



山东大学

2024年量子计算与机器学习研讨会

STCF DTOF上基于CNN/QCNN的粒子鉴别

姚志鹏，黄性涛，李腾

山东大学

2024.08.08 长春

- I. 研究背景
- II. 基于卷积神经网络的Pion/Kaon鉴别
- III. 量子计算和量子机器学习
- IV. 量子卷积神经网络用于PID的可行性研究
- V. 总结

1 超级陶粲装置 (STCF)

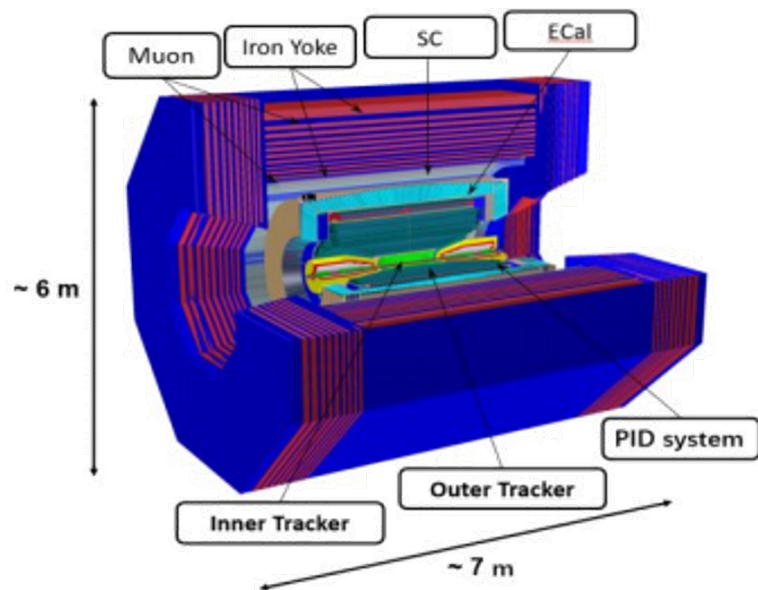


超级陶粲装置 (STCF) 是中国新一代的正负电子对撞机。

- 质心能量 2-7 GeV
- 峰值亮度 $0.5 \times 10^{35} \text{ cm}^{-2}\text{s}^{-1}$
- 具备未来进一步提升峰值亮度和实现束流极化的空间

丰富的物理研究

- c夸克和 τ 轻子的物理研究
- 非微扰强相互作用和强子结构
- 新物理寻找



从相互作用点向外，STCF包括循迹系统 (ITK 和 MDC)，粒子鉴别系统 (PID)，电磁量能器 (EMC)，超导线圈 (SCS)，缪子探测器 (MUD)

ITK <ul style="list-style-type: none">• $< 0.25\% X_0 / \text{layer}$• $\sigma_{xy} < 100 \mu\text{m}$	Cylindrical μ RWELL CMOS MAPS
MDC <ul style="list-style-type: none">• $\sigma_{xy} < 130 \mu\text{m}$• $\sigma_{\eta/p} \sim 0.5\% @ 1 \text{ GeV}$• $dE/dx \sim 6\%$	Cylindrical Drift chamber
PID <ul style="list-style-type: none">• 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	pCsl + APD
MUD <ul style="list-style-type: none">• 0.4 - 2 GeV• π suppression > 30	RPC + scintillator

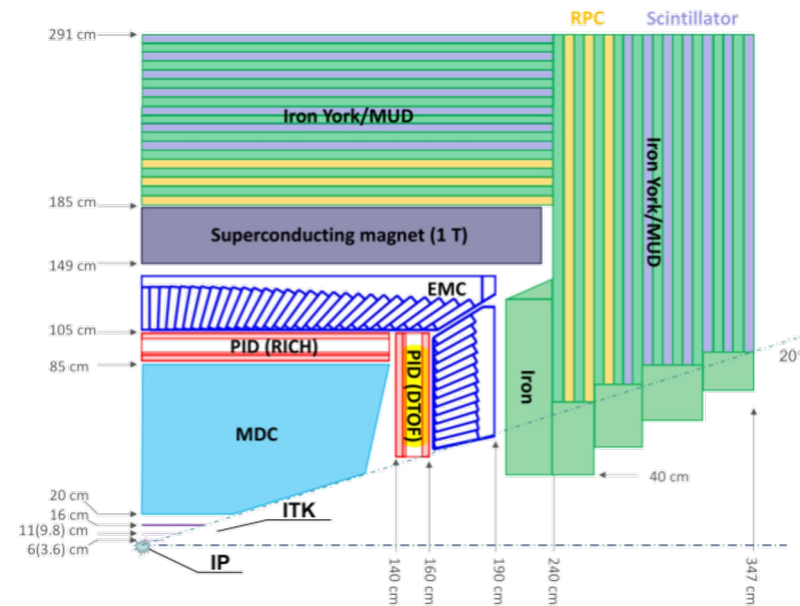
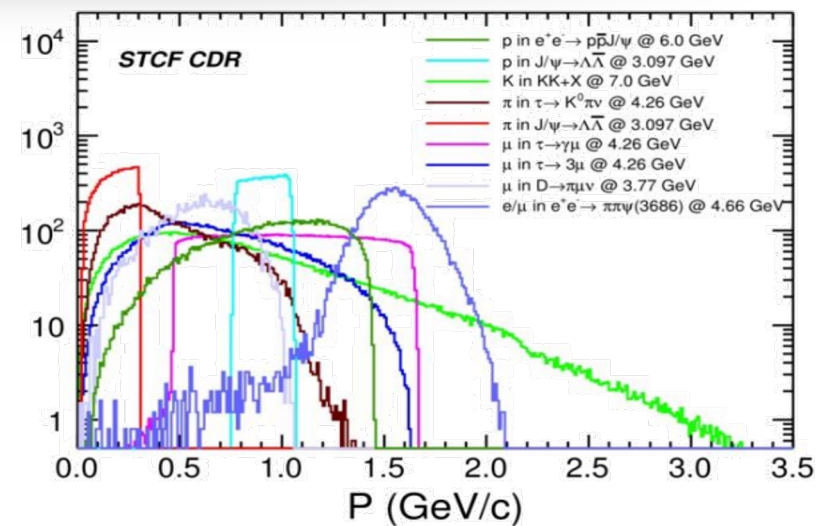
粒子鉴别（**PID**）是对撞机实验中进行物理研究的重要工具，为了更好的满足物理需求，STCF对高动量下PID的性能提出了更高的要求。

- $p < 2 \text{ GeV}/c$ 时， π/K 误判率 $< 2\%$ 时，相应鉴别效率 $> 97\%$
- $p > 0.7 \text{ GeV}/c$ 时， μ 子鉴别效率 $> 95\%$ ； $0.5 < p < 0.7$ 时，鉴别效率 $> 70\%$
- 良好的中性粒子鉴别能力

为了进一步提高PID的效率，STCF还设计了一个PID系统，使用两种切伦科夫探测技术：

- 环形成像的切伦科夫探测器(**RICH**)，位于桶部
- 基于探测内部全反射切伦科夫光的飞行时间测器(**DTOF**)，位于端盖

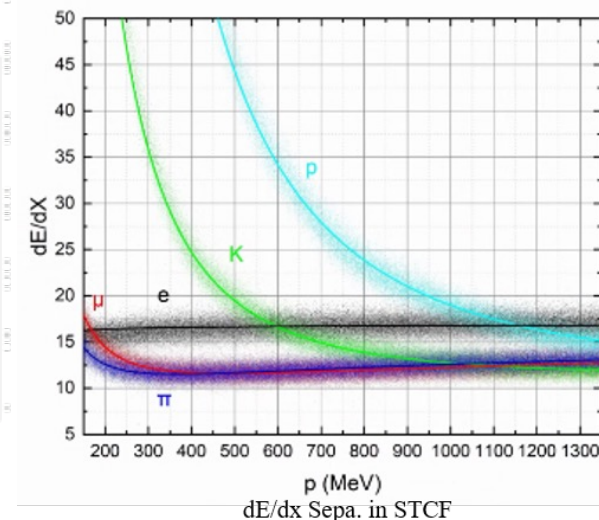
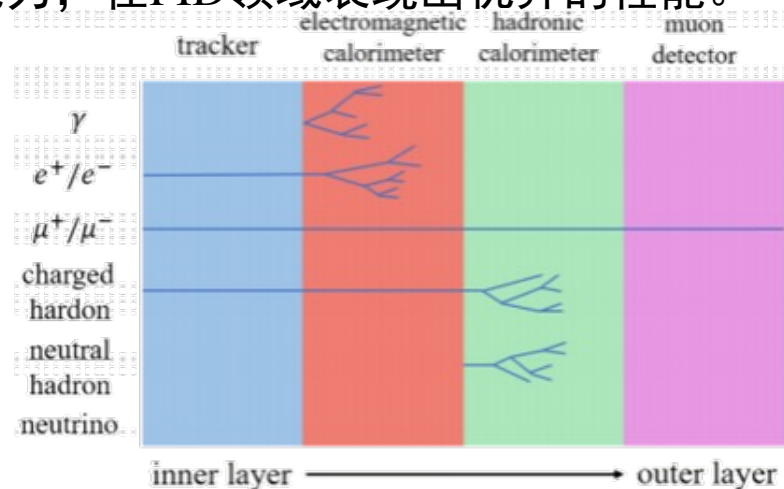
主要针对 π/K 的鉴别，也能为分辨 μ/π 提供一定的能力



仅依赖单个探测器的信息很难准确地分辨粒子，传统方法对于高PID性能的要求面临一定的挑战。

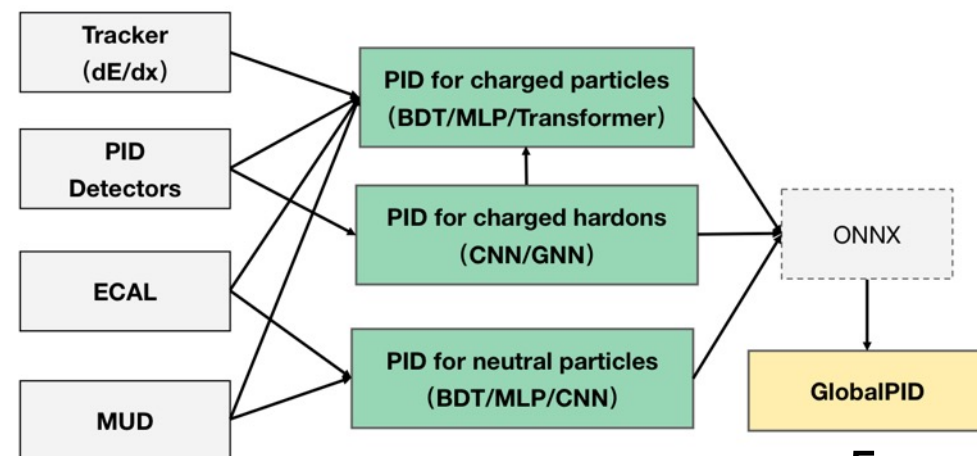
机器学习凭借在高维数据空间中提取有用特征的能力，在PID领域表现出优异的性能。

- 联立多个探测器的信息进行粒子鉴别
- 能够充分利用探测器的原始响应
- 增强决策树,神经网络和支持向量机等技术
都已在PID领域发挥重要作用。



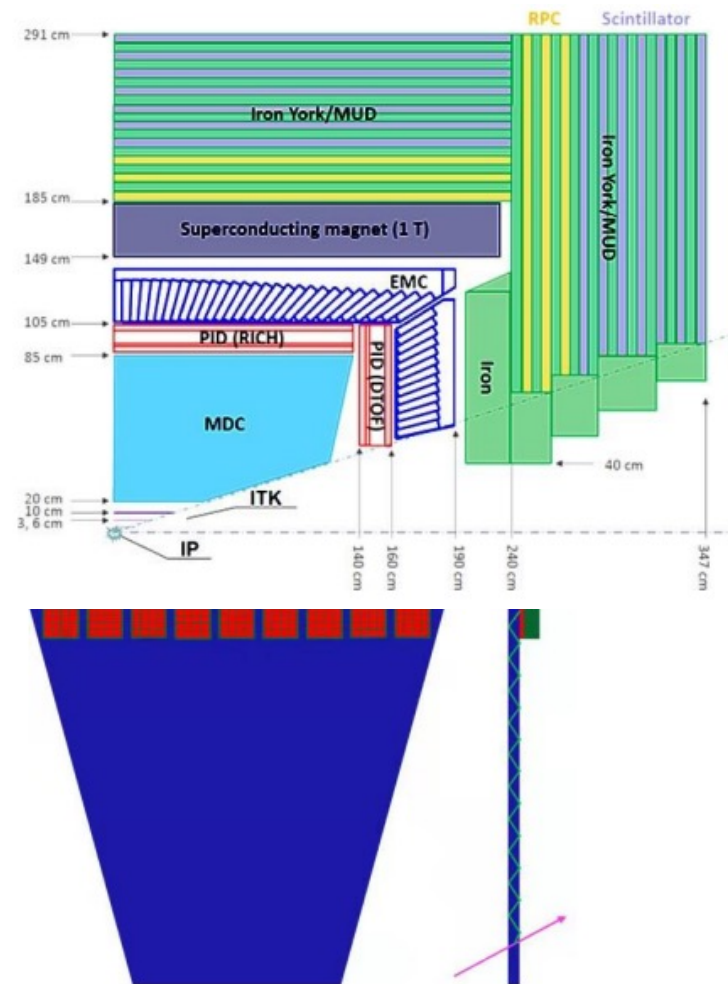
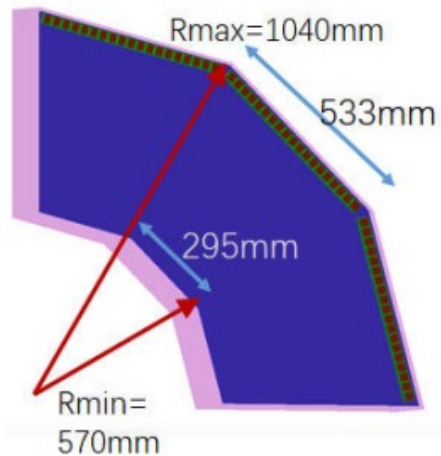
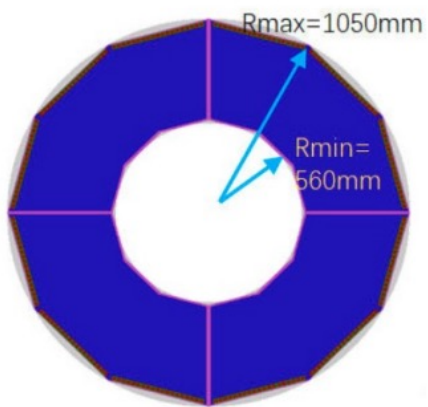
目前已基于机器学习算法开发了多种粒子鉴别技术，以适应不同类型的探测器。后续可基于ONNX移植到OSCAR框架。

本报告针对DToF中的 π/K 鉴别问题，利用探测器的原始响应和重建特征来开发卷积神经网络（CNN）来实现高PID能力。



DTOF作为STCF的端盖PID探测器，是基于探测内部全反射切伦科夫光的技术，最早由BaBar引入。

- 由两个相同的端盖圆盘组成，含多个相同的扇区
- 探测器覆盖的极角范围约为 $23\text{-}36^\circ$
- 熔融二氧化硅板作为切伦科夫辐射体产生光子
- 辐射体外侧耦合多阳极微通道光电倍增管（MCP-PMTs）来接收切伦科夫光子

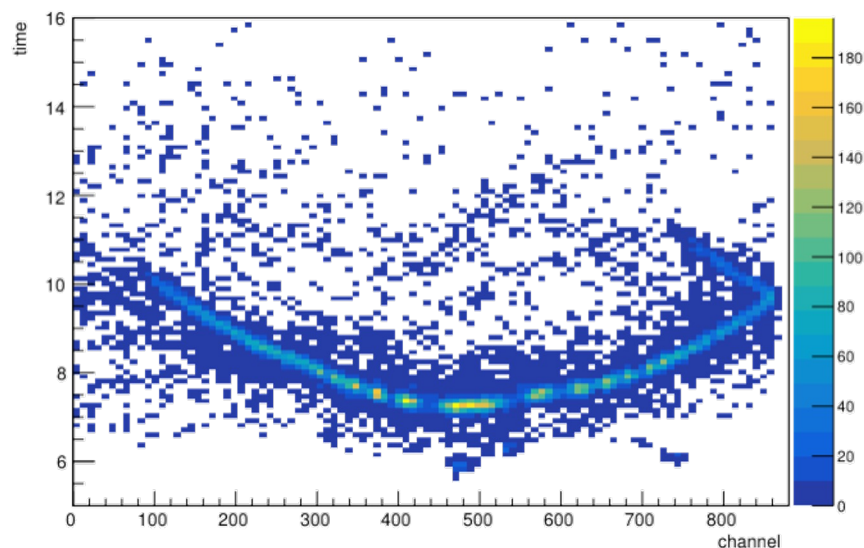


传统似然法：基于重建TOF分布构建似然概率密度函数

$$\text{TOF} = T - \text{TOP} - T_0 = T - \frac{L\bar{n}_g}{c} - T_0.$$

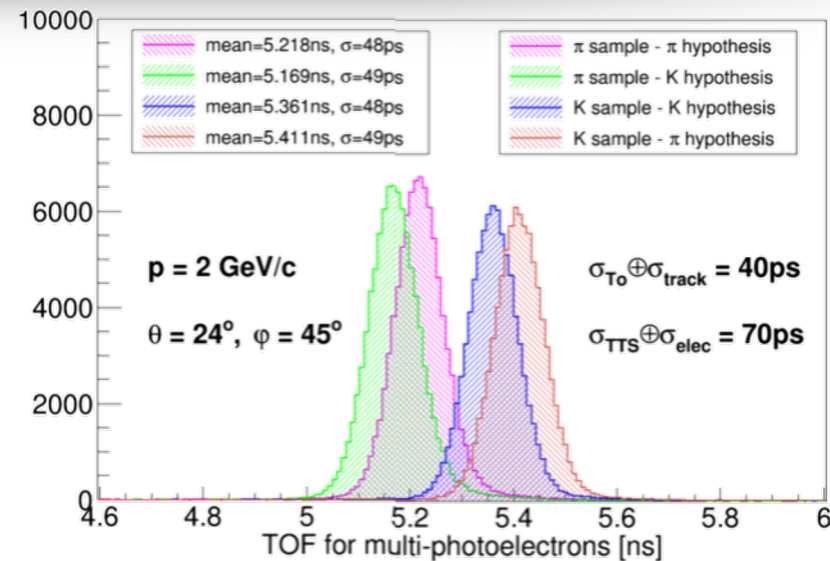
$$\mathcal{L}_h = p_h(N_{p.e.}) \prod_{i=0}^{N_{p.e.}} f_h(\text{TOF}_i)$$

可能忽略了PMT pattern上的空间信息（光子读出通道）？



利用DTOF探测器原始响应，构建二维pixel map：

- X-label: PMT所接收的切伦科夫光子的读出通道
- Y-label: PMT所接收的切伦科夫光子的到达时间
- Value: bin里的光子数目

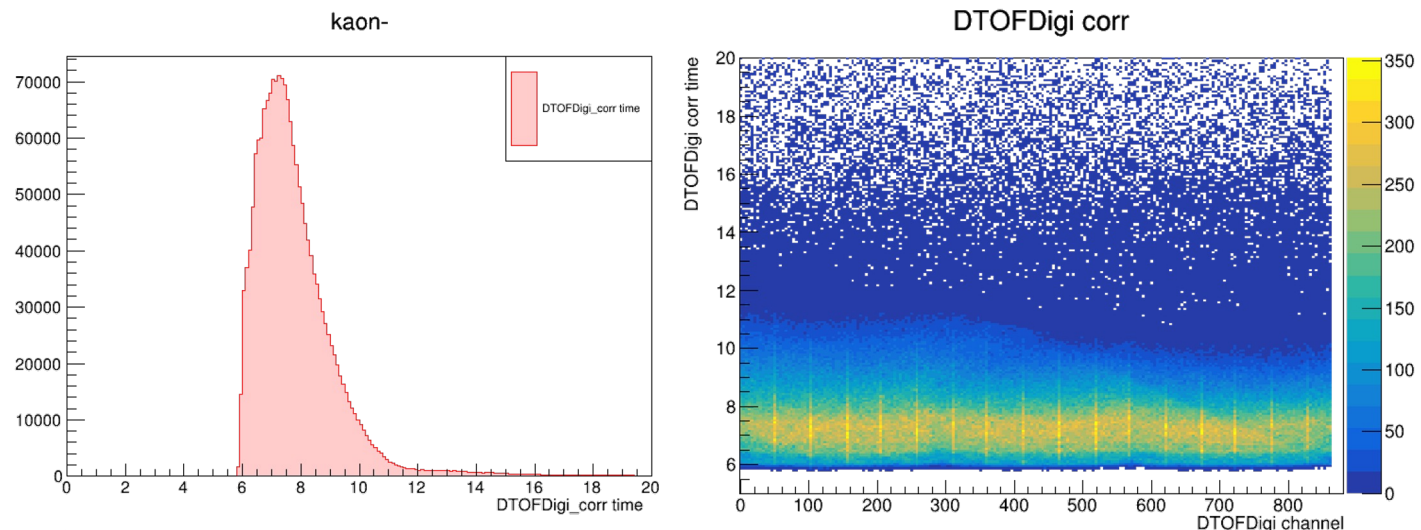


构建的图像反映出光子的拓扑结构，像素值更大的地方表示粒子有更大概率在当前时间击中对应通道。

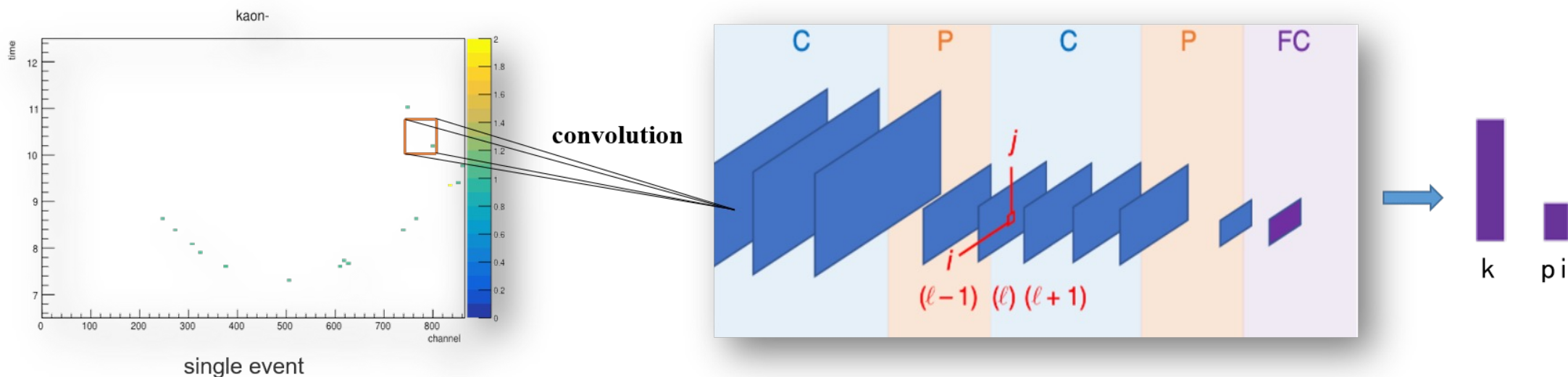
使用OSCAR模拟单粒子Pion/Kaon的MC样本，动量0.5-2.4 Gev，极角23-36°，数据量达800w

根据大量数字化数据的Channel和Time分布：

- $0 \leq 4 * \text{Channel}_x + \text{Channel}_y \leq 864$
- $5.5 \leq \text{Time} \leq 15.5 \text{ ns}$ (时间分辨~ 50 ps)
- 分bin: $\text{Channel} * \text{Time} = 216 * 200$



对每个event构建的2D pixel map输入CNN，进行训练与分类



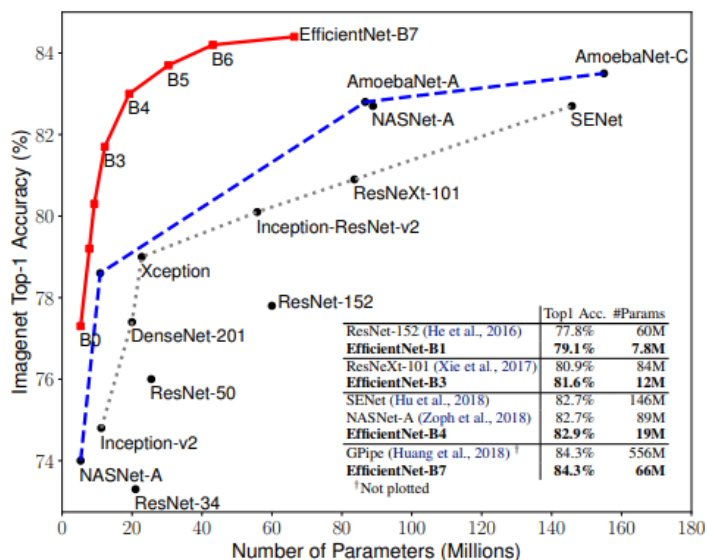
2 CNN模型：EfficientNetV2



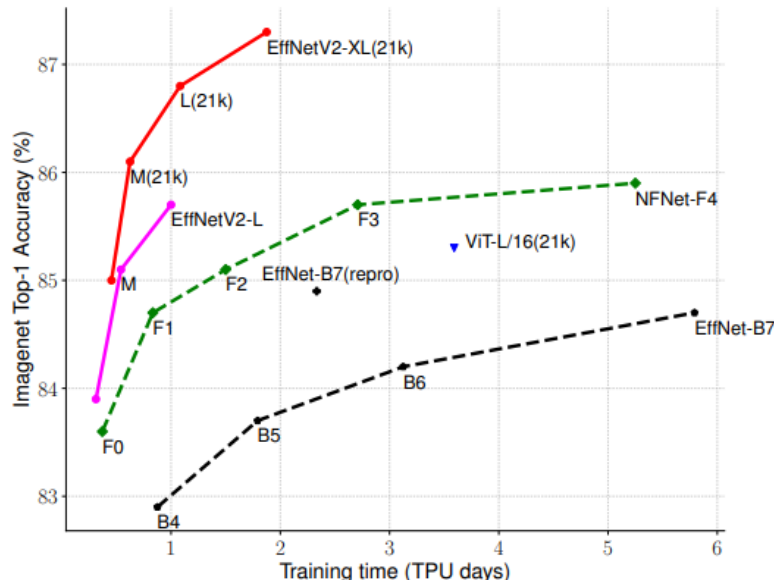
EfficientNet

V1: 采用复合系数均匀缩放网络深度/宽度/输入图像分辨率的大小，能力十分强大
V2: 更轻量化的网络，相较于V1，减少了参数量，加快了计算速度

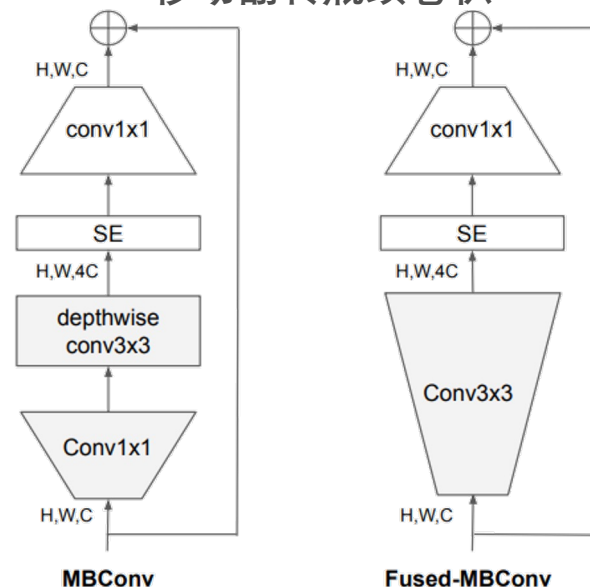
V1准确率



V2改进

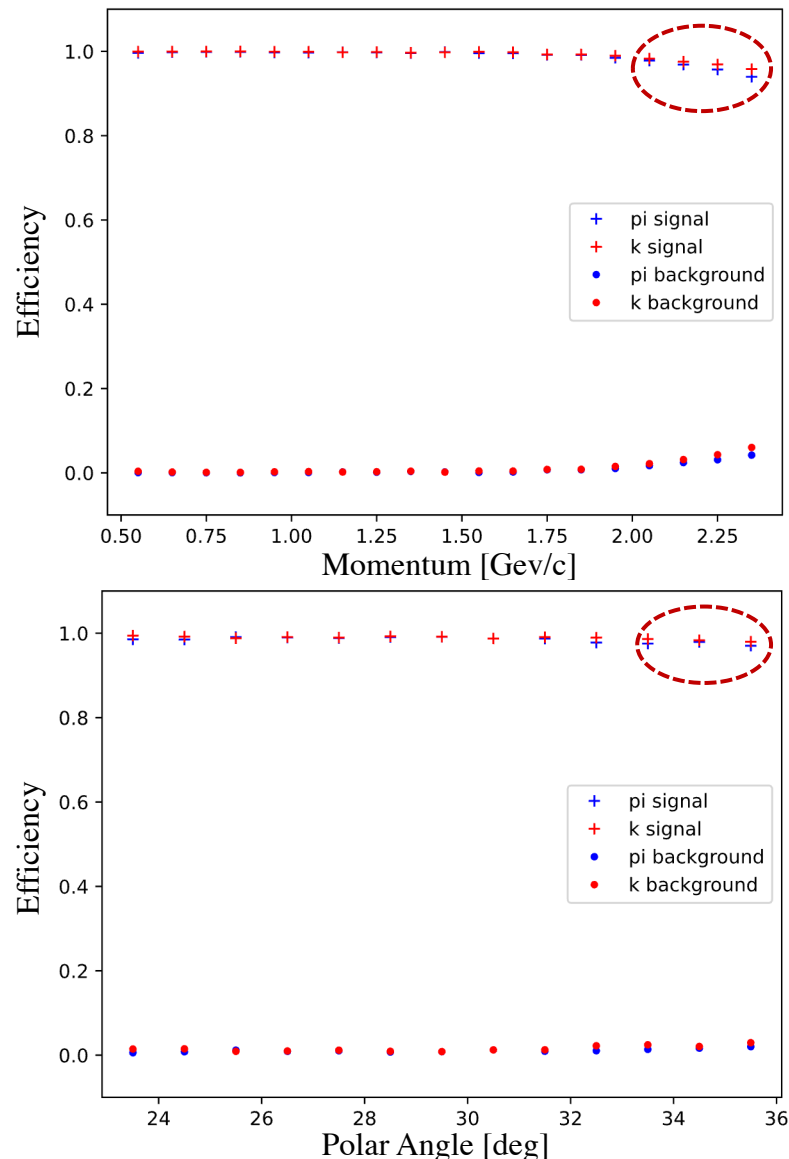


移动翻转瓶颈卷积

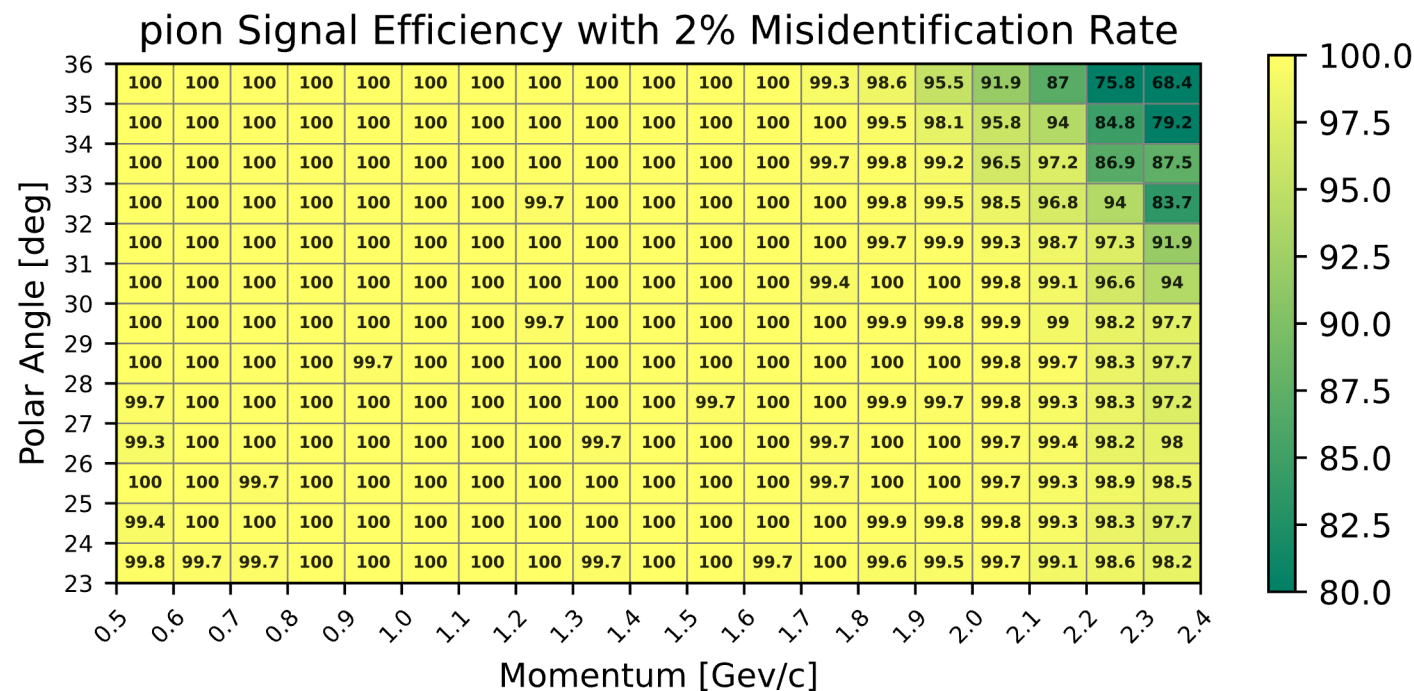


- 使用 EfficientNetV2-S 作为基准模型并进行调整优化
- 依赖光子击中通道和到达时间构建二维图像作为模型输入
- 全连接层额外添加重建得到的粒子击中DTOF时的动量和位置

Pion和Kaon在不同动量和极角范围下的信号效率和本底误判率



通过对模型预测概率设置不同的阈值，将kaon作为本底控制其误判率 $<2\%$ ，pion的信号效率分布如下：



- 在动量达2Gev时，能满足STCF对 π/K 鉴别效率($>97\%$)的要求
- 高动量、大角度区域效率有所衰减

3 量子计算与量子机器学习



- 未来更高的能量的HEP实验将极大挑战着传统方法，当前机器学习算法也将消耗更多计算资源。
- 量子计算技术的发展和成功应用有希望为高能物理提供新的算法可能。

量子计算利用量子力学特性(例如量子态叠加和纠缠)来解决传统计算机无法解决或进行计算加速。

量子比特:

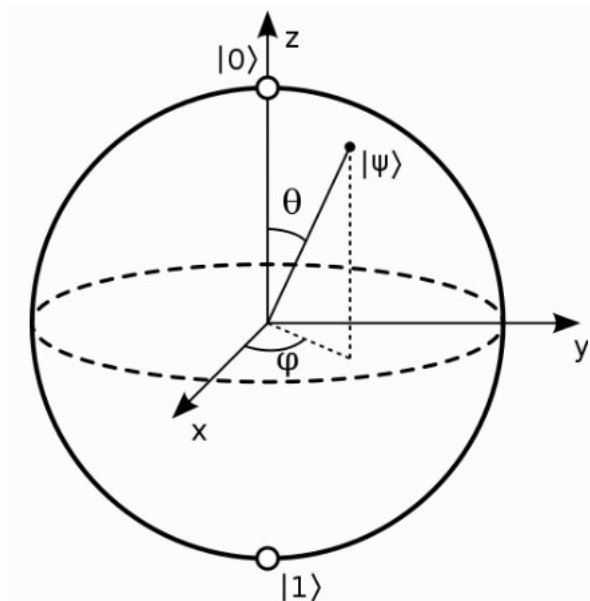
量子比特 (qubit) 有两种基态 $|0\rangle$ 和 $|1\rangle$ ，还可处于两者之间的叠加态。
从原点到布洛赫球面上任何一点的单位向量都可以是一个量子比特的状态。

量子逻辑门:

用于改变量子比特的状态。如Pauli-X门，类似经典逻辑非门， $|0\rangle$ 经过变为 $|1\rangle$ 。



量子算法，如Shor的因数分解算法和Grover的搜索算法，已经显示出在某些特定问题上超越传统算法的潜力。



布洛赫球

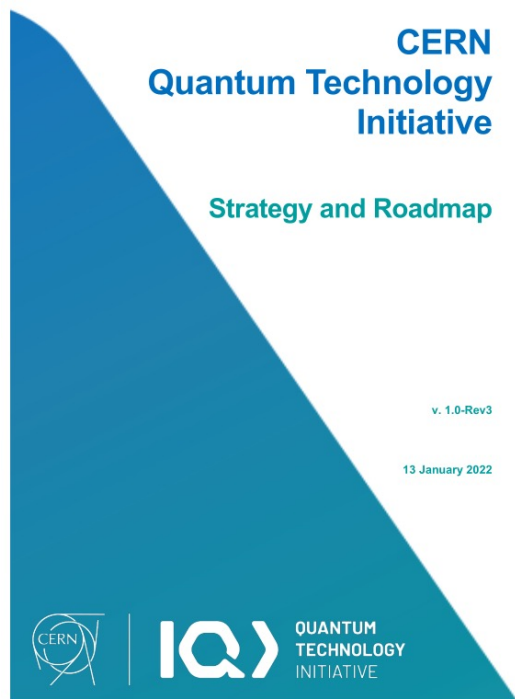
3 量子计算与量子机器学习



量子计算与机器学习的交叉，形成了新的领域—量子机器学习（QML）。

QML有望简化计算复杂度并加快计算速度，并且通过叠加和纠缠来利用高维希尔伯特空间更多的信息。

2021年CERN量子技术计划公布首份定义其中长期量子研究计划的路线图，促进量子技术在高能物理学领域的应用。



- 用于物理模拟的量子生成对抗网络（QGAN）
- 用于径迹重建的量子图神经网络（QGNN）
- 用于粒子鉴别的量子支持向量机和变分量子分类器（QSVM、VQC）

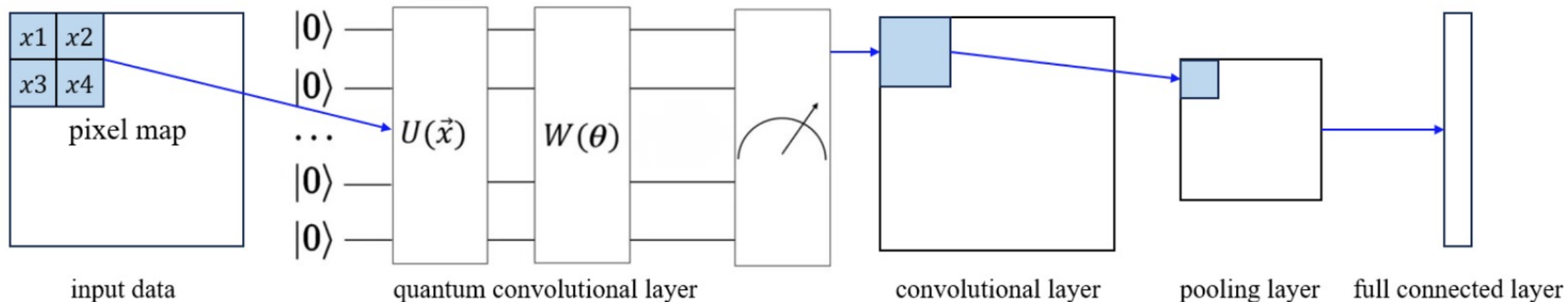
由于NISQ (noisy intermediate-scale quantum) 限制，目前QML算法大多为混合量子-经典算法。构建参数化的量子线路作为机器学习模型中的结构层使用，以及依赖经典优化。

基于经典CNN，使用模拟器开发了一个量子卷积神经网络（QCNN）来对 π/K 鉴别问题进行可行性研究，探索潜在的量子优势。

1. 数据编码线路

经典数据无法直接在量子线路上进行处理，需通过数据编码线路来将经典数据映射到量子态空间。受限于量子资源，无法一次性映射整个图像（缩放至 $32*32$ ），而是将其作为 $2*2$ 的区域像素进行遍历处理。通过应用RX旋转门来实现的该泡利角度编码操作。

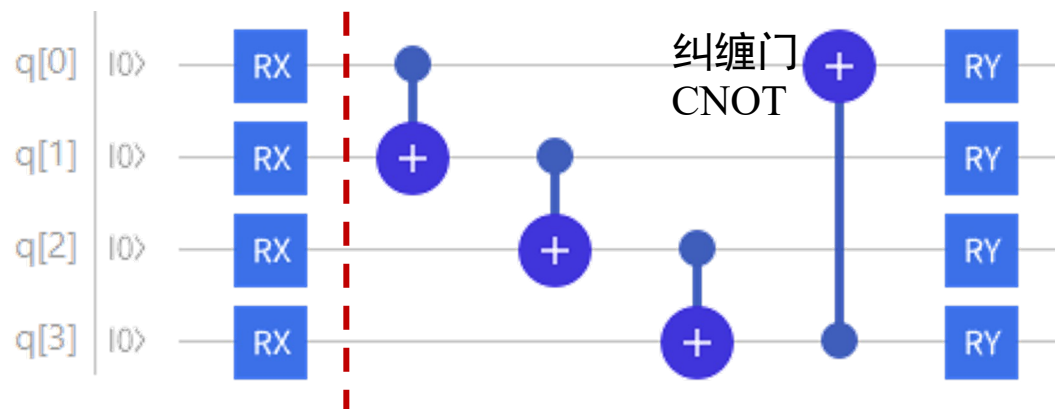
$$RX(x_i) = \exp\left(-i\frac{x_i}{2}X\right) = \begin{pmatrix} \cos\frac{x_i}{2} & -i\sin\frac{x_i}{2} \\ -i\sin\frac{x_i}{2} & \cos\frac{x_i}{2} \end{pmatrix} \quad x_i \text{由经典特征固定, 不随训练更改}$$



2. 量子卷积线路

利用参数化变分量子线路处理上一步图像子区域生成的量子态。在CNN的标准术语中，这相当于使用一个 2×2 的经典卷积核进行卷积，并可以设置步长为1。

θ 为自由参数，训练中根据梯度进行优化



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

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

3. 测量和解码

解码过程是通过测量量子态实现的。根据量子力学的基本原理，在可观测量上测量的期望值是确定性的。基于此经典值，后续量子卷积层可以耦合经典层。

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

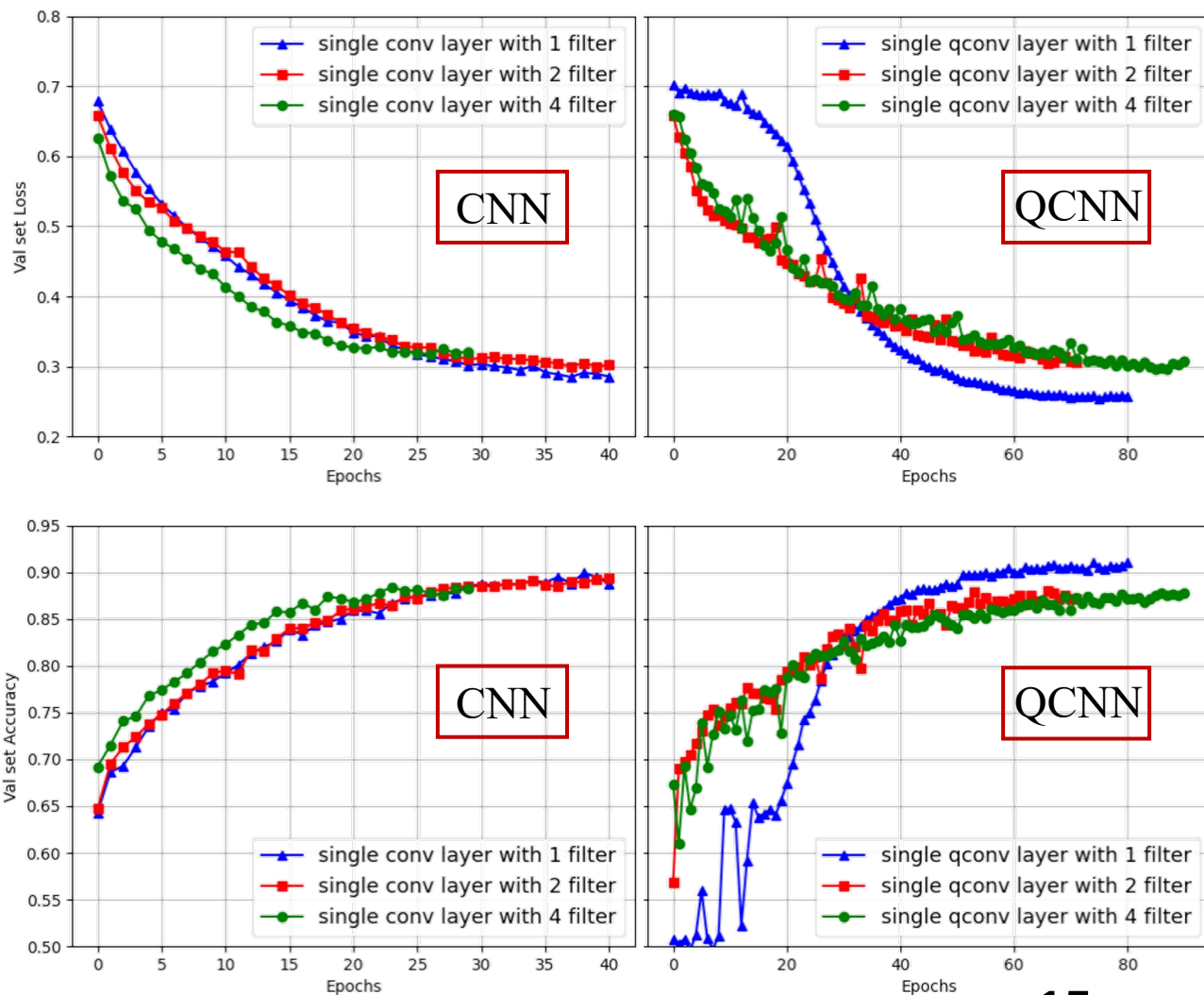
4 QCNN与CNN的性能对比



使用TensorFlow Quantum和Google Cirq的理想模拟器，目前初步开发了一个可训练的量子卷积层，在小的数据集上来替代CNN中的传统卷积层。

- ▶ **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

- 量子线路共使用含四个参数的旋转门和，与2*2经典卷积核保持**同一参数水平**。
- 在单个经典卷积层和量子卷积层情况下，进行多个卷积核下的性能比较。
- 总体而言，在小的数据集上，QCNN展现出与经典CNN相近的特征提取和学习能力。





- 对于D_{TOF}上的 π/K 鉴别，利用原始探测器响应和重建特征开发了一个CNN模型。
- 目前CNN的性能满足了STCF的物理需求，但在高动量和大角度范围还存在继续优化的空间。
- 量子机器学习有望凭借潜在的量子优势在高能物理实验中发挥重要作用
- 进行了关于QCNN的可行性研究，结果显示在小型数据集上，QCNN和CNN的表现相似。
- 由于量子资源的限制，目前只在小规模数据集上进行比较。对于经典机器学习，能处理大规模数据仍然是其关键优势。

2024年量子计算与机器学习研讨会

Thanks