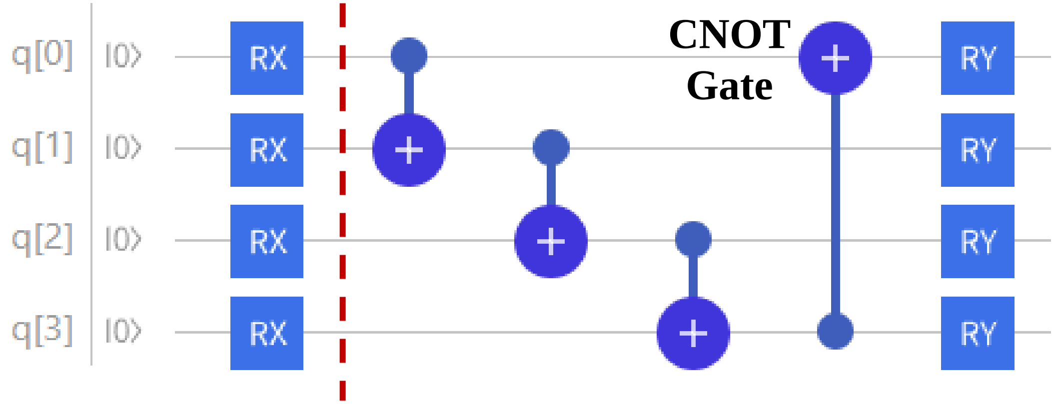
$$RX(x_i) = \exp\left(-i\frac{x_i}{2}X\right) = \begin{pmatrix} \cos\frac{x_i}{2} & -i\sin\frac{x_i}{2} \\ -i\sin\frac{x_i}{2} & \cos\frac{x_i}{2} \end{pmatrix}$$

x_i is fixed by the classical feature



2. Quantum convolution circuit

Use parameterized variational quantum circuits to process the quantum states generated from the image sub-regions in the previous step.



θ is a free parameter that is optimized during training

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

$$\mathrm{CNOT} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

3. Measurement and decoding

The decoding process is achieved by measuring the qubits on the observables $\hat{\boldsymbol{o}}$. The expectation value from multiple measurements is deterministic. Based on this classical value, quantum convolutional layers can naturally couple with the subsequent classical layers.

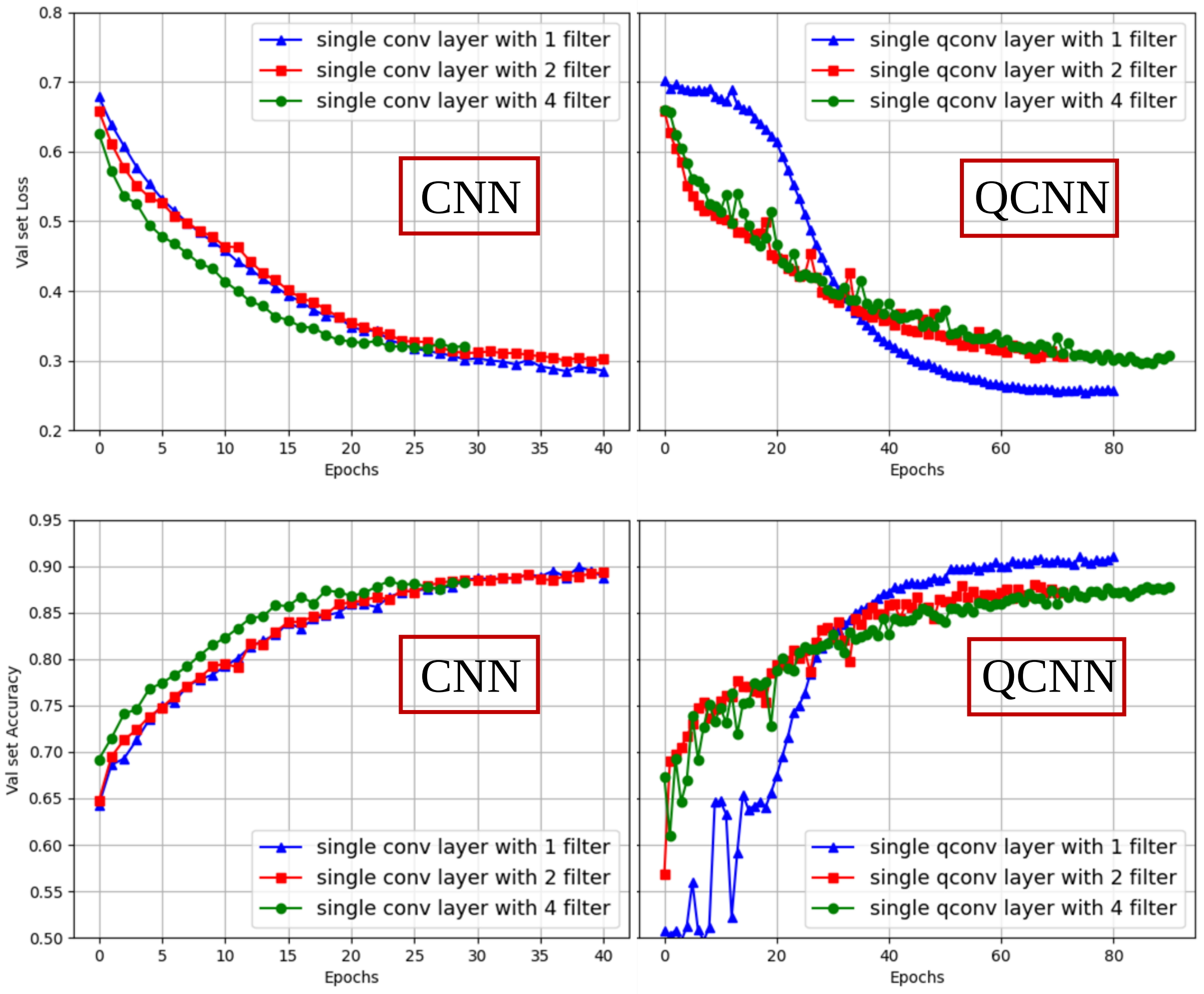
$$f(\boldsymbol{\theta}) = \left\langle \psi | W^{\dagger}(\boldsymbol{\theta}) \hat{O} W(\boldsymbol{\theta}) | \psi \right\rangle$$

4 Performance Comparison between QCNN and CNN

The Google quantum simulator based on Cirq and Tensorflow_quantum

Test a performance comparison between a single QCNN/CNN layer with 1, 2, and 4 convolutional kernels

- **Single Q/C ONV2D**(**filters = n**, (2, 2), shape= (32*32)),
- Flatten()
- Dense(128, activation='relu'), Dense(2)
- Adam learning_rate = 0.0001, batch_size = 16



- The quantum circuit utilizes rotation gates with four parameters, maintaining the same parameter level as a 2*2 classical convolutional kernel.
- On small datasets, the QCNN demonstrates feature extraction and learning capabilities comparable to those of classical CNNs.

Summary

- For the π/K identification on DTOF, we developed a CNN model based on original detector responses and reconstructed features.
- The CNN performance currently meets the physical requirements of STCF, but there is still room for optimization at high momentum and large angle ranges.

- Quantum machine learning is expected to provided new algorithm possibilities in high-energy physics experiments due to potential quantum advantages.
- A feasibility study on QML showed that QCNN and CNN perform similarly on small datasets.
- Due to quantum resource limitations, comparisons have only been made on small-scale datasets. Handling large-scale data remains a key advantage for classical machine learning.

27th International Conference on Computing in High-Energy & Nuclear Physics

Thank you for listening