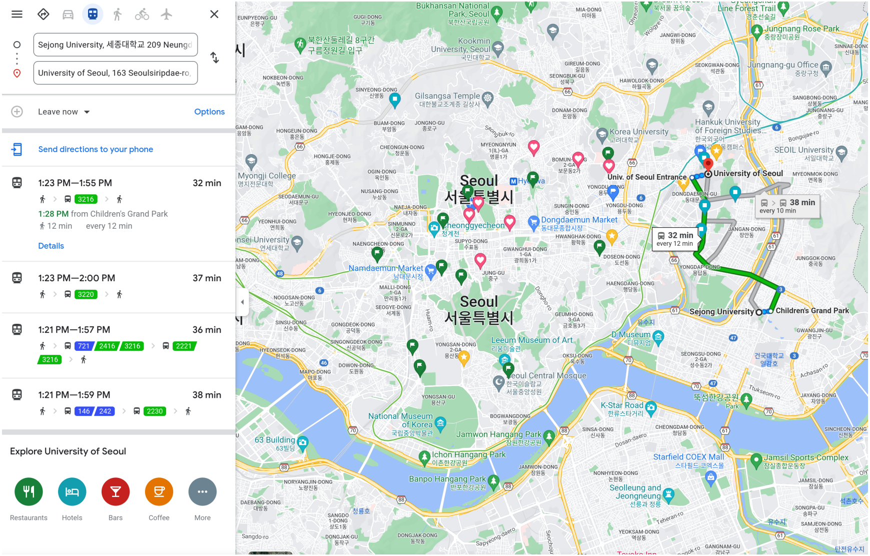
Machine Learning at the CPLUOS Laboratory at the University of Seoul

Ian J. Watson for the CPLUOS Lab Team

University of Seoul

Workshop for Korea-UK AI/ML Research in Fundamental Sciences
Sejong University, Seoul, November 4, 2022

Sejong University to the University of Seoul



연구실 현원 : 교수 / 연구원

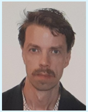
교수


민현수 박동수

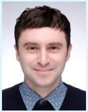

박인규 이상훈

연구교수 / 포스트닥 / 병역특례연구원

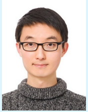
박상남
2018



구스타프슨
2018



Cris Sabiu
2020

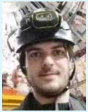


노창동
2020

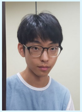
김동현
2020

Ian Watson
2017

노연정
2020



J. Merlin
2021



이윤재
2017

장우진
2017

스태프

장세덕
2016

김휘영
2021

엔지니어 (검출기)

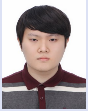
박성호
2015

정영균
2012

이협우
2019

엔지니어 (컴퓨팅)

조영권
2019



김남수
2018

연구실 현원 : 대학원생

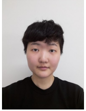
우주론


지한나 2019


구현모 2019

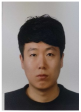

주영 2021


황세연 2021


김수미 2020

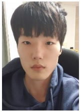
입자물리 (+ 천체)


고병학 G2016


양승진 G2018


손영완 2018


김도영 2020


조백선 2021

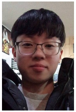
입자물리 (+ 검출기)


송동현 G2017

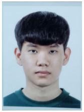

강예찬 G2017


김슬기 G2018


강다영 G2018


최명훈 G2021


김정우 2015


허지원 2020

응용물리


황인 G2021


제태성 G2021

학부생 인턴

김연주 2020

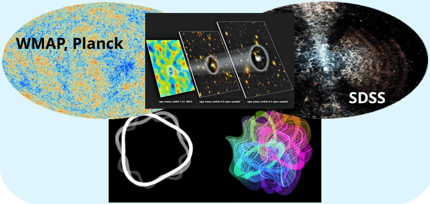
박진혁 2021

조성휘 2021

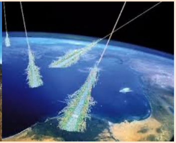
이정민 2021

주요 연구 내용

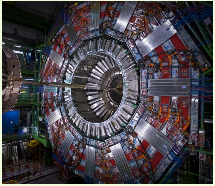
우주론 , 장론 (끈이론)

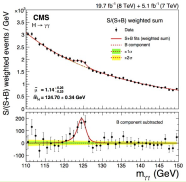


천체 물리 실험

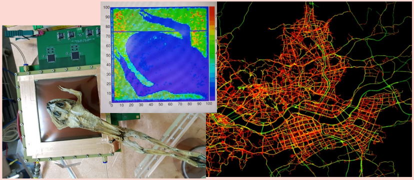


핵 / 입자물리 실험 (CMS 국제공동연구)



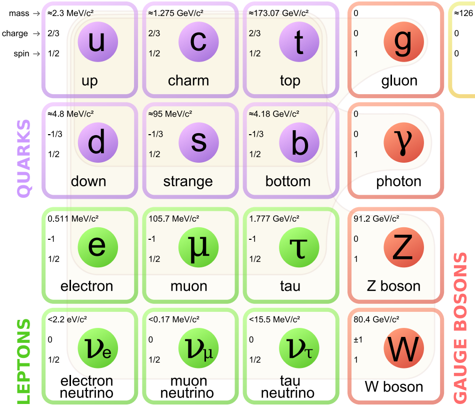


검출기 개발 , 응용물리 (빅데이터 AI)



Overview of Projects for Today

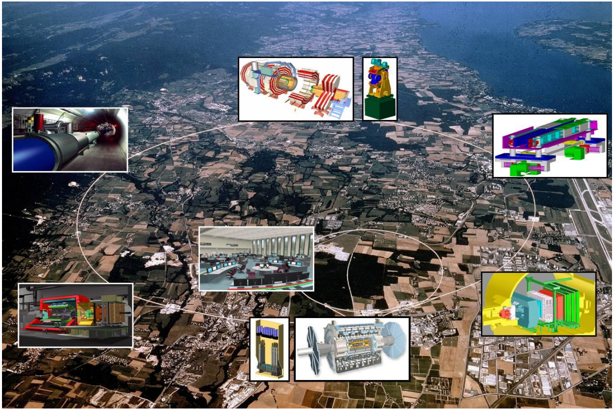
- LHC
- Future Colliders
- HAWC
- Machine Learning for Cosmology

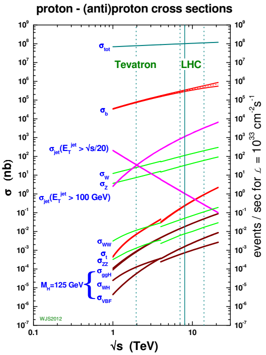


- Forces (exchange particle):
  - Strong (gluon)
  - Electromagnetism (photon)
  - Weak (W/Z)
- Higgs Boson - Responsible for mass of the particles
- Quarks: Interact via the strong, weak and EM forces
- Leptons: Interact via the weak and EM forces

- The "Standard Model" of particle physics
- Some issues with this model:
  - What is the Dark Matter that seems necessary from observation?
  - Why is there more matter than anti-matter in the universe?
- We are probing nature to try to break the model with the LHC

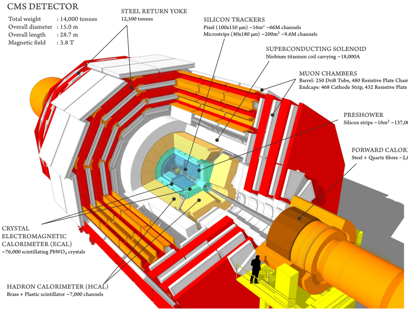
Context: LHC

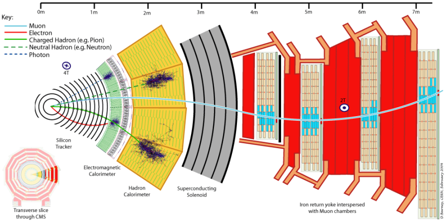




- 4 experiments on the LHC proton-proton collider ring: CMS, ATLAS, LHCb, ALICE
- At UoS we work at CMS: Bunch crossing rate of ~ 40 MHz, from which typically ~ 1 kHz stored to disk for further analysis
  - Each bunch crossing has dozens of "pileup" collisions
  - From this data haystack, searching for extremely rare processes

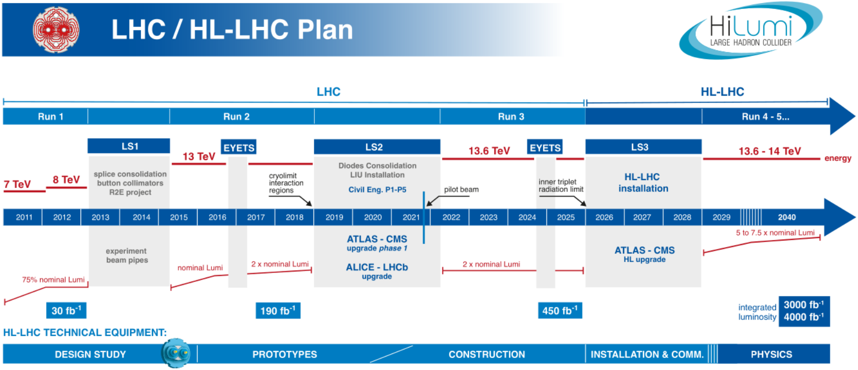
Context: CMS





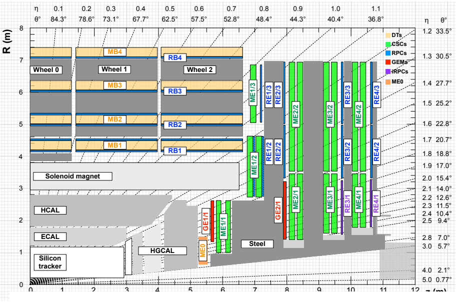
- 14,000 tonnes, 28.7 meters long, 15 meters in diameter
- Instrumented with detectors arranged in an onion structure
  - Provides particle reconstruction and identification ($e^{\pm}/\gamma$ vs hadronic vs $\mu^{\pm}$), particle momentum and energy measurements
- Used to find the Higgs boson (along with ATLAS), probing the Standard Model of Particle Physics and searching for new physics beyond the Standard Model

Context: HL-LHC

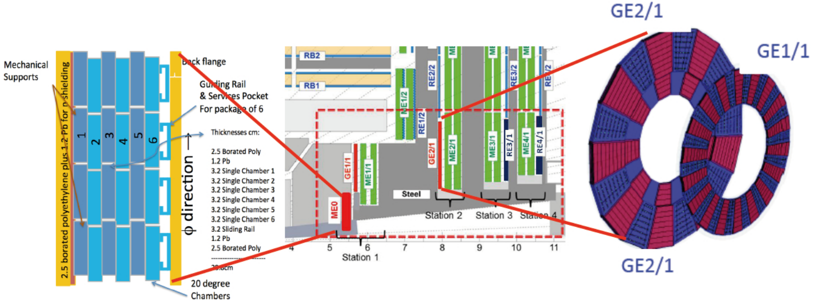


- Over the decades we willl increase luminosity
- Expect to collect about 30x more data than we currently have
  - Go from $\approx$ 500 PB of stored data total, to several EBs
- We will need a commensurate increase in the amount of simulation, ability of data acquistion, speed of user analysis, etc.
  - Could say similarly for e^+e^- (B2), neutrinos (DUNE), etc.
- HL-LHC potentially followed by FCC, takes the roadmap to the 2090s

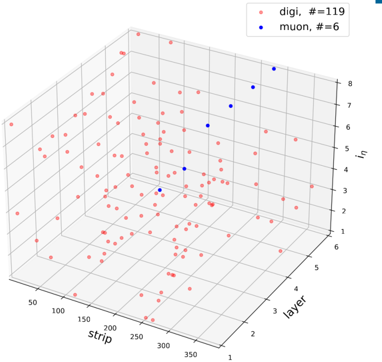
ME0 + GE[12]/1 Geometry Overview

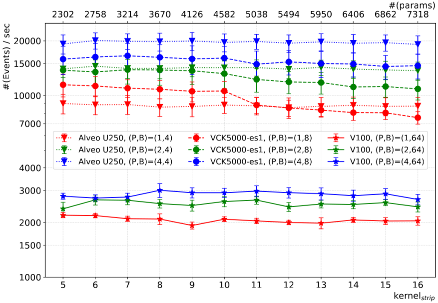


- UoS works on the GEM (Gas-Electron Multiplier) detectors
- Being installed for the LHC upgrade
- In particular, ME0 is a new, very forward detector
    - Forward = close to the beam line, where most of the "pileup" particles will spray in



ME0 Segment Finder

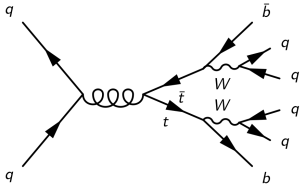


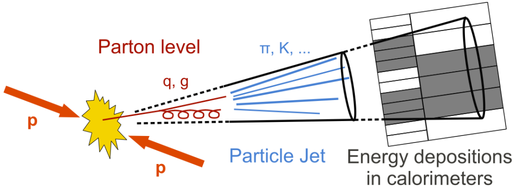


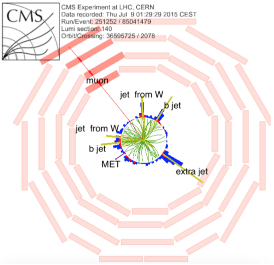
- Muons are very important for many studies, and a very clean signal of an "interesting" event
- Need to be able trigger on events in μs
  - Need to go from 40 MHz $\rightarrow$ O(1 kHz)
- We want to use image segmentation CNN on detector in the time budget, so we tried putting it on an FPGA
  - Using the Vitis-AI framework, which also allows quantization

Youngwan Son

Jet Assignment in Top Production



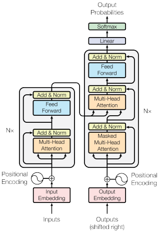


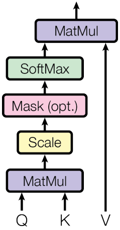


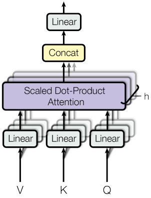
From CMS Top Events Displays

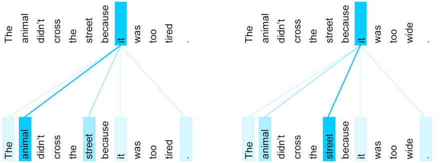
- t is the heaviest (known) fundamental particle
- It decays into a quark and W, the W can decay to a quark pair or lepton-neutrino pair
  - Quarks and gluons found as a collimated *hadronic jets* in data
- Usually produced in top–anti-top pairs, but we don't know which jets to assign to which top
- The all quark decay is hard to distinguish from simple quark/gluon production

Attention Mechanism





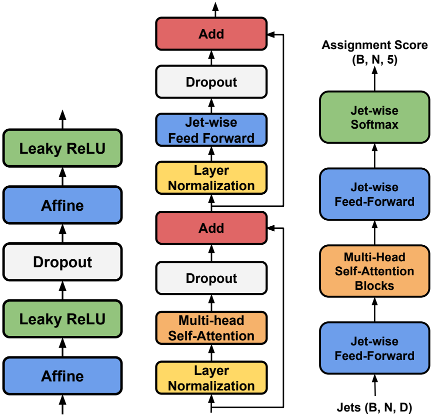




- The transformer model produces output sequence from input sequence
- Uses self-attention layers to model the dependencies between elements of the sequences: each word is trained to "attend" to related words
  - This replaces the previous recurrent models which processed single tokens at a time to produce a single context vector
  - Example showing a transformer which connects grammatical elements
- The transformer model is at the heart of modern deep learning language processing
  - Ever since it was argued that Attention is All You Need in 2017
- Achieved astonishing performance in applications like GPT-3

https://ai.googleblog.com/2017/08/transformer-novel-neural-network.html

SAJA (Self-Attention for Jet Assignment) Network



Jet-wise FF Attention SAJA

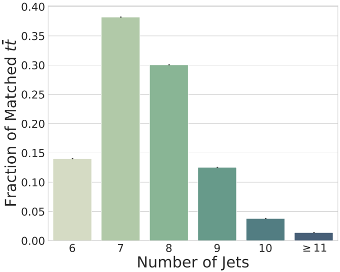
B is *batch size*, N is number of jets **which can be different for each event**

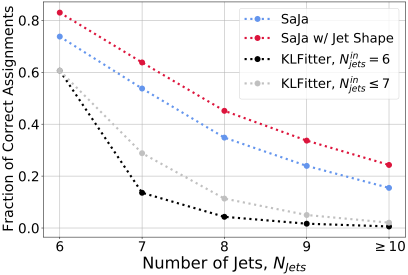
Seungjin Yang

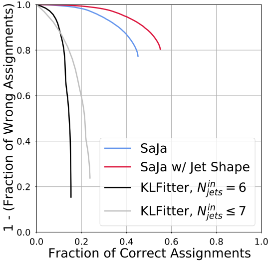
- Jet-wise network acts per jet
- Its main component is the Attention block
  - Multihead attention layer is the only place where the jets interact with each other
- SAJA takes the D-length vector of N jets applies the attention mechanism and output a score for each jet
  - Classifies each jet according to the role in the underlying process
  - We will do $t\bar{t}$, so is it b-jet, jet from W, or *other*, e.g. gluon radiation

arxiv:2012.03542

Results compared to classical method

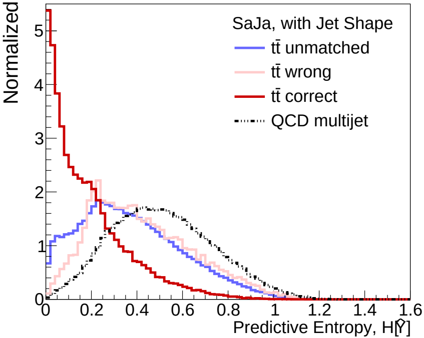






- Here correct means classifies **all** the jets in the event correctly
- Consistently beating the classical likelihood based method (restricted in # of jets)
- Even in very high number of jets, where likelihood fails entirely, is able to classify some subset of events

Jet assignment to event interpretation



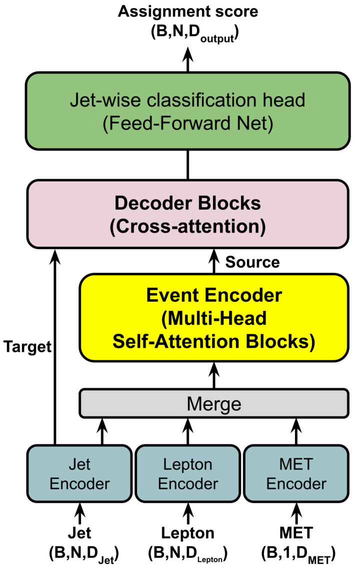
Jet#	b_1	W_1	Score b_2	W_2	other	assign
1	0.4	0.3	0.2	0.1	0	b_1
2	0.2	0.1	0.4	0.2	0.1	b_2
3	0.1	0.0	0.1	0.2	0.6	other
...						

$$H = \frac{1}{N}\sum_{j=1}^{N}\left(-\sum_{c\in\text{classes}} \hat{y}_c^{(j)} \log \hat{y}_c^{(j)}\right)$$

$\hat{y}_c^{(j)}$: prediction of network for j^{th} jet to be in the c^{th} category

- The output of SAJA is a table: for each jet a score is assigned to each category, normalized to 1 over all the categories
- For each jet, we assign the jet as the category with the highest output
- Take the average of the entropy of the output of each jet, H
- High entropy implies the network is uncertain about the assignment
- The network is only trained to reconstruct $t\bar{t}$ but we get QCD rejection for free!

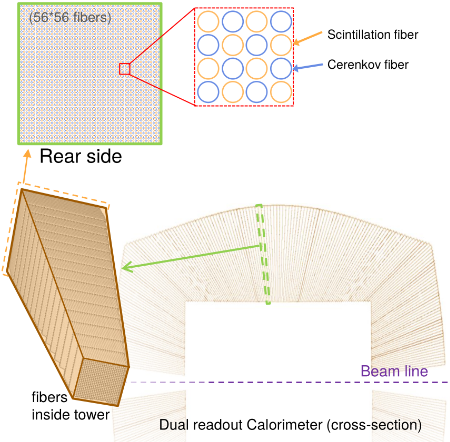
SAJA in Dilepton Events

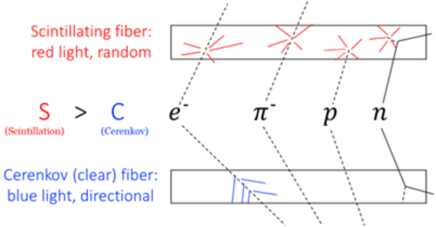


Jeewon Heo

- Also trying building a SAJA model for jet assignment in dilepton events $tt \to b\ell\nu b\ell\nu$
- Here we combine the heterogenous reconstructed object types together, apply the attention to all of them, then extract the assignments
- Work in progress for another of our projects trying to find the rare decay $t \to sW$ [Woojin Jang (arxiv:2112.01756), Jeongwoo Kim]

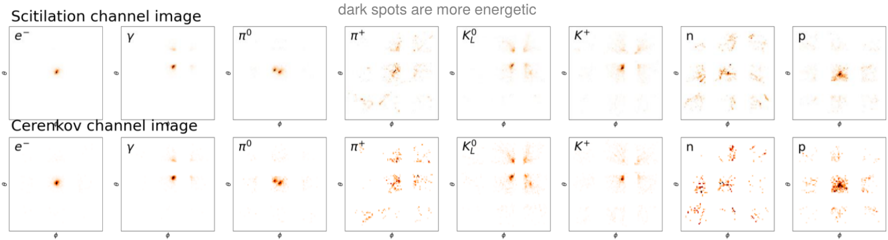
Intro to the Dual Readout Calorimeter

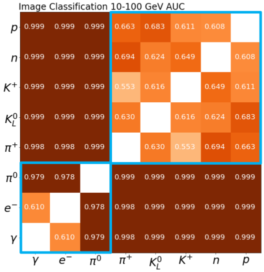




- The Dual-Readout Calorimeter is a new detector concept we are working on in Korea
  - Novel design using two different types of readout fibers to distinguish EM from hadronic particles
  - In design and testing phase, being considered for the FCC

Particle Id for the DRC for FCC

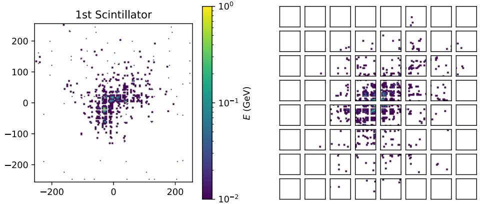




Yunjae Lee

- We are working on trying to apply deep learning to distinguish particles using just the DRC
- Separate into "EM"-type and "hadronic"-type based on interactions in the calorimeter
- Excellent separation between EM and hadronic showers, some separation within classes

Tau Identification for FCC



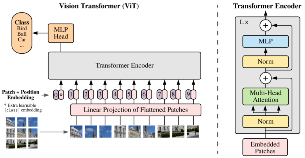
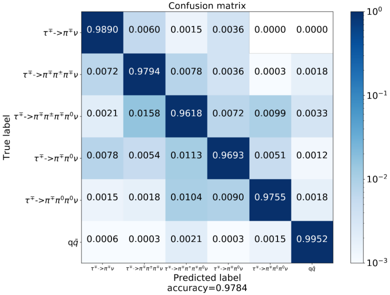


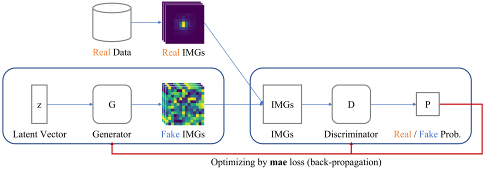
Figure 1: Model overview. We split an image into fixed-size patches, linearly embed each of them, add position embeddings, and feed the resulting sequence of vectors to a standard Transformer encoder. In order to perform classification, we use the standard approach of adding an extra learnable "classification token" to the sequence. The illustration of the Transformer encoder was inspired by Vaswani et al. (2017).



- Tau's or τ's are essentially very heavy electrons, they decay into electrons muons or hadrons
- Hadronic taus look very similar to hadronic jets
- Trying to build a specialized τ finder using a Vision Transformer-based network
- Show very good rejection in Z boson simulation

Youngwan Son

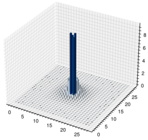
GANs for FCC



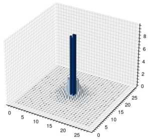
Generative adversarial network (GAN)

- Simulation is taking increasing amounts of time as the detector size and simulation granularity increases
- We are trying to create a GAN to do a fast simulation of the DRC
- Produces detector response output image from information about the incoming particle
- Orders of magnitude speed up using a GAN on GPU vs GEANT4 (the full, detailed simulation)

3D plot of average img

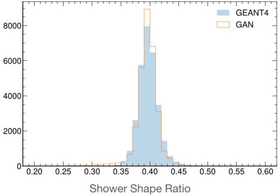


Average image of **gan** output



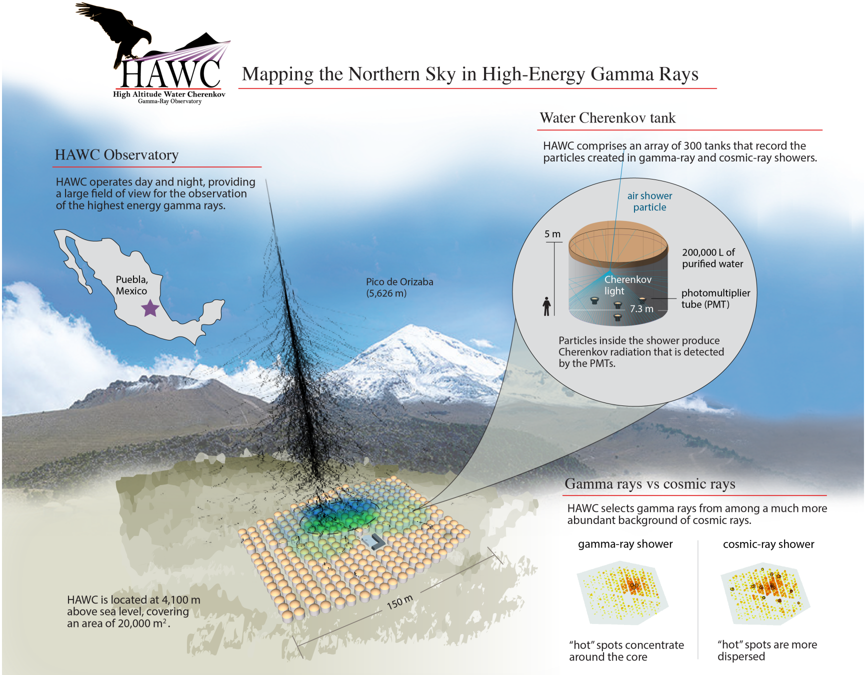
Average image of **geant4**

Doyeong Kim

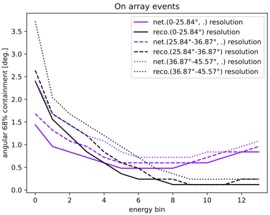


Energy	Type	Machine	Latency (s)
10 GeV	Geant4	CPU	1023.67
	GAN	GPU	0.45
50 GeV	Geant4	CPU	4671.63
	GAN	GPU	0.45
100 GeV	Geant4	CPU	9564.91
	GAN	GPU	0.45

Computing Performance

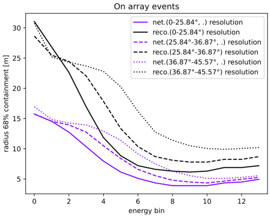

HAWC
High Altitude Water Cherenkov
Gamma-Ray Observatory
Mapping the Northern Sky in High-Energy Gamma Rays
HAWC Observatory
HAWC operates day and night, providing a large field of view for the observation of the highest energy gamma rays.
Puebla, Mexico
Pico de Orizaba (5,626 m)
Water Cherenkov tank
HAWC comprises an array of 300 tanks that record the particles created in gamma-ray and cosmic-ray showers.
air shower particle
5 m
200,000 L of purified water
Cherenkov light
photomultiplier tube (PMT)
7.3 m
Particles inside the shower produce Cherenkov radiation that is detected by the PMTs.
150 m
HAWC is located at 4,100 m above sea level, covering an area of 20,000 m².
Gamma rays vs cosmic rays
HAWC selects gamma rays from among a much more abundant background of cosmic rays.
gamma-ray shower
cosmic-ray shower
"hot" spots concentrate around the core
"hot" spots are more dispersed

Deep Learning for HAWC




```
# unit: log(Energy/GeV)
0: 0 <= mclogEnergy < 2.5
1: 2.5 <= mclogEnergy < 2.75
2: 2.75 <= mclogEnergy < 3.0
3: 3.0 <= mclogEnergy < 3.25
4: 3.25 <= mclogEnergy < 3.5
5: 3.5 <= mclogEnergy < 3.75
6: 3.75 <= mclogEnergy < 4.0
7: 4.0 <= mclogEnergy < 4.25
8: 4.25 <= mclogEnergy < 4.5
9: 4.5 <= mclogEnergy < 4.75
10: 4.75 <= mclogEnergy < 5.0
11: 5.0 <= mclogEnergy < 5.25
12: 5.25 <= mclogEnergy < 5.5
13: 5.5 <= mclogEnergy < 10.0
```

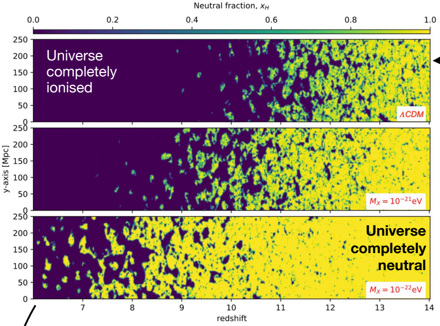
mclogEnergy: True energy of primary gamma-ray



- Working on two distinct problems:
  - Identify γ rays from the overwhelming cosmic ray background
  - Identify the direction and energy of the incoming γ ray
- Trying out a ViT trained on these two problems separately
- Showing some results from the incoming angle regression output
  - At lower energies, we do better than the standard offline reconstruction, but worse at higher energies
  - γ events have a steeply falling spectrum with energy (and are also simulated this way), so this is likely a data imbalance issue
  - We are trying reweighting the simulation w.r.t. energy to try to improve this overall, but so far no luck

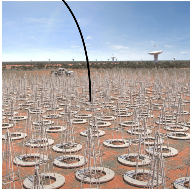
Baeksun Cho, Myeonghun Choi, see also Baeksun's talk yesterday for particle ID

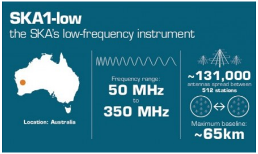
Constraining Ultra Light Axion-like Dark Matter in the Early Universe from Future Radio Survey Data



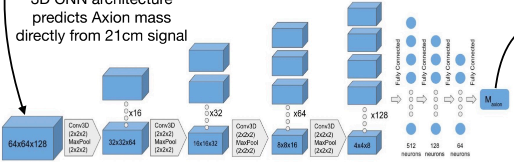
Lighter Axions delay the formation of ionised 'bubbles'

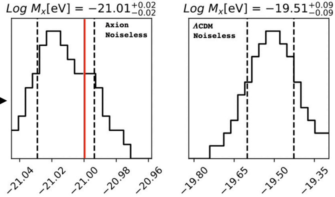
SKA will detect 21cm signal in the early universe, probing the density of neutral hydrogen





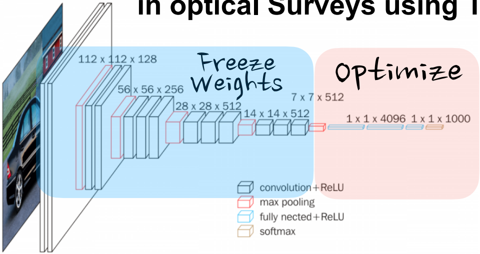
3D CNN architecture predicts Axion mass directly from 21cm signal

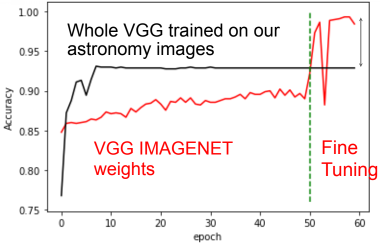


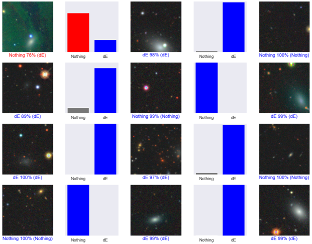


(Sabiu, Kadota, Asorey, Park 2022 JCAP)

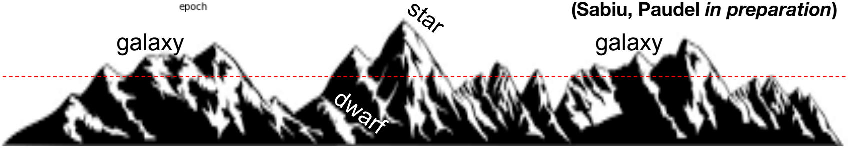
Detecting Low Surface Brightness Galaxies in optical Surveys using Transfer Learning



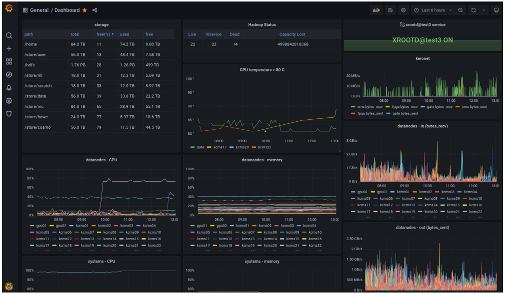




(Sabiu, Paudel *in preparation*)



Summary



- The CPLUOS lab at University of Seoul is working on several ML topics using deep learning