Deep learning implementation for Dual-readout calorimeter

YunJae Lee (University of Seoul)

on behalf of the Korean Dual-Readout Calorimeter Team

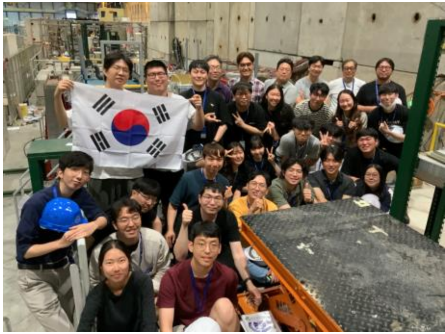
2023 NSRI Workshop

9 January 2024

E-mail: yunjae.lee@cern.ch

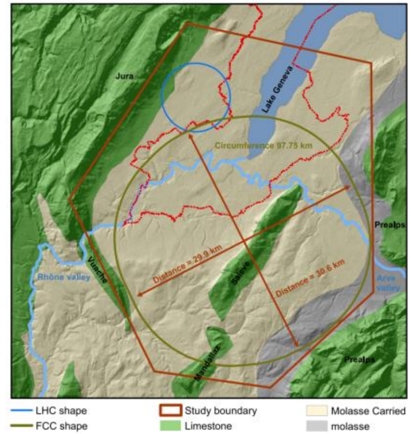
Activities

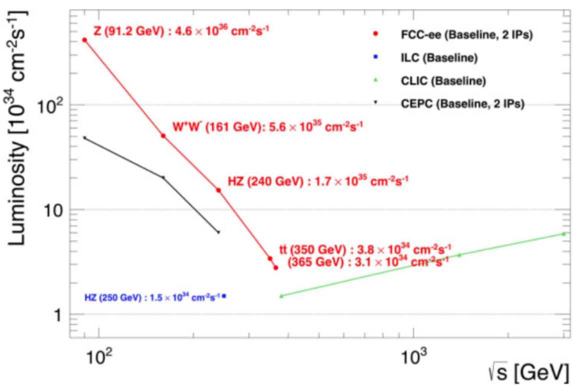
- KPS Fall 2023 - Readout bundling optimization of dual-readout calorimeter for particle identification using deep learning
- KPS Spring 2023 - Particle identification for dual-readout calorimeter using deep learning
- Korean DRC workshop 2022 at GWNU
- KPS Fall 2022 - Status of the Fast Simulation for Dual-Readout Calorimeter in Future e+e- Colliders using Generative Adversarial Network
- Test beam 2022 at CERN
- CALOR 2022 - Particle Identification for Dual-readout Calorimeter using Deep Learning
- KPS Spring 2022 – Hadronic Tau Identification for the Dual-Readout Calorimeter using Vision Transformer with Hyperparameter optimization
- KPS Fall 2021
  - Fast Simulation for the Dual-Readout calorimeter using GAN
  - Particle identification for Dual-Readout calorimeter
  - Vision Transformer based Hadronic Tau Identification for the Dual-Readout Calorimeter
- KPS Spring 2021
  - Fast simulation for Dual-Readout calorimeter using GAN
  - Machine Learning application on particle identification for dual-readout calorimeter
- APS April 2021 - Deep learning implementation for Dual-Readout calorimeter

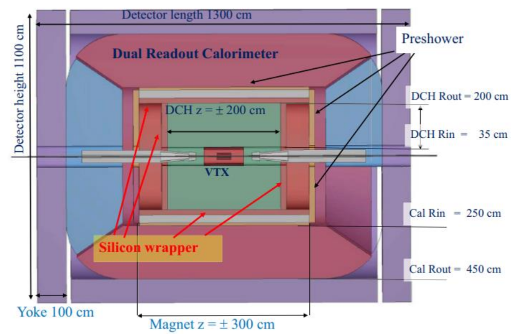


Dual-Readout Calorimeter

- Dual-readout calorimeter is at future circular collider(FCC) conceptual design report.
  - Dual-readout calorimeter is main component of IDEA detector
- FCC is 100 km circular collider on plan.
  - Planning to produce electron positron collision from 2045
- Korean dual-readout calorimeter team is doing main role of R&D.
- UOS contribution is mainly about deep learning study.

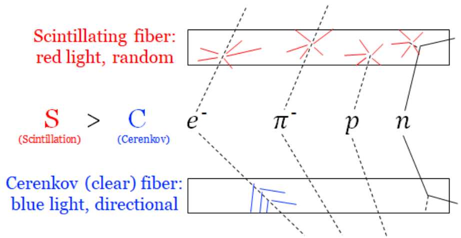


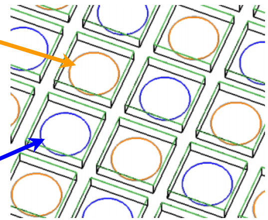


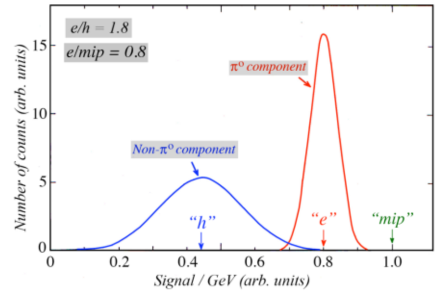


IDEA detector

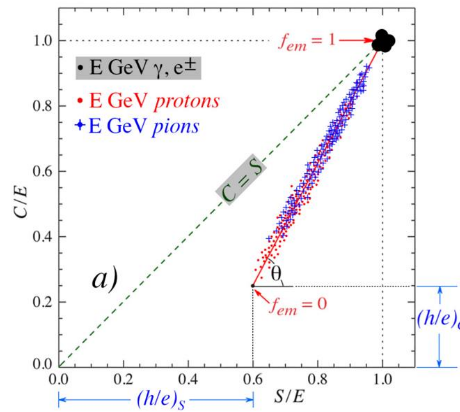
Dual-Readout Calorimeter





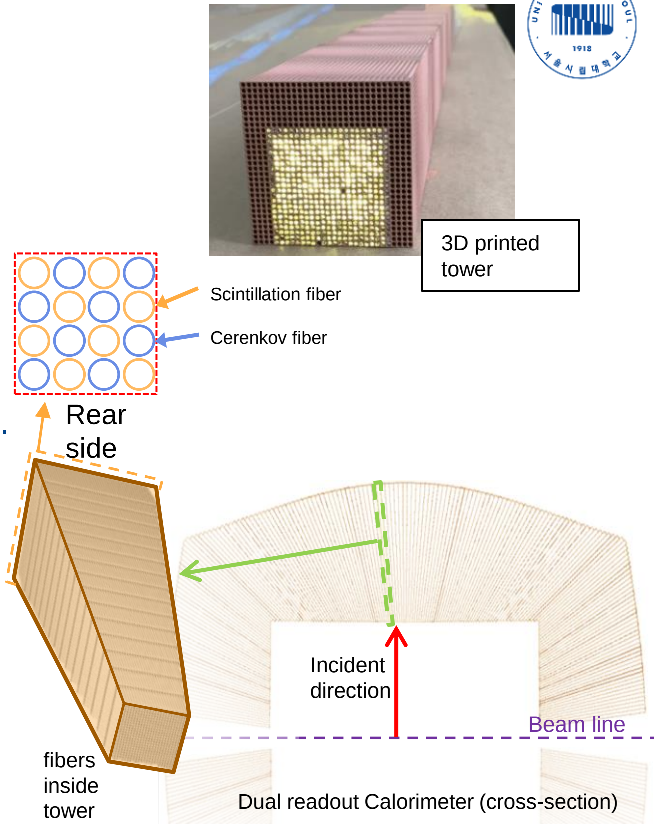


- Dual-readout calorimeter has two different, Scintillation and Cerenkov fibers components.
  - Scintillation fibers react to both EM and hadronic particle, Cerenkov fiber reacts to EM particle only.
  - Ratio of hadronic component and EM component $\boldsymbol{h/e}$ is differed by Scintillation part $(\boldsymbol{h/e})_S$ and Cerenkov part $(\boldsymbol{h/e})_C$.
- Hadronic energy can be measured with better resolution utilizing signals of hadronic component and EM component.



Simulation setup

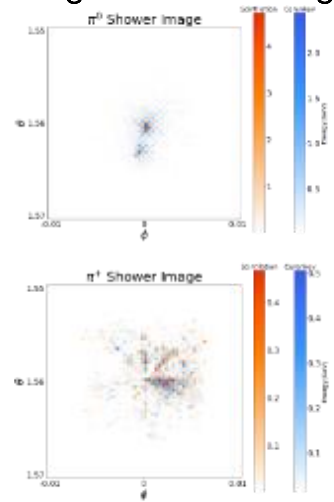
- Wedge tower geometry for simulation study
  - Cerenkov fiber and scintillation fibers are implanted in copper towers.
  - SiPM readout for end of each fiber.
- Monte Carlo Simulation
  - Calorimeter and shower are simulated by GEANT4.
  - Dual-readout calorimeter only – no other structure.
  - No magnetic field applied.
- Shower particles are generated at center of calorimeter with certain energy with directed to middle of barrel.
  - e^-, gamma, π^+, π^0 particles are generated with random energy in 10 - 100 GeV uniform range.



Preprocessing of reco energy

- Reconstructed scintillation and Cerenkov energy distribution of shower or jet are processed to 2 channel image.
  - Reconstructed energy pixelized by relative position of readout.
  - Image center is center of tower where mostly energy deposited.
  - Shower image covers 168x168 readouts range $\Delta\phi \in (-0.03, 0.03)$, $\Delta\theta \in (-0.03, 0.03)$.
- π^0 shower shows separated cluster, according to opening angle of its decays while π^+ shower images show wide spreads.
- Point cloud format is proposed as different input format.
  - Position of pixel and its value became point component.
  - Data size is reduced by sparsity of image.
  - Extended positional range covering hemisphere.
  - Timing(depth) information is used as additional channel.

Single shower image



Scintillation channel image

0	0	0	0	0	0
0	0.2	0.7	0	0	0
0	0.3	1.2	0.8	0	0
0	0	0.5	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0

Cerenkov channel image

0	0	0	0	0	0
0	0	0.6	0	0	0
0	0.4	1.0	0.7	0	0
0	0.2	0.5	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0

Point cloud

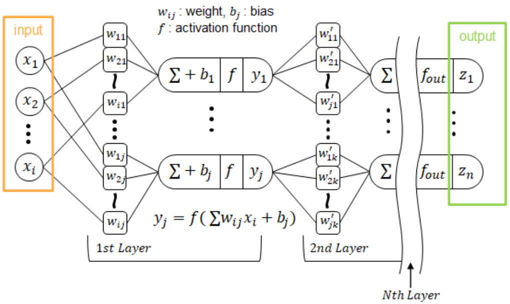
θ	3	4	3	3	2	2	2
ϕ	4	4	5	3	4	5	3
S	1.2	0.8	0.7	0.5	0.3	0.2	0
C	1.0	0.7	0.6	0.5	0.4	0	0.2

Width x width x 2 → # of hit x 4

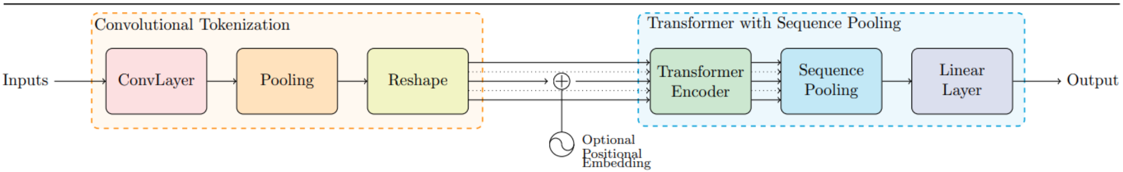
Deep learning model

- One of ML methods, which are based on neural networks
  - In each layer, weighted sum of inputs and bias are passed through activation function to outputs for subsequent layer.
  - Neural networks model can fit arbitrary dataset to necessary output.
- Transformer model applied with convolutional layer.
  - Transformer captures dependencies and relationships between distant region on image.
  - Applying convolution prior to transformer reduces input size to transformer and summarizes features.

Neural Networks



Compact Convolutional Transformer (CCT)[1)]



1) Ali Hassani et al., arXiv:2104.05704

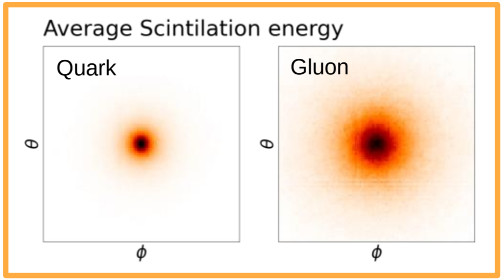
Deep Learning Application

Application areas	Technical goals	Works ongoing
Particle identification	Shower object identification, jet tagging.	Discrimination between quark/gluon or τ/quark discrimination, EM/hadronic shower identification.
Jet reconstruction	Mass and energy reconstruction.	Applying regression for quark/gluon jet discriminant variables.
Fast simulation(GAN)	GAN based shower generator.	Applying GAN for electron shower image generation.

- Event generation and hadronization are simulated with Pythia8.
- GEANT4 simulation and DD4hep was used for MC data of dual-readout calorimeter.
- Reconstructed energy deposits are being used for deep learning inputs.
- Target outputs(particle type, energy) are obtained from generation level.

Jet Identification

- Classification of quark and gluon jets.
  - CNN(input image) and PointNet(input list) model were applied.
  - Performed for 50 GeV u quark and gluon jets with only calorimeter information.
  - Model responses show separation between quark and gluon.



Jet discrimination ROC curve

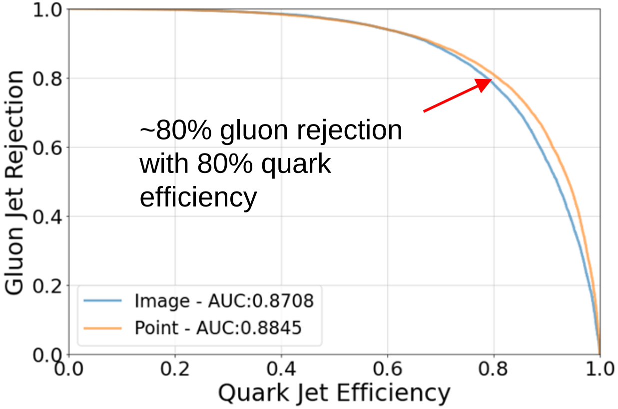
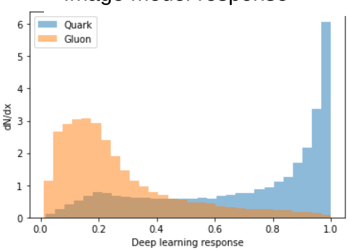
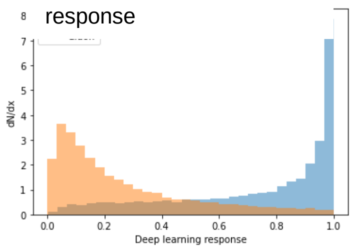


Image model response

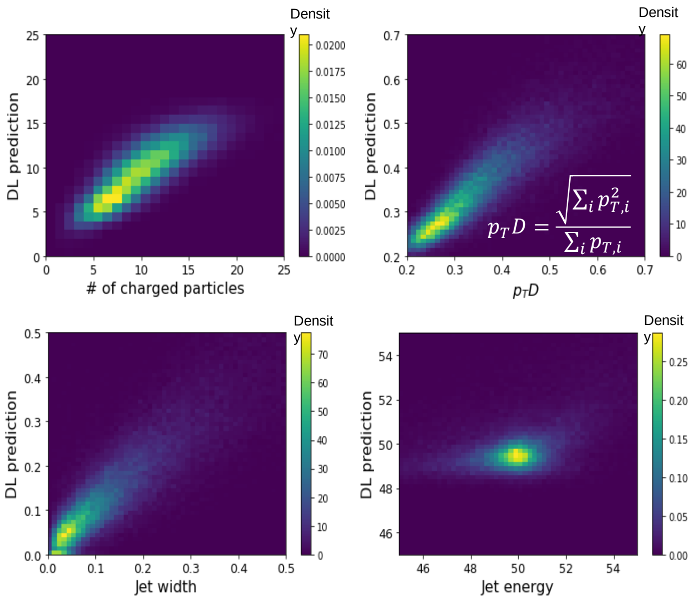


Point cloud model response



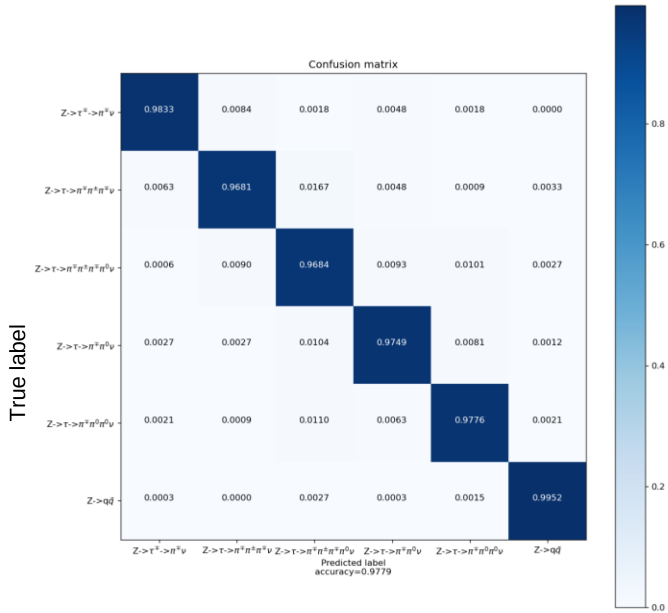
Jet Variable Regression

- Regression for jet variables of quark, gluon jet.
  - CNN model with jet image is applied.
  - Particle multiplicity, jet width, $p_T D$ (jet fragments variable) are used in jet discrimination.
  - Trained with images of 50 GeV u quark and gluon jets.
  - DL predictions follow tendencies with generated variables.



Tau identification

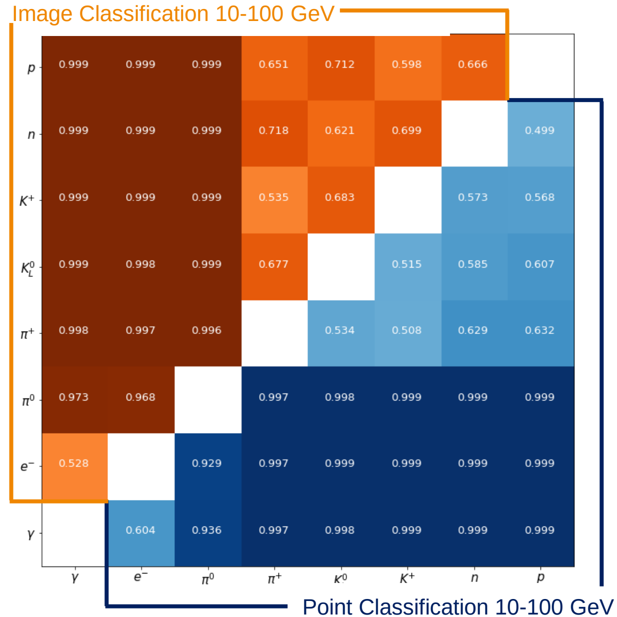
- Vision transformer methods is applied for τ jet identification compared to quark jet background.
- Identification rate in each section. Each row and column represents actual class and predicted class.
- $Z \rightarrow \bar{q} q$, main background of hadronic τ is identified about 99%.
- $Z \rightarrow \tau^{+}\tau^{-} \rightarrow$ (hadronic) are identified truly over 95%.
- Vision Transformer can be performed to hadronic τ classifier



Studied by Youngwan Son

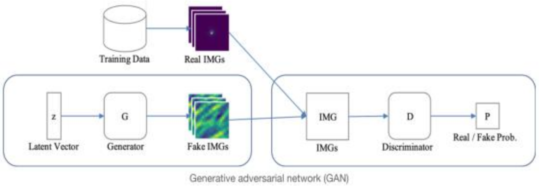
Particle Identification

- Binary Classification performed between EM or hadronic shower particles labeled at row and column.
- Shower from particle gun energy between 10~100 GeV.
- Left upper side shows AUCs of image classification, right lower side shows point cloud classification.
- EM showers and hadronic showers are well discriminated.
- Hadronic shower discrimination varied by charge and between meson and baryon.



Fast simulation using GAN

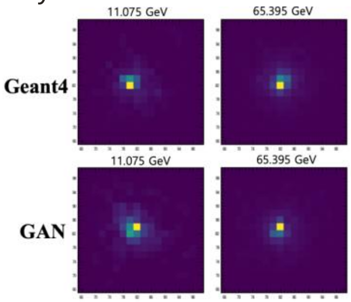
- Energy deposit image of electron shower(10~100 GeV) is trained with GAN model to generate fake image mimicking the Geant4 simulated shower image.
- Geant4 simulation takes times and it's proportional with energy.
- GAN model outputs almost instantly.

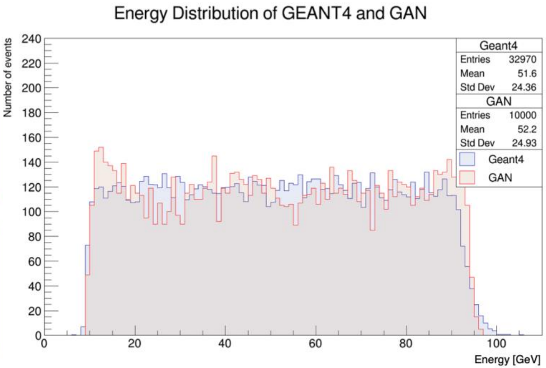


Energy	Type	Latency (s)
10 GeV	Geant4	1023.67
	GAN	0.45
100 GeV	Geant4	9564.91
	GAN	0.45

The latency time to simulation 10,000 events

Shower image

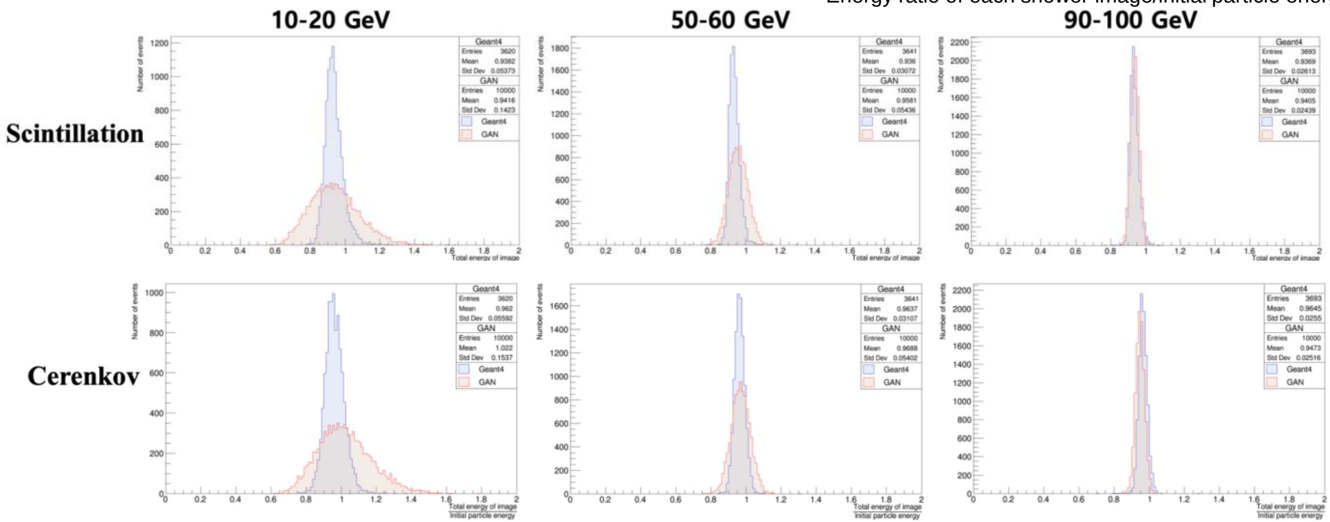




Studied by Doyeong Kim

Fast simulation using GAN

Energy ratio of each shower image/initial particle energy

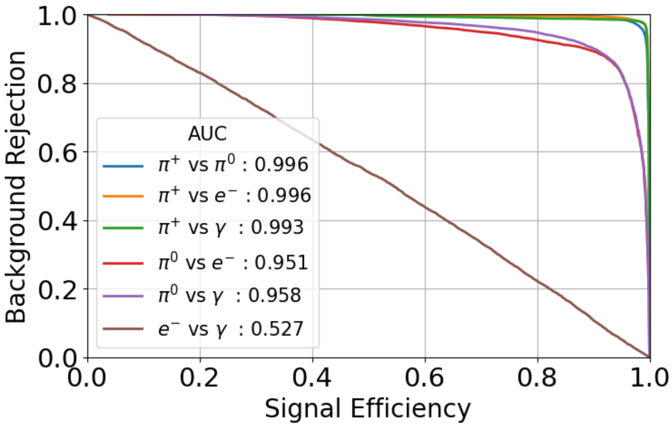


- Energy of shower energy between Geant4 GAN generated shower energy fits at higher energy range.

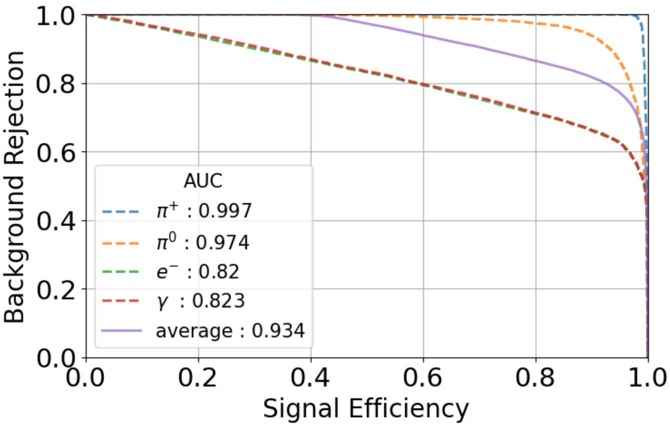
PID performance

- CCT model with point cloud is trained to identify multiple classes π^+, π^0, e^- and gamma.
  - Though π^+ and π^0 have different feature to discriminate, model didn't get biased to specific label.
  - π^0 identification performance reduced than binary classification model..

One vs one ROC curves



One vs rest ROC curves



Readout bundling

- Shower image with 1pixel/1fiber(readout) is ideal case.
  - Massive number of channels is also problem.
  - Readout bundling with multiple fibers applied at test beam.
- Bundling readout to multiple fibers affects positional resolution.
  - positional resolution is crucial to discriminate pi0 and gamma.
- Summed readout value assuming n x n fibers to a readout.
  - Number of pixel decreases at shower image.
  - ϕ and θ index of point cloud are rounded as bundling size
  - Bundling size selected in 1,2,4,8,14 and 28(divisor of 56)

Bundled fibers to readouts[1)]

Bundling at point cloud

θ	3	4	3	3	2	2	2
ϕ	4	4	5	3	4	5	3
S	1.2	0.8	0.7	0.5	0.3	0.2	0
C	1.0	0.7	0.6	0.5	0.4	0	0.2

θ	2	2	2	2	1	1	1
ϕ	2	2	3	2	2	3	2
S	1.2	0.8	0.7	0.5	0.3	0.2	0
C	1.0	0.7	0.6	0.5	0.4	0	0.2

Scintillation channel image

0	0	0	0	0	0
0	0.2	0.7	0	0	0
0	0.3	1.2	0.8	0	0
0	0	0.5	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0

0.2	0.7	0
0.3	2.5	0
0	0	0

Cerenkov channel image

0	0	0	0	0	0
0	0	0.6	0	0	0
0	0.4	1.0	0.7	0	0
0	0.2	0.5	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0

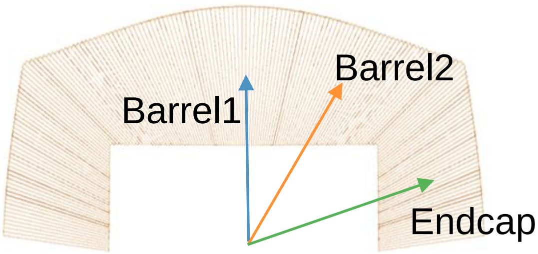
0	0.6	0
0.6	2.2	0
0	0	0

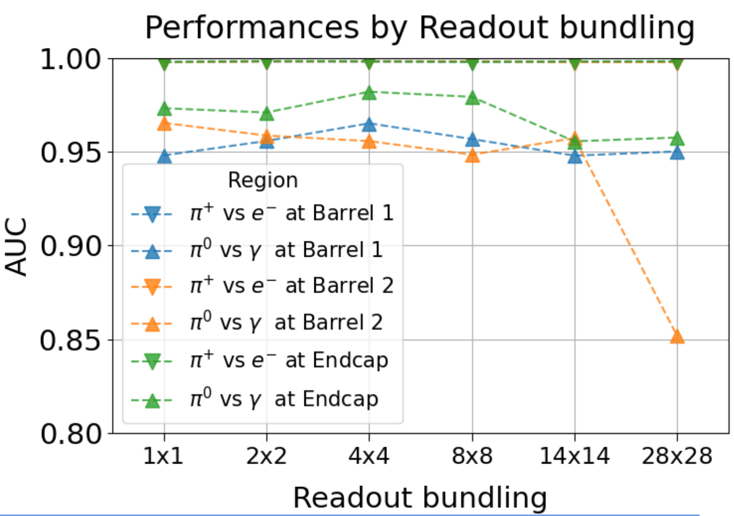
Bundling at shower image

1) Sehwook Lee et al. arXiv:1712.05494

Performance over geometry

- Particle direction for 3 different θ regions.
  - Barrel 1: $\theta = \pi/2$
  - Barrel 2: $\theta = 0.3\pi$
  - Endcap: $\theta = 0.12\pi$
- Compared one vs one score for different fiber bundling
- π^+ vs e^- performances aren't different at readout bundling between 1x1 and 28x28
- π^0 vs gamma performances get unstable at 14x14 and 28x28 bundling.





Summary

- Deep learning implementation has been studied for the dual-readout calorimeter using deep learning.
- Hadronic τ classification can be performed with vision transformer model.
- Jet variable regression will be developed for better jet reconstruction.
- GAN shower fastsim can generate shower image of certain energy.
- PID for π^+,π^0, e^- and gamma with same model showed capable performance.
- Significant performance reduction not found at 8x8 readout bundling which can reduce building cost.
- With these deep learning application, the physics potential of the dual-readout calorimeter will be maximized.

Backups

19

Dual-Readout Calorimeter

- EM shower fraction(f_{em}) is directly measured by scintillation and Cerenkov responses.
  - $f_{em} = 1$ for EM shower, $f_{em} < 1$ for hadron shower
  - Hadronic energy can be measured with better resolution.

$$f_{\rm em} = \frac{(h/e)_C - (C/S)(h/e)_S}{(C/S)[1-(h/e)_S] - [1-(h/e)_C]}$$

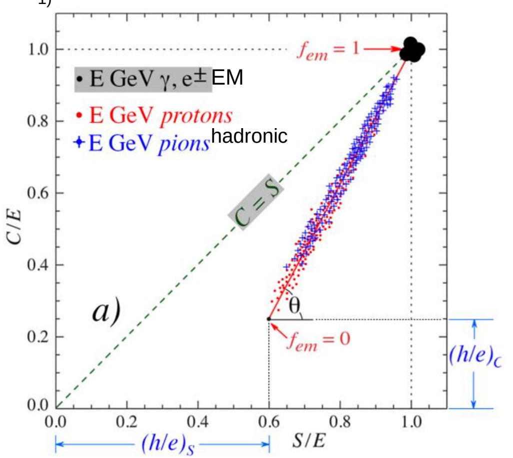
$$E = \frac{S - \chi C}{1 - \chi}$$

$$S = E\left[f_{\rm em} + \frac{1}{(e/h)_S}(1 - f_{\rm em})\right]$$

$$C = E\left[f_{\rm em} + \frac{1}{(e/h)_C}(1 - f_{\rm em})\right]$$

$$\cot\theta = \frac{1-(h/e)_S}{1-(h/e)_C} = \chi$$

Signal from a dual-readout calorimeter [1)]



Signal is calibrated with electron of known Energy E.
Cerenkov signal is smaller than scintillation signal with hadron showers.

1) Lee, S., Livan, M. and Wigmans, R. (2018) Nucl. Instr. and Meth. A882, 148

Model Details

- Environment
  - OS : CentOS7
  - Training with GPU : V100 Nvidia GPU
  - DL library : Keras 2.4.0 with Tensorflow 2.4.0
- Data set division
  - Training 50%, validation 20%, test 30%
- Loss function
  - Categorical cross-entropy for classification
  - Nadam optimizer(learning rate=0.0002)
  - Early stopping applied till 50 epochs training.
- Model
  - 64 channel 4 kernal Conv - 2x2 MaxPool - 64 channel 4 kernal Conv - 2x2 MaxPool – 128 channel – AveragePool : resolution >3x3
  - 64 channel 3 kernal Conv - 2x2 MaxPool - 64 channel 1 kernal Conv - 2x2 MaxPool – 128 channel – AveragePool : 3x3 resolution
- Process time
  - Image 0.7ms/shower 45 mins 50epochs to train enough around 15 epochs get best model

Challenges of realization

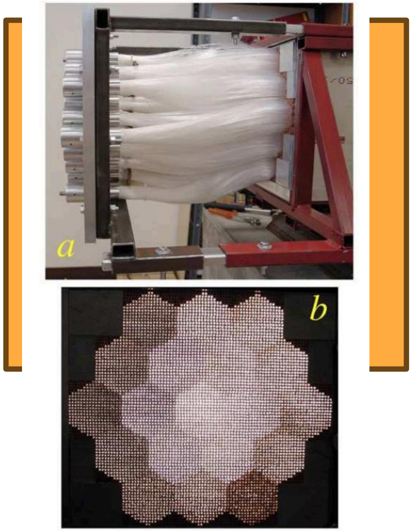
- There are massive number of fibers.

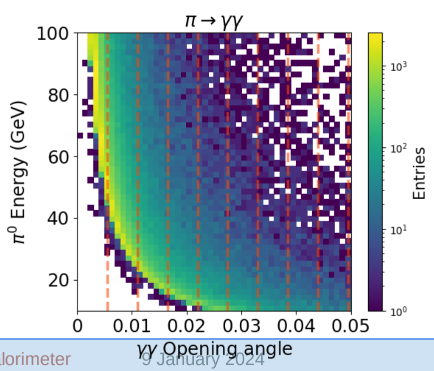
- It is almost impossible attaching single readout to each fiber.
  - Assembling fibers are very hard works but readouts cost more, and we need to deal with massive number of channel also.

- Bundled fiber to single readout is also used at test beam.

- Readout bundling affect positional resolution.

- We need to optimize readout bundling(cost) and performance.





π^0 vs gamma performance

- Binary classifications are performed between π^0 and gamma showers.
  - Model responds to input data between 1.0(true) to 0.0(false).
- Receiver operating characteristics(ROC) curve drawn with signal and background efficiencies.
  - Area under ROC curve(AUC) implies how efficiency is close to 1.0
- CCT model with point cloud format has better performance with than with image format.