Machine learning in nuclear and particle physics

Christophe Rappold

IEM – CSIC

27 January 2025

Reminder from Part 1

- Machine learning is
  - Extracting semantic information with parametric models
    → Learning = tuning the many parameters of the model
  - Finding patterns / associations → predictions
  - 3 types of methods:
    - Reinforcement / Unsupervised / Supervised learning
- Main Algorithms for supervised learning:
  - Decision Tree: Decision trees / Random forests / XGBoost
  - Neural networks: FFNN, CNN, RNN, Transformers …
- How ?
  - optimization problem: find optimal weights
    - → careful of Underfitting / Overfitting

Outlook

- ML : S-B separation & particle identification
- CNN : alpha decay in emulsion
- GNN : particle tracking
- GAN : simulate emulsion reaction
- Mask R-CNN : hypernuclei finding in emulsion
- Segmentation: full digitization of nuclear emulsion data

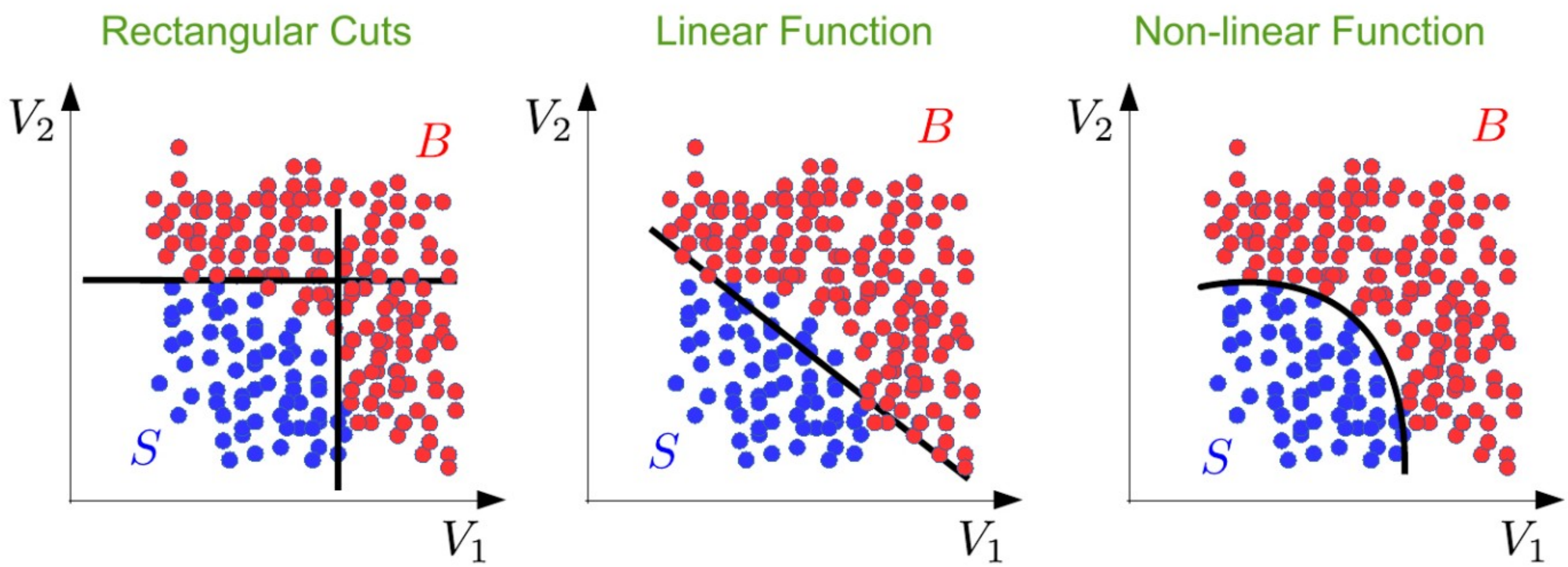
ML for
1. S-B separation
2. Particle Identification

Multivariate analysis (MVA)

- We have data S and B described by discrete variables

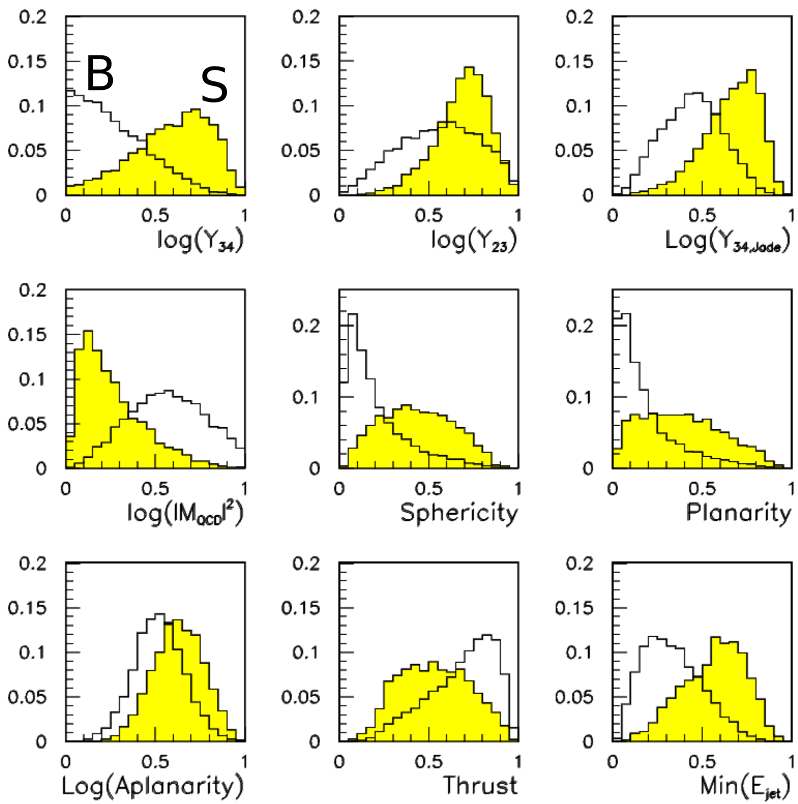
  → Separating S and B
  - Classification of measurements using a set of observables (V_1 … V_n)

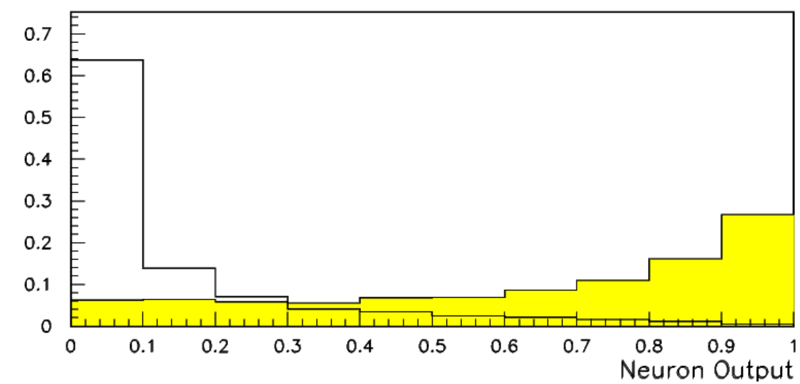
  → Find optimal separation conditions considering correlations

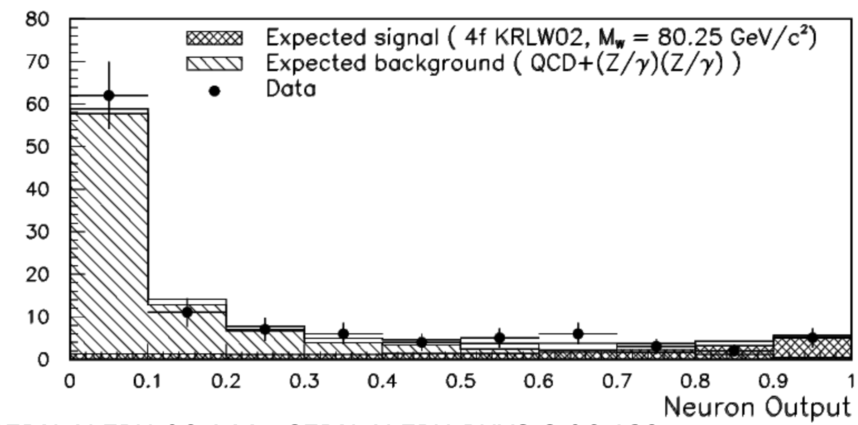


Signal / Background separation: Neural network

- WW cross section (1996): S: e+e− → W+W− | B: e+e− → qqgg
  - Input variables based on jet structure, event shape etc



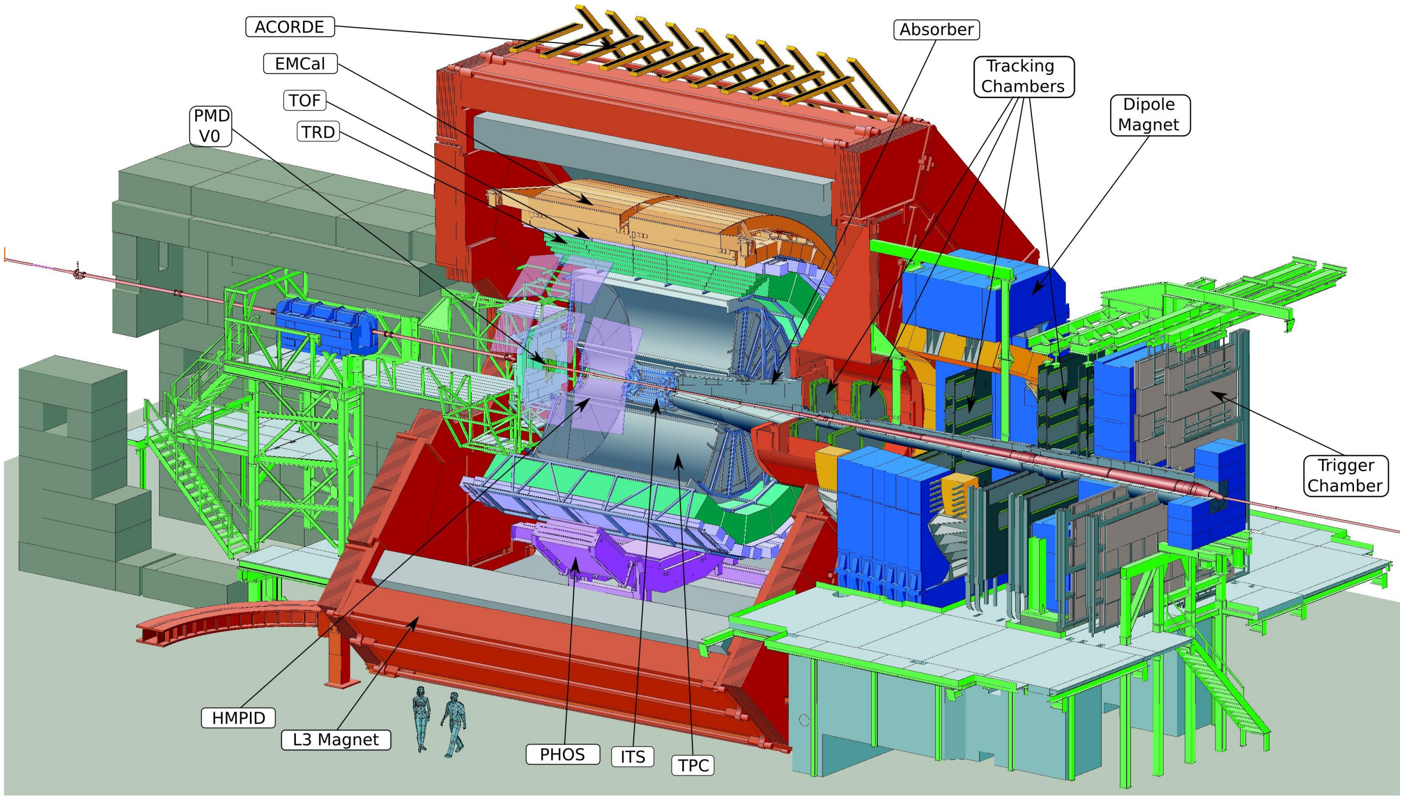




CERN-ALEPH-96-144 ; CERN-ALEPH-PHYSIC-96-132

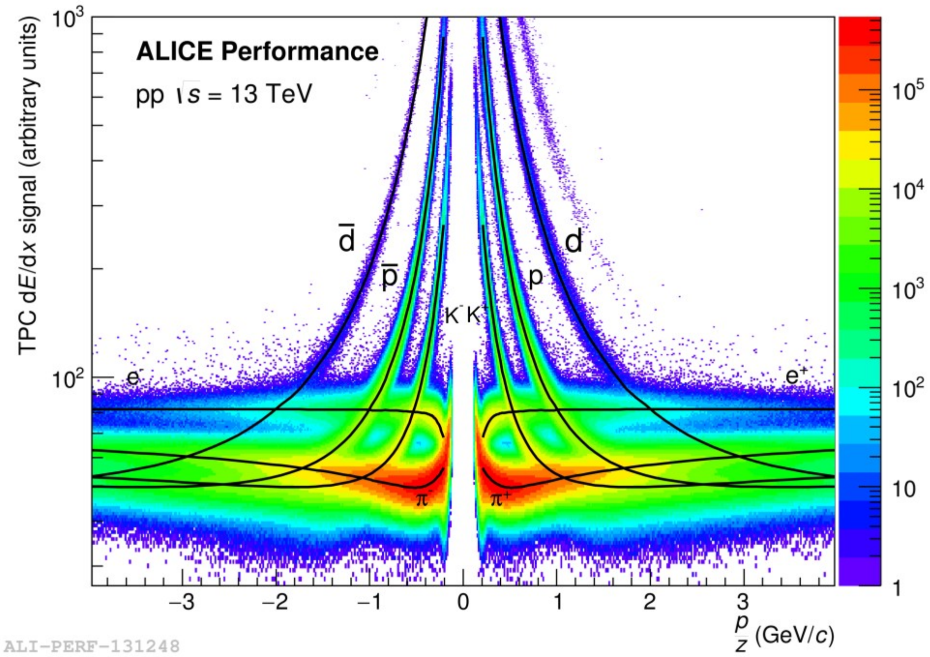
Particle identification : Random Forest

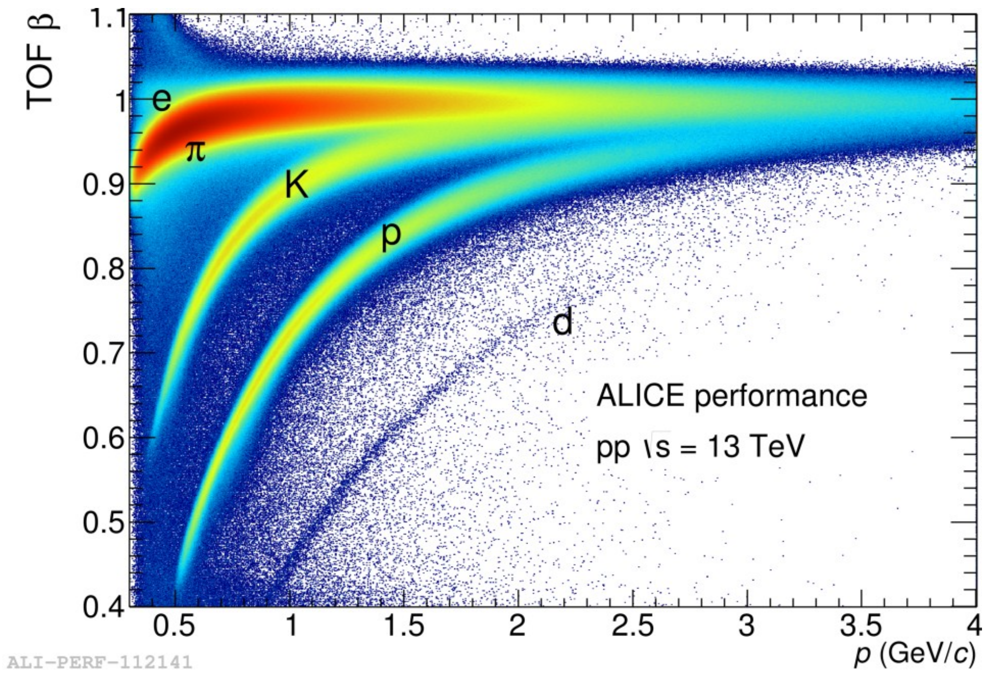
- ALICE experiment: Study of QGP → conditions just after Big Bang
  - PID with TPC and TOF



Particle identification : Random Forest

- ALICE experiment :
  - PID with TPC and TOF



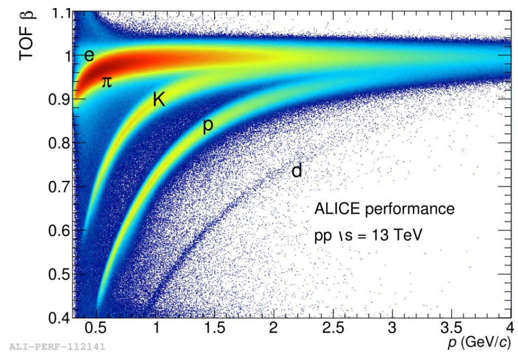


Particle identification : Random Forest

- Models of dE/dx vs p/q & β vs p/q

$$\langle -\frac{dE}{dx}\rangle = K z^2 \frac{Z}{A}\frac{1}{\beta^2}\left[\frac{1}{2}\ln\left(\frac{2 m_e c^2 \beta^2 \gamma^2 Wmax}{I^2}\right)-\beta^2-\frac{\delta(\beta\gamma)}{2}\right]$$

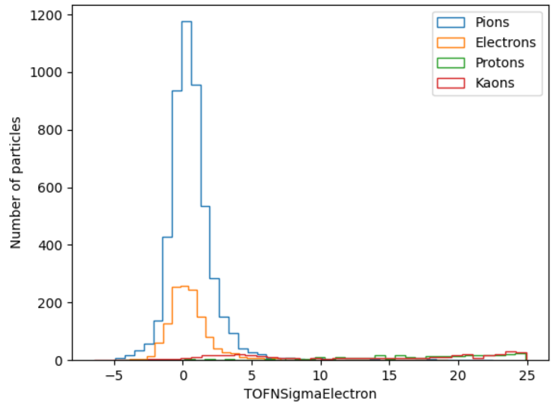
$$\beta=\frac{1}{\sqrt{m^2/p^2+1}}$$

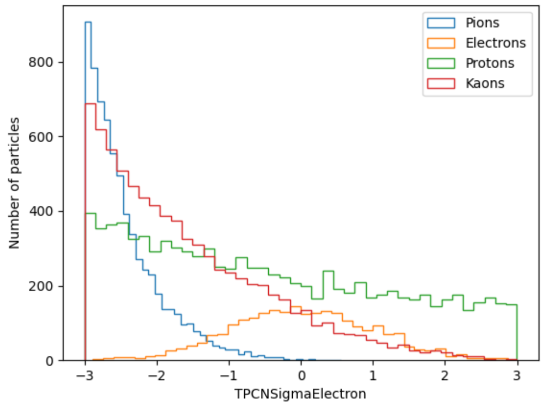


- Considered features:

$$TOF\ N\sigma=\frac{TOF^{measured}-\langle TOF^{particle}\rangle}{\sigma_{TOF}} \qquad dE/dx\ N\sigma=\frac{dE/dx^{measured}-\langle dE/dx^{particle}\rangle}{\sigma_{dE/dx}}$$

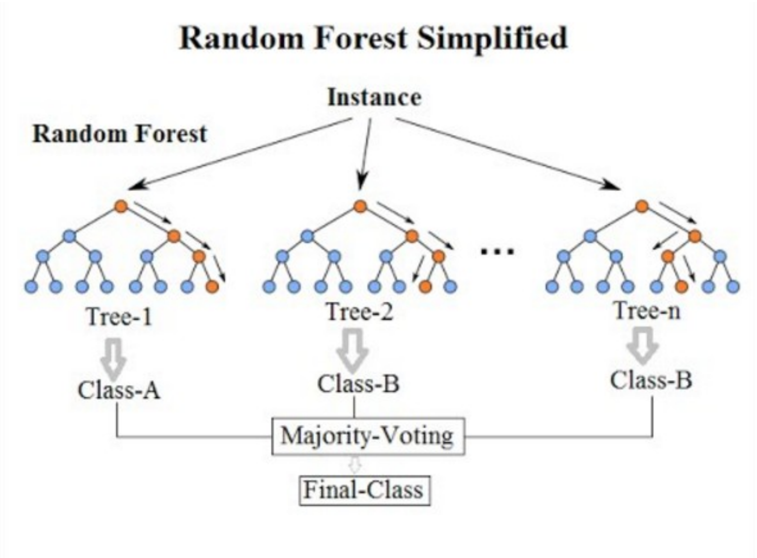
  - Multiplicities in detectors
  - DCA to primary vertex





Particle identification : Random Forest

- Random Forest :
  - Create Decision Trees :
    - Each decision tree → optimized on a random subset of features & *only* access to a random set of the training data
    - increases diversity in the forest → more robust prediction
  - Final classification → vote

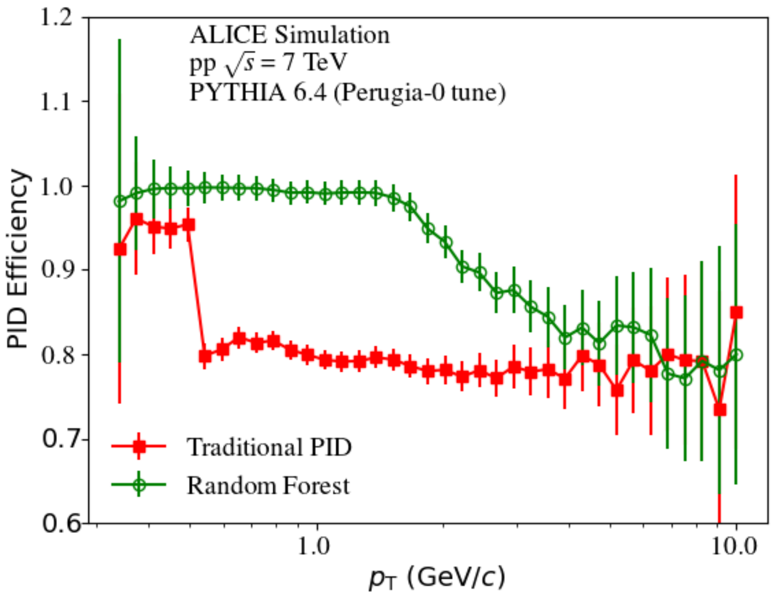


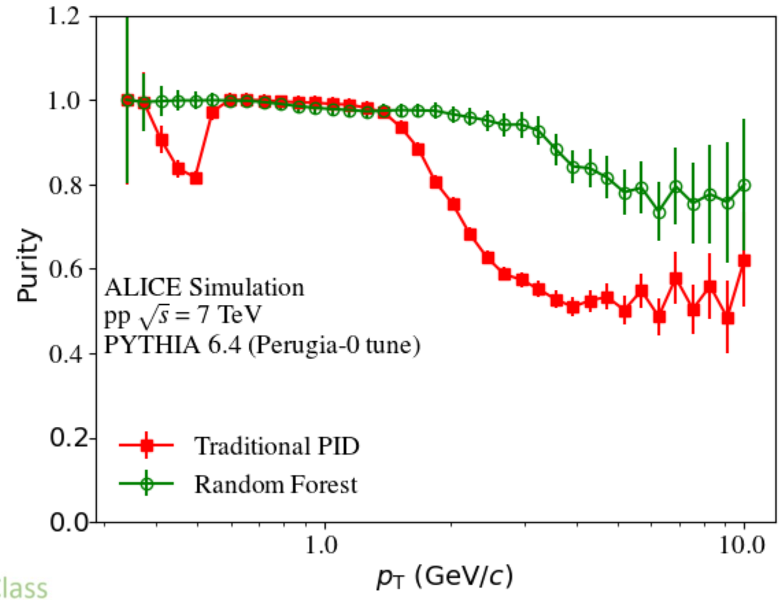
Particle identification : Random Forest

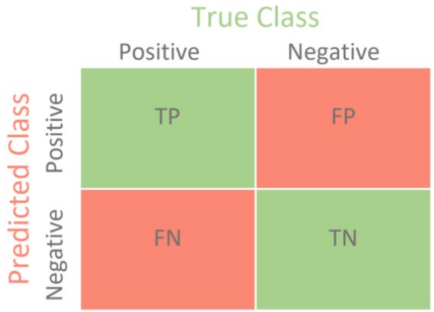
- Results:

ALICE, T. Trzciński et al., AISC 945 (2019) 3

Kaon class







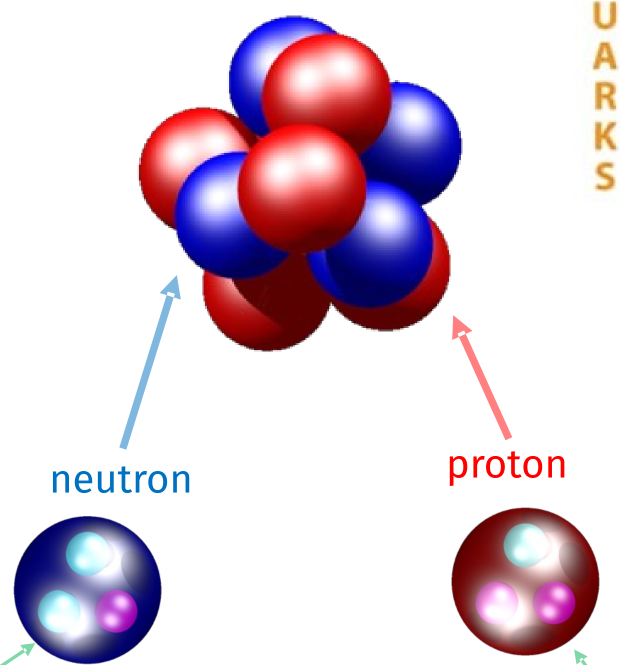
Efficiency = TP / P

Purity = TP / (TP+FP)

CNN for
alpha decay in nuclear emulsions

My main research topic

- Hypernuclear study:



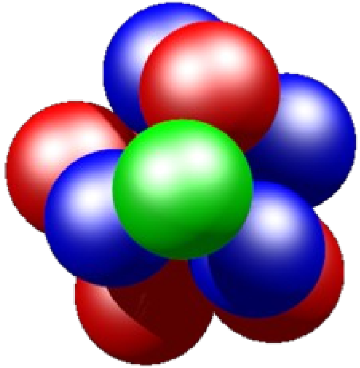
Hyperon	Quarks	$I(J^P)$	Mass (MeV)
Λ	uds	$0(1/2^+)$	1115
Σ^+	uus	$1(1/2^+)$	1189
Σ^0	uds	$1(1/2^+)$	1193
Σ^-	dds	$1(1/2^+)$	1197
Ξ^0	uss	$1/2(1/2^+)$	1315
Ξ^-	dss	$1/2(1/2^+)$	1321
Ω^-	sss	$0(3/2^+)$	1672

hyperon (Λ)

Lifetime~ 10^{-10} ps

s-quark : distinguishable from u- and d-quark

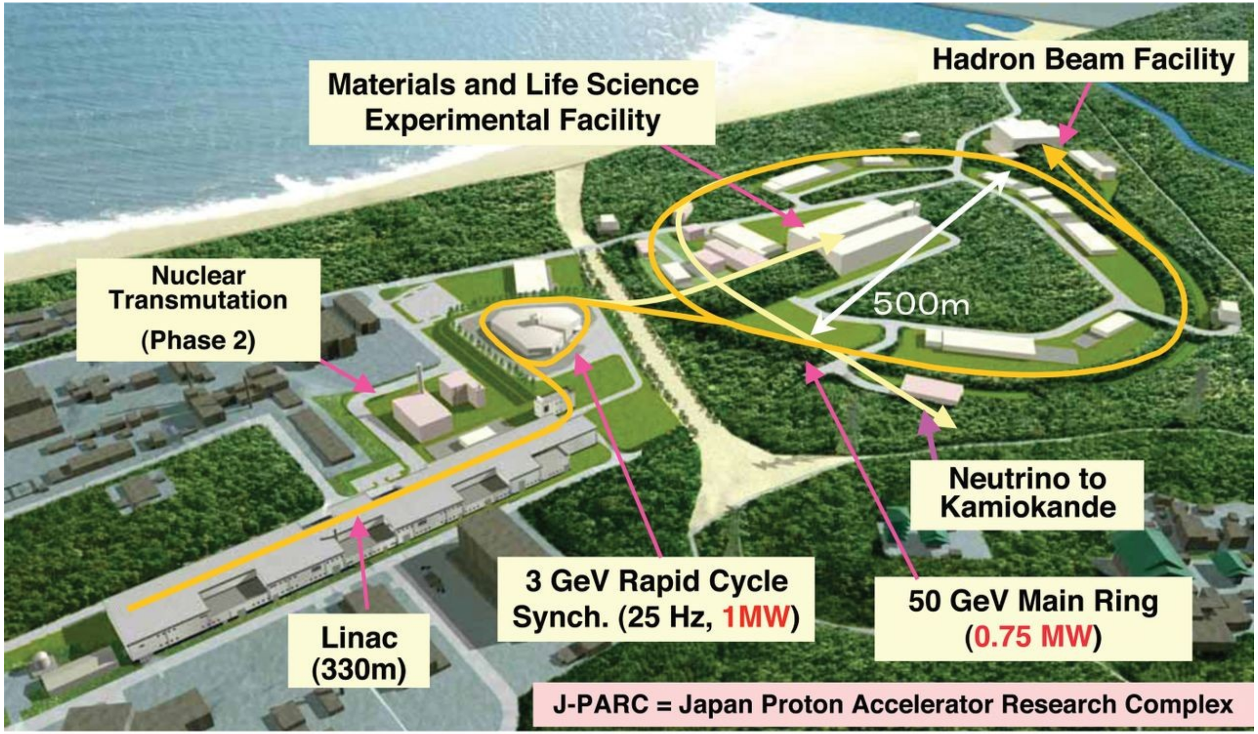
Hypernucleus



Micro-lab for study Baryon interactions

Nuclear emulsion for double strangeness hypernuclei

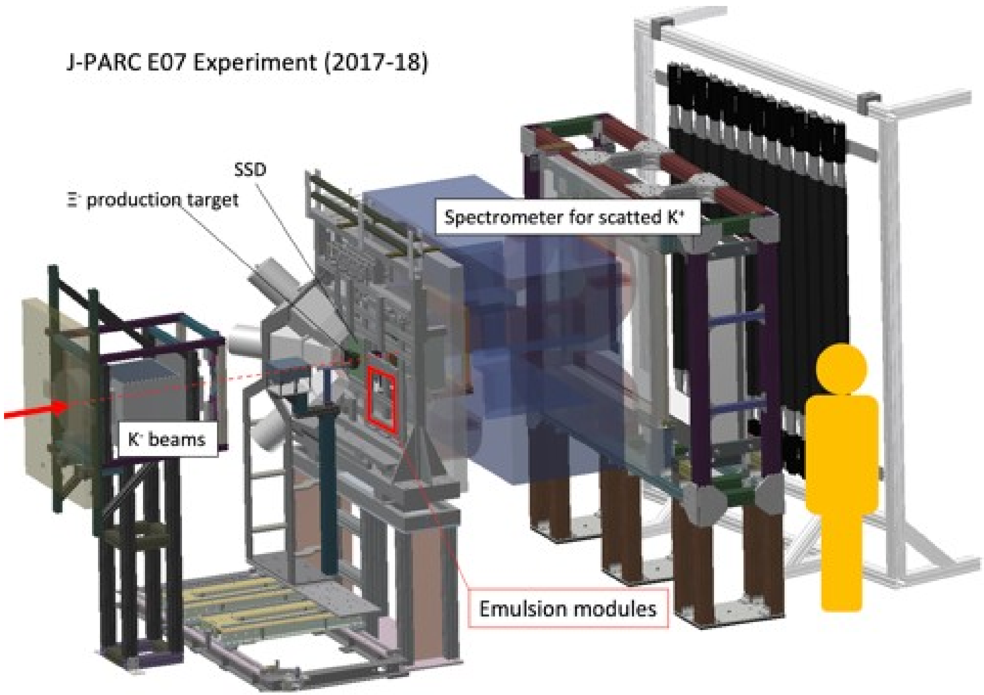
- **J-PARC E07 experiment :**
  - J-PARC : Japan Proton Accelerator Research Complex



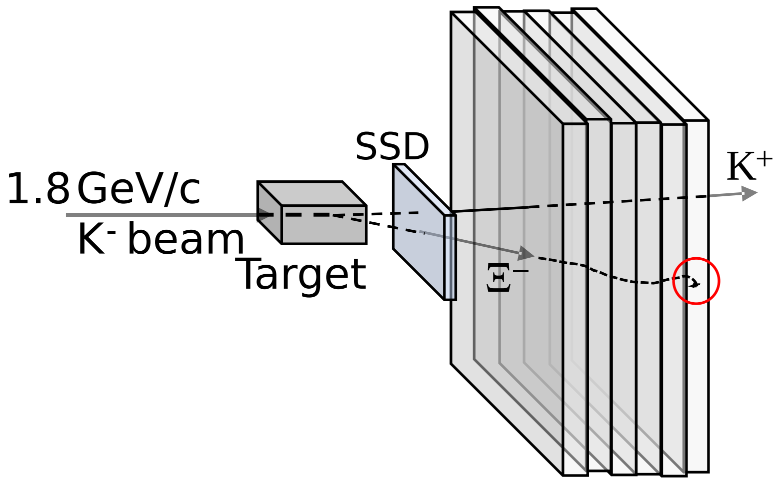
Joint Project between KEK and JAEA

Nuclear emulsion for double strangeness hypernuclei

- J-PARC E07 experiment
  - Study of double-strangeness hypernuclei
  - **Hybrid method**: Triggered detectors + nuclear emulsions

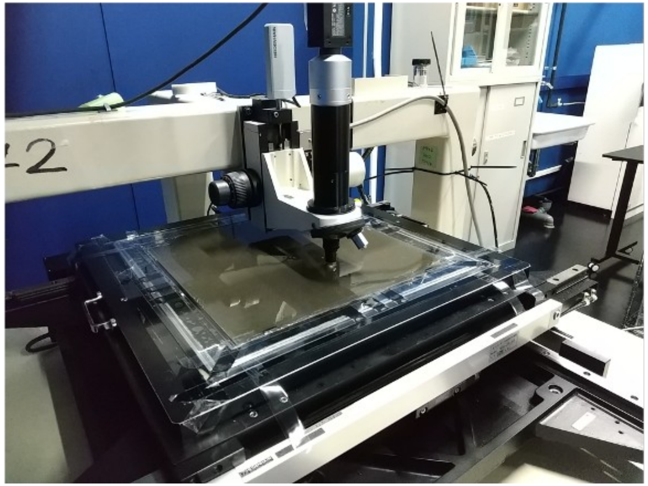


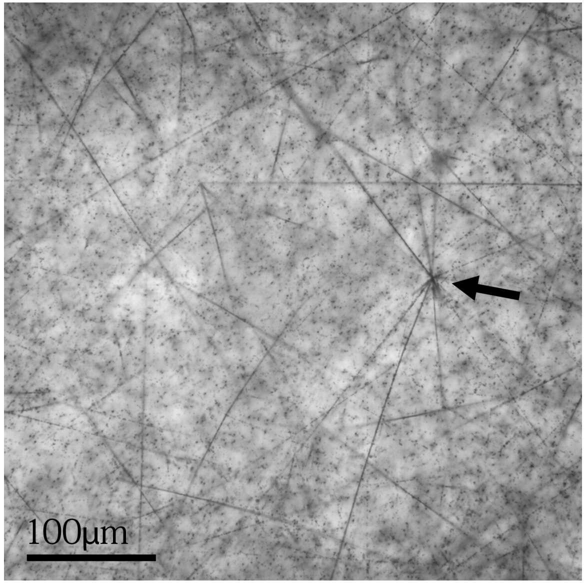
Triggers by the observation of (K-, K+) reactions



Nuclear emulsion for double strangeness hypernuclei

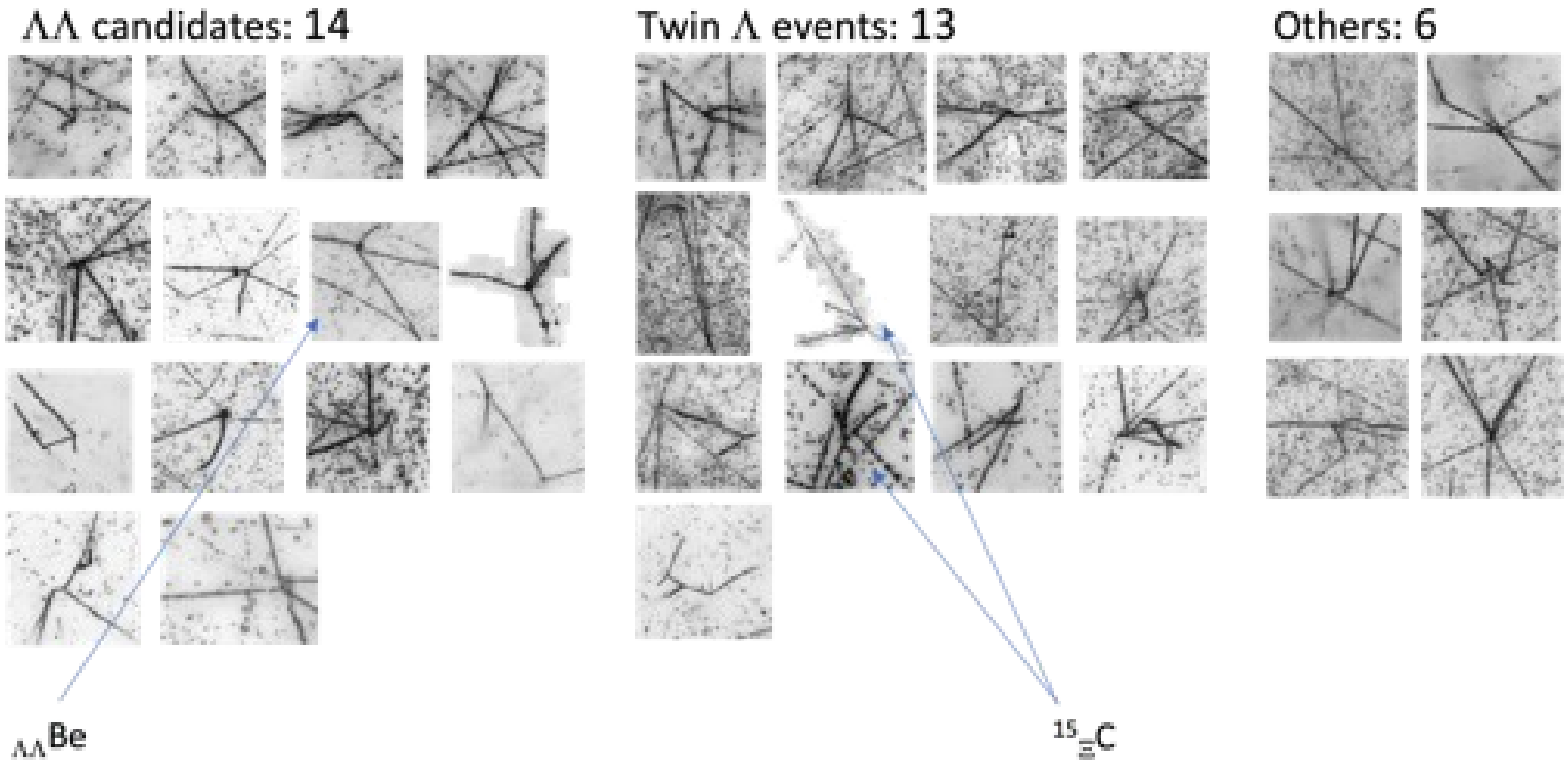
- Scanning methods :





Nuclear emulsion for double strangeness hypernuclei

- Current outcome of E07:



H. Ekawa et al., Prog. Theor. Exp. Phys. 2019, 021D02

Nuclear emulsion for double strangeness hypernuclei

- Current outcome of E07:
  - Triggered events : Ξ^- identified and tracked by detectors + outgoing K+ → estimation of the position of stopped Ξ^- in emulsion
  - Visual inspections by an optical microscope → around the estimated stop position
  - Small portion of emulsion plates analyzed → too much human workload !

Nuclear emulsion for double strangeness hypernuclei

- Still in those 1300 emulsion plates :
  - K- beam interacted directly with the nuclei of the emulsions
    → produce hypernuclei (single & double)
  - It was proposed to search for hypertriton ($^{3}_{\Lambda}H$)
  - But : no additional information → need to scan everything !
    → 1.4 billion images / emulsion : 110 TB x 1300 → 140 PB
    → 560 years to analyze this
  - Background :
    - Beam tracks & Nuclear fragmentation : 10000 & 1000 / mm^2
- Use of machine learning to find those events !
  → To be done in 3 years

Finding alpha decay for nuclear emulsion calibration

- Study of emulsion:
  - Measure ranges of the particles & fragments

    → With range – kinetic energy relation: measure of decay kinematics
- Alpha decay events:
  - Spontaneous decay chain of long-lived radioisotopes such as uranium and thorium in the emulsion
  - Calibration for density / space homogeneous

    → Absolute calibration of the range – kinetic energy relation
- Convolutional Neural Network: ResNet-50

Finding alpha decay for nuclear emulsion calibration

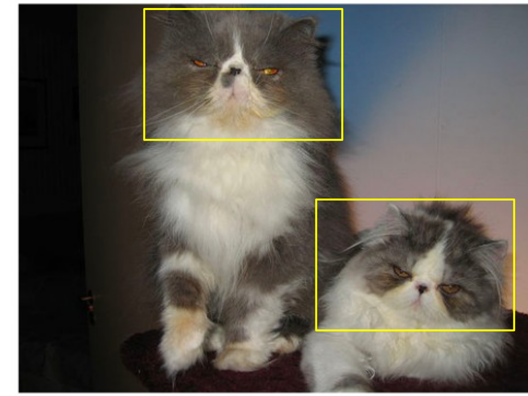
- What is a CNN :
  - Data with some “invariance to translation”

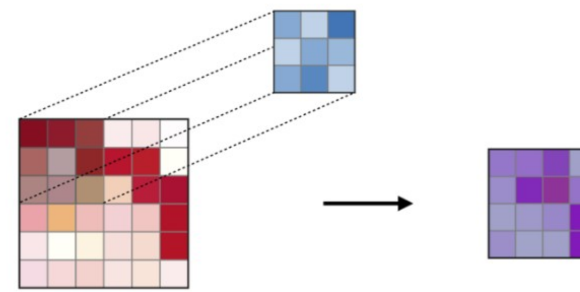
    → A representation at a certain location can be used elsewhere
  - Convolutional layers, build on this idea:
    - A same “local” transformation applied everywhere

      → preserve structure of signals
  - 1D Discrete Convolution:

$$\mathbf{x}\in\mathbb{R}^M, u\in\mathbb{R}^n, \forall i\in[0\ldots M-n+1]:(x*u)_i=\sum_{j=0}^{n-1} x_{i+j}u_j$$

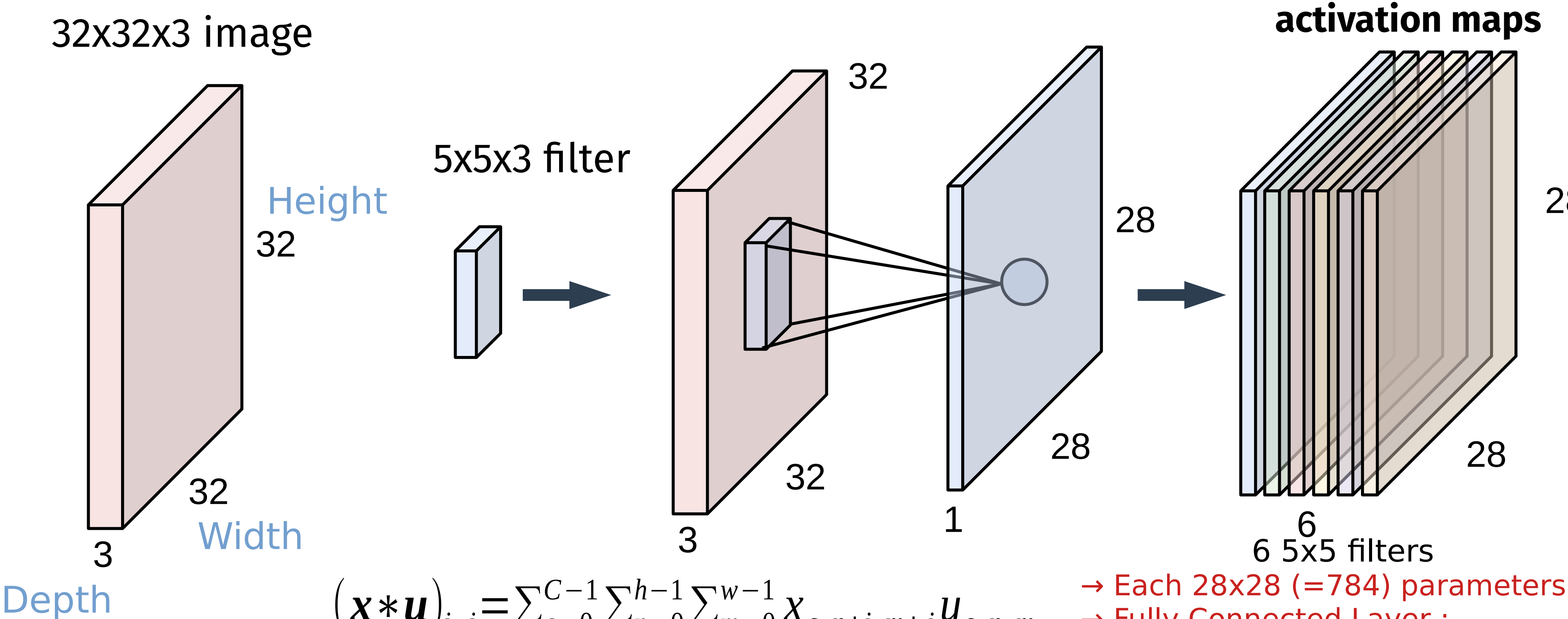
    - u is called Convolutional kernel of width k
  - Scan across data and multiply by kernel elements





Finding alpha decay for nuclear emulsion calibration

- Convolution Layer: preserve spatial structure



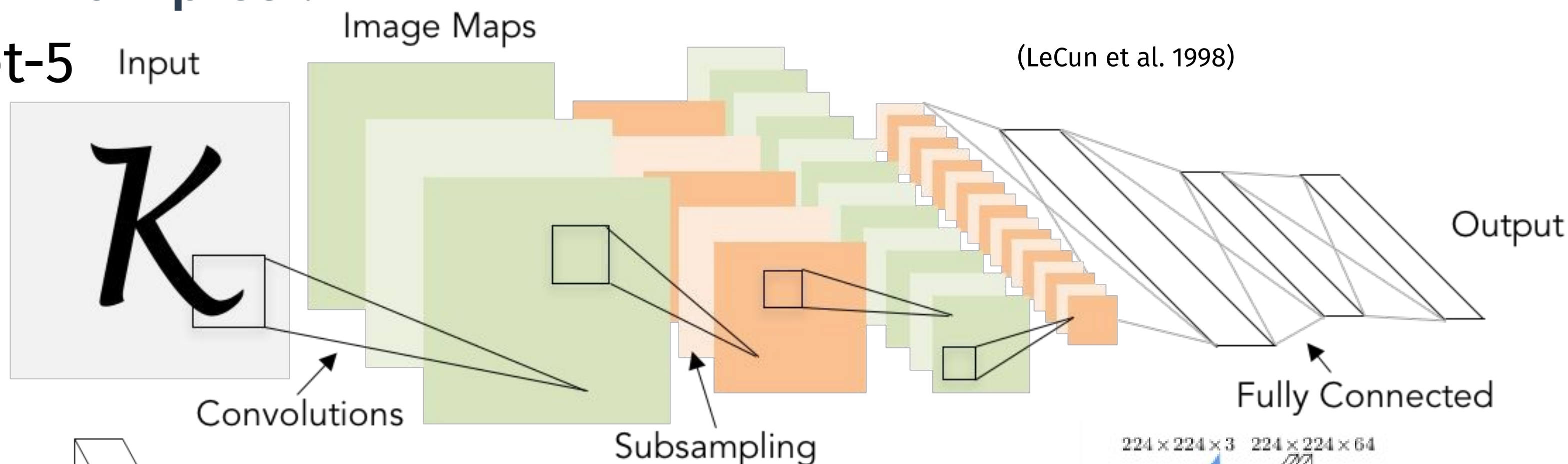
$$\left(\boldsymbol{x}*\boldsymbol{u}\right)_{i,j}=\sum_{c=0}^{C-1}\sum_{n=0}^{h-1}\sum_{m=0}^{w-1}x_{c,n+i,m+j}u_{c,n,m}$$

$$i\times j\in\left(H-h+1\right)\times\left(W-w+1\right)$$

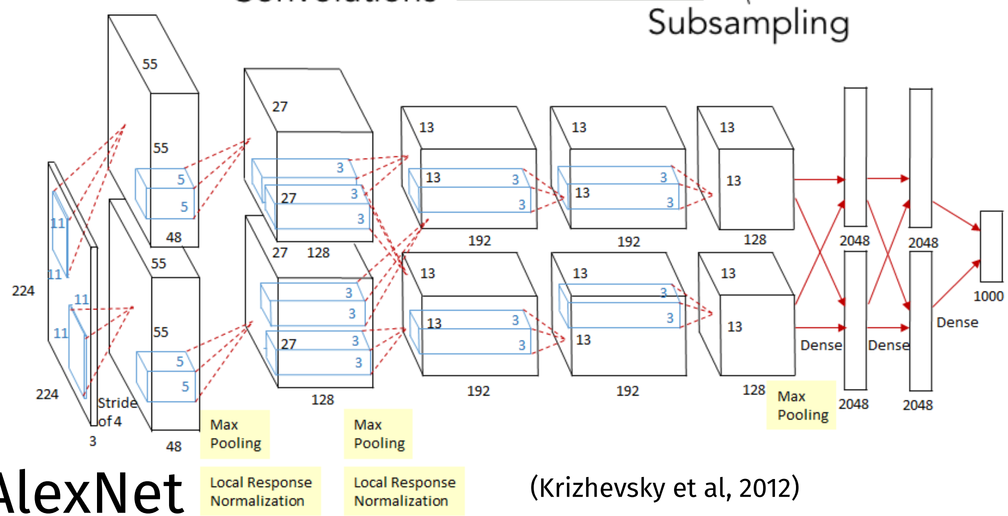
→ Each 28x28 (=784) parameters
→ Fully Connected Layer :
32x32x3 x size hidden (784)= 2.4M

- Examples :

LeNet-5



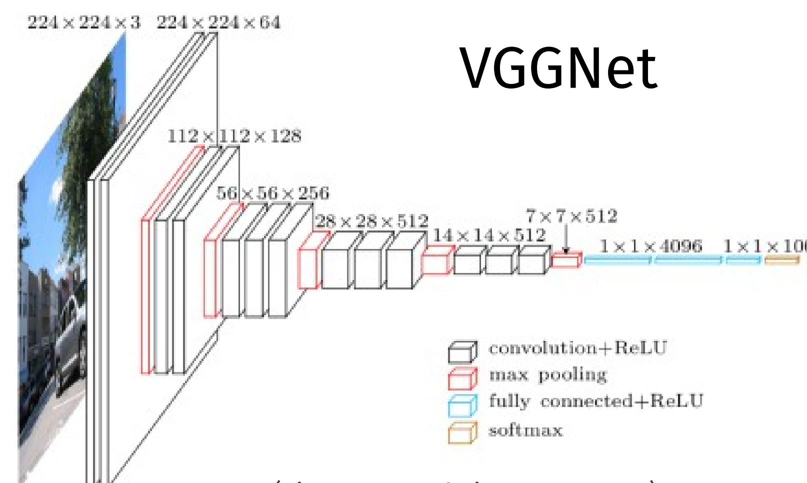
(LeCun et al. 1998)



AlexNet

(Krizhevsky et al, 2012)

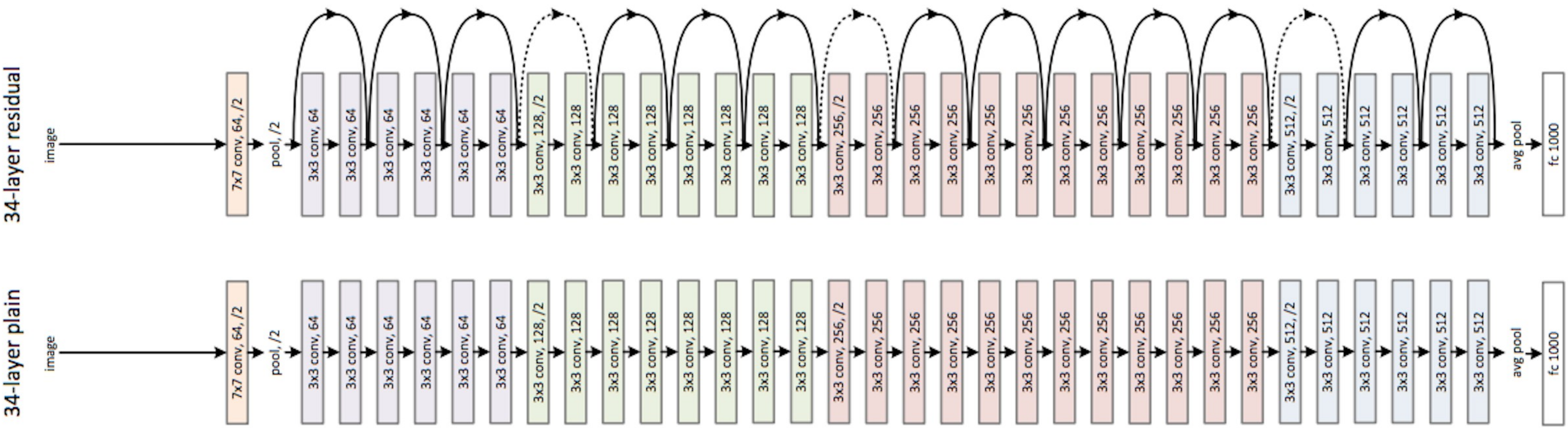
VGGNet



(Simonyan and Zisserman, 2014)

Finding alpha decay for nuclear emulsion calibration

- ResNet:
  - 34 layers:



  - Classics : ResNet – 18, 34, 50, 101, 152 (layers)

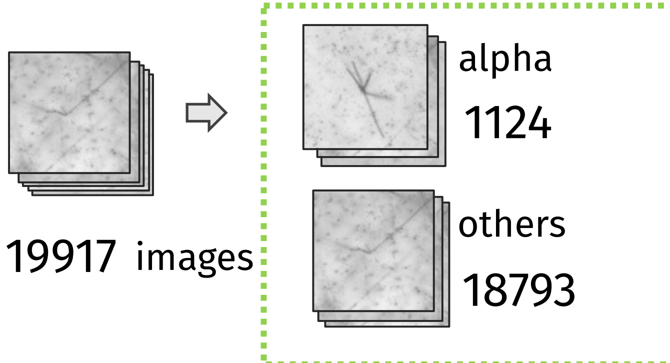
Params: 25M

Params: 60M

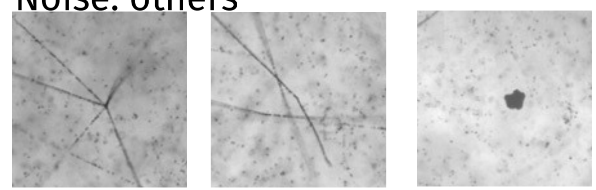
Finding alpha decay for nuclear emulsion calibration

- Alpha decay events (calibration) : CNN

Training data (real images)



Noise: others

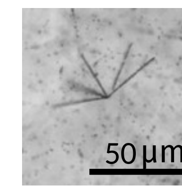


Other interaction

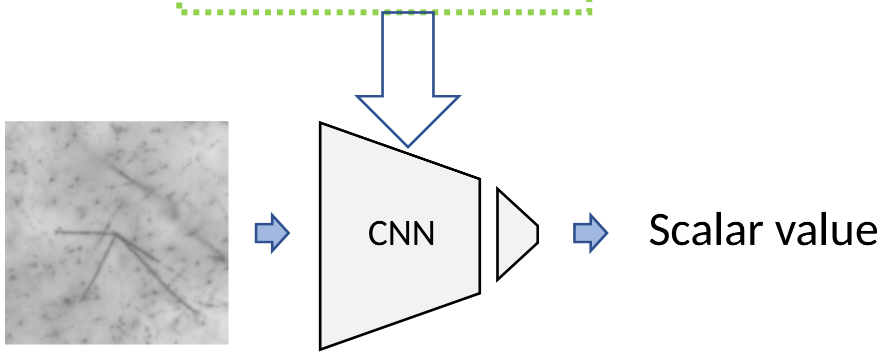
Cross

Dust

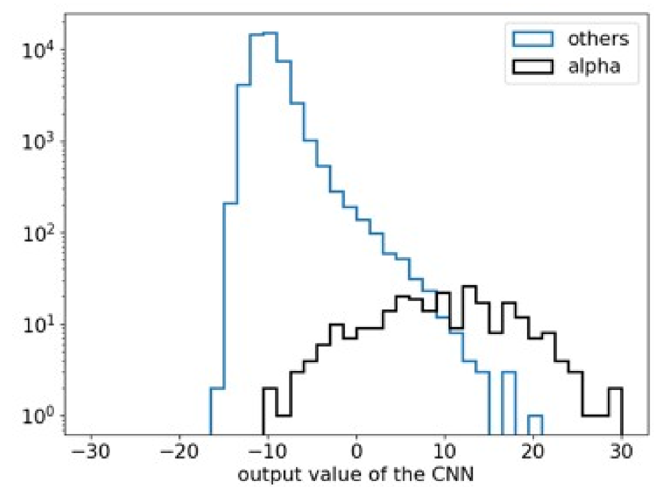
Target



α decay



300x300 pixels



J. Yoshida, et al., NIM A, 989 (2021) 164930

Finding alpha decay for nuclear emulsion calibration

- CNN classifier : Alpha decay detection

	Precision	Recall	# of candidates
Conventional method	0.081 +- 0.006	0.788 +- 0.056	2489
CNN classifier	0.547 +- 0.025	0.788	366 +- 18

- Precision = TP / TP + FP   Recall = TP / TP + FN
- 7 times more precision !

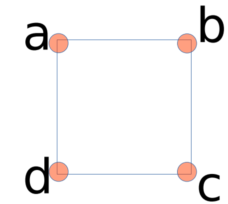
model's ability to detect Positive samples

- Conventional :
  - 2489 out of 46948 events, including 201 true alpha decay
- CNN classifier:
  - 350 alpha-decay candidates, including 201 true alpha-decay

GNN for
Particle tracking

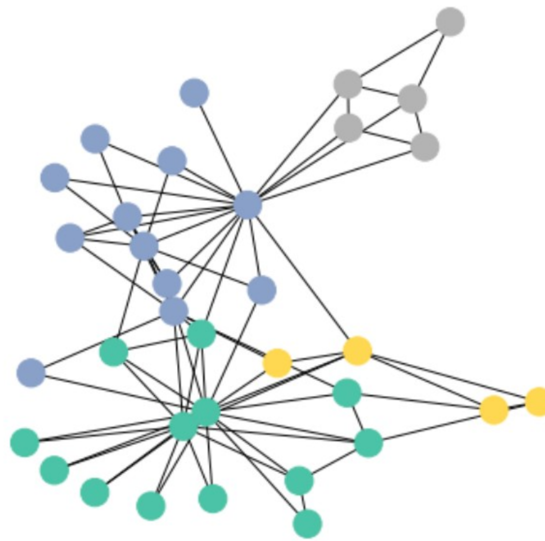
Particle tracking : Graph Neural Network

- Graph neural network:
  - Operate on graph structured data
  - Graph = nodes that can be connected by edges →



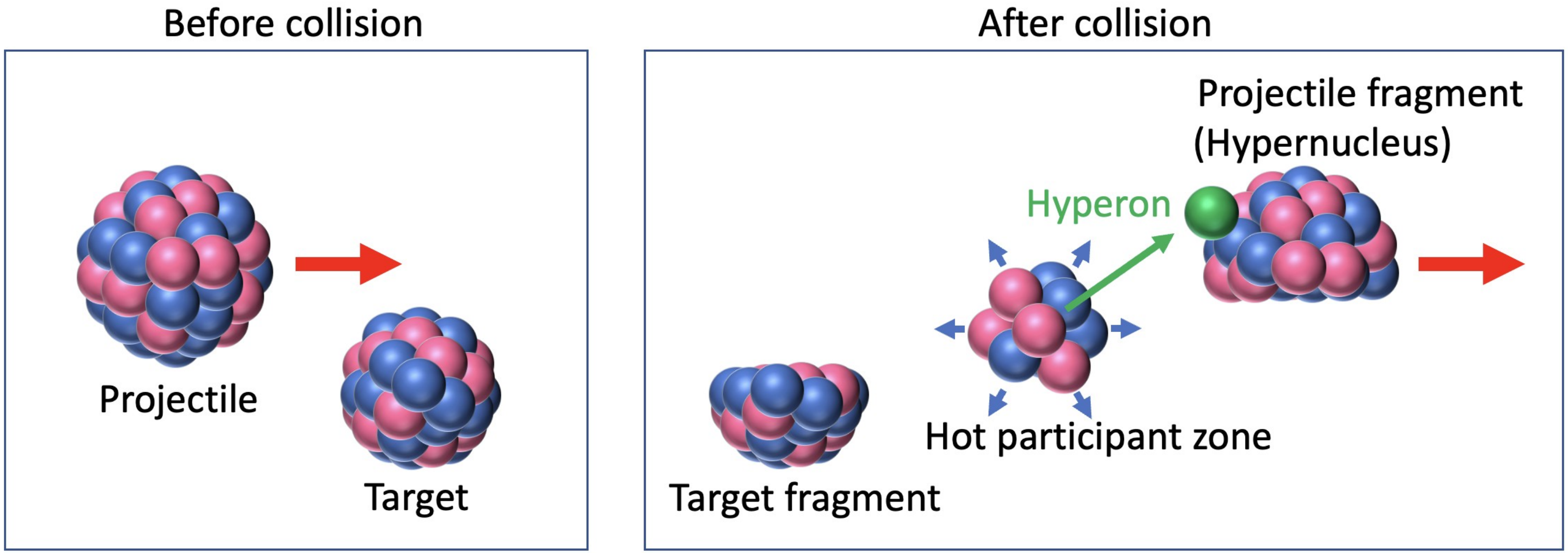
$$\begin{array}{c c} & \begin{array}{cccc} a & b & c & d \end{array} \\ \begin{array}{c} a \\ b \\ c \\ d \end{array} & \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix} \end{array}$$

- GNN = CNN + network embedding
  - Generalization of convolutional neural network
    - CNN → work on local spatial features
    - Networks & graph → generalize to arbitrary object
  - CNN : conv filter → locality
  - Graph : adjacency matrix → object relationships
  - No grid structure: adjacency matrix
    → Operations pass information from neighborhood



Particle tracking : Graph Neural Network

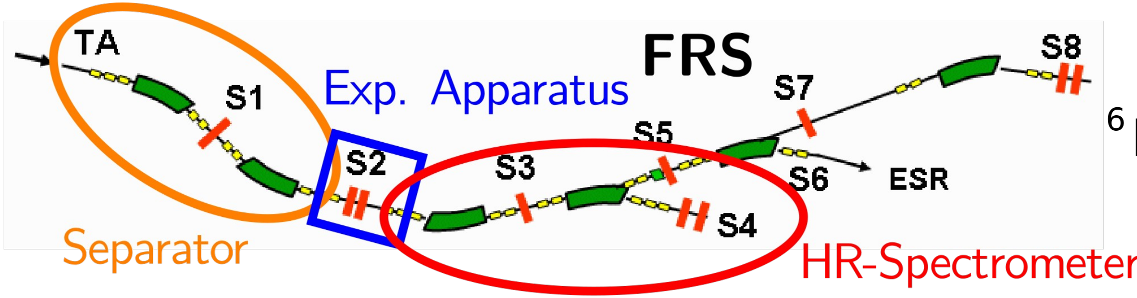
- Hypernuclear production in heavy ion collisions:
  - $NN \rightarrow \Lambda KN$ $E_{th} \sim 1.6$ GeV : Beam > E_{th} : available at GSI (2 *A*GeV)



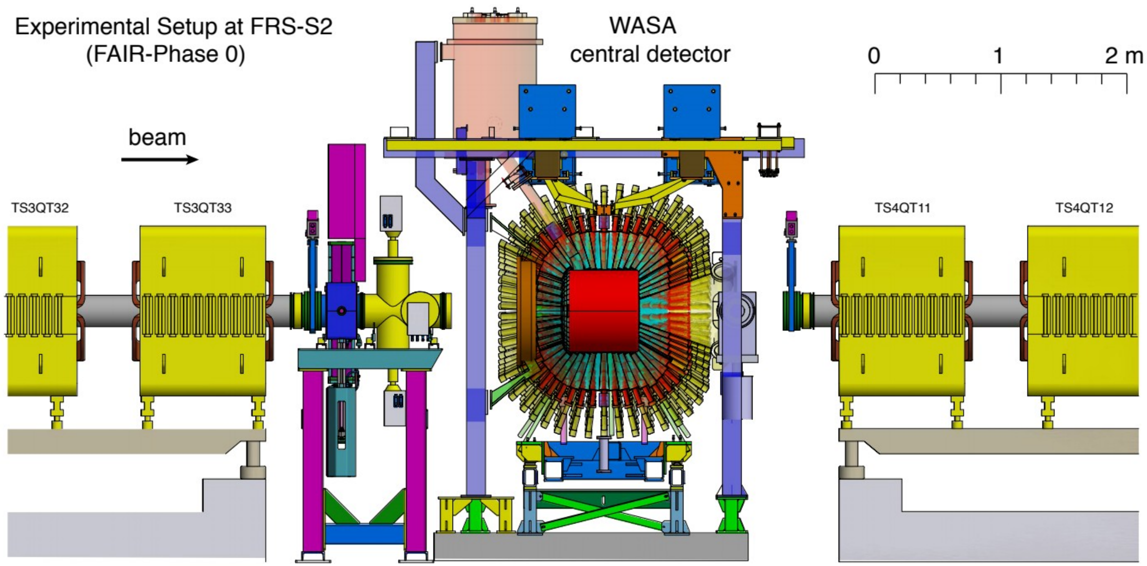
  - Coalescence of Λ in spectator fragment
    - same velocity than projectile: Lorentz Boosted
    - study Hypernuclei in flight

Particle tracking : Graph Neural Network

- Hypernuclear study in WASA-FRS experiment:



$$^{6}\mathrm{Li}(@2\,\mathrm{GeV/u})+{}^{12}\mathrm{C}\rightarrow{}^{3}_{\Lambda}\mathrm{H},{}^{4}_{\Lambda}\mathrm{H},\mathrm{nn}\Lambda+\mathrm{K}^{+}$$
$$\mathrm{Fragments}+\mathrm{Hadrons}$$



$${}^{3}_{\Lambda}\mathrm{H}\rightarrow{}^{3}\mathrm{He}+\pi^{-}$$
$${}^{4}_{\Lambda}\mathrm{H}\rightarrow{}^{4}\mathrm{He}+\pi^{-}$$
$$\mathrm{nn}\Lambda\rightarrow\mathrm{n}+\mathrm{d}+\pi^{-}$$

Particle tracking : Graph Neural Network

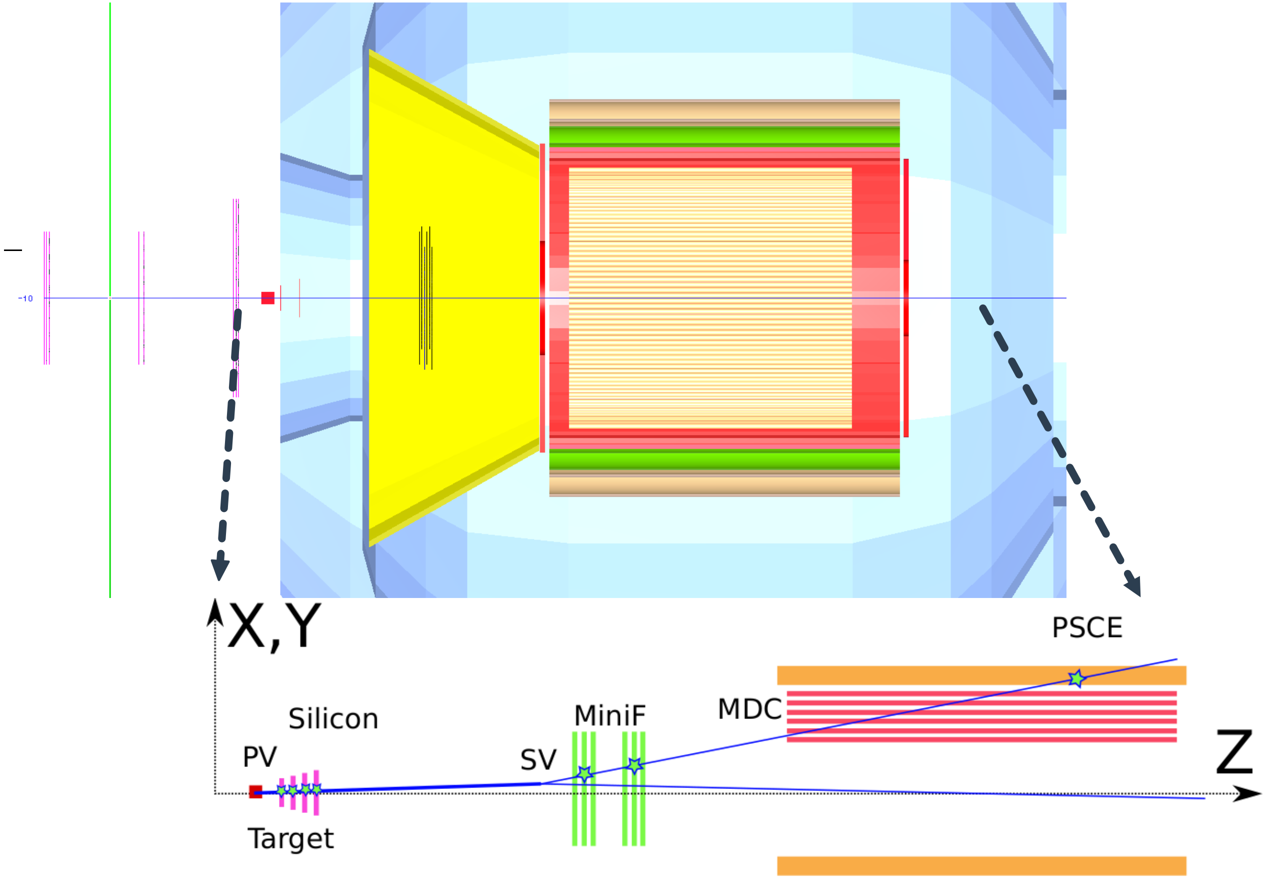
- Hypernuclear study in WASA-FRS experiment:

$$^{6}\mathrm{Li}(@2\,\mathrm{GeV/u}) + {}^{12}\mathrm{C} \rightarrow {}^{3}_{\Lambda}\mathrm{H}, {}^{4}_{\Lambda}\mathrm{H}, nn\Lambda + K^{+} + \mathrm{Fragments} + \mathrm{Hadrons}$$

$${}^{3}_{\Lambda}\mathrm{H} \rightarrow {}^{3}\mathrm{He} + \pi^{-}$$

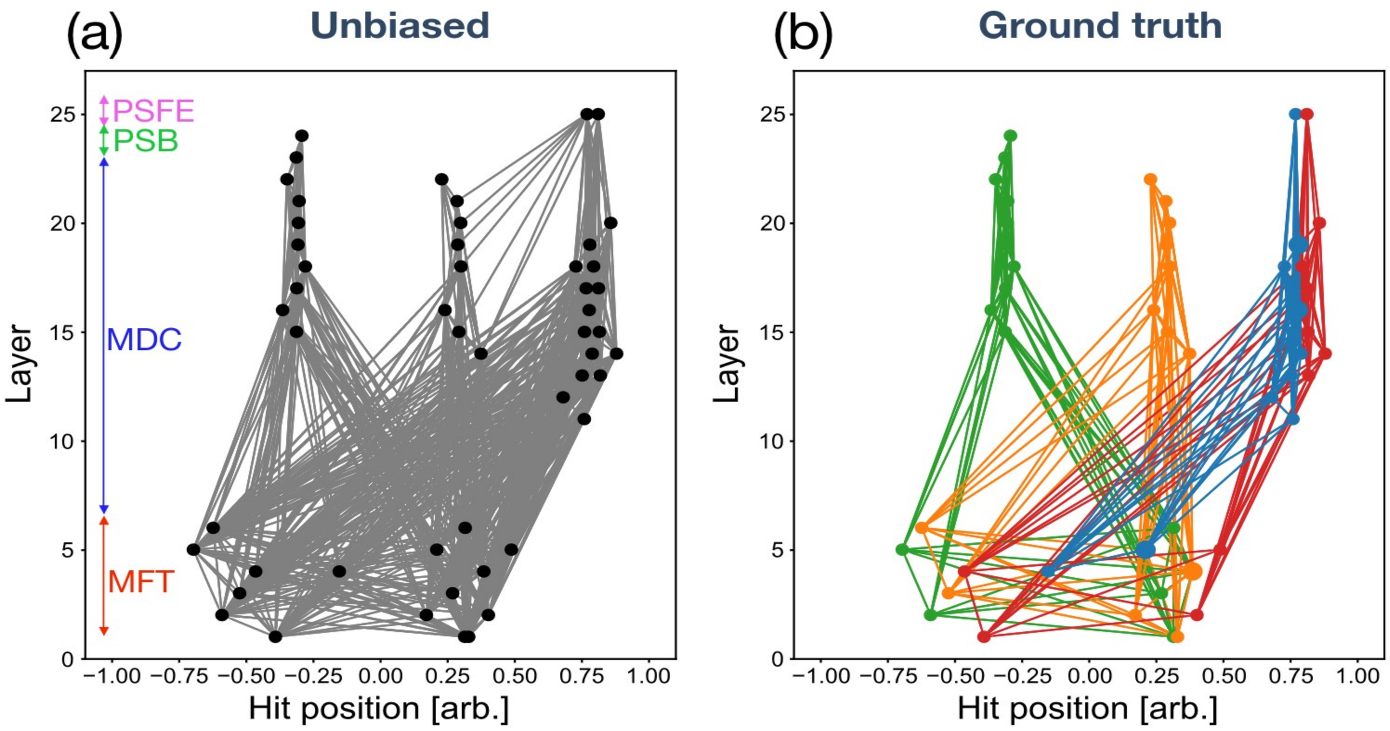
$${}^{4}_{\Lambda}\mathrm{H} \rightarrow {}^{4}\mathrm{He} + \pi^{-}$$

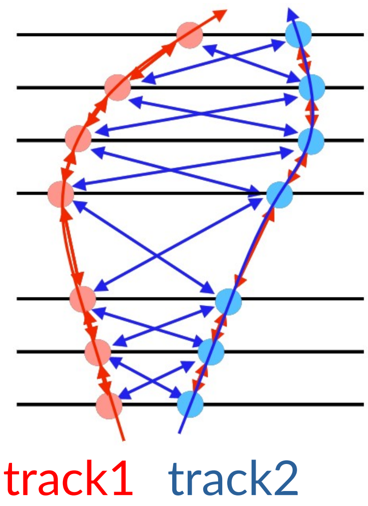
$$nn\Lambda \rightarrow n + d + \pi^{-}$$



Particle tracking : Graph Neural Network

- Study of Hypernuclei in our WASA-FRS experiment:



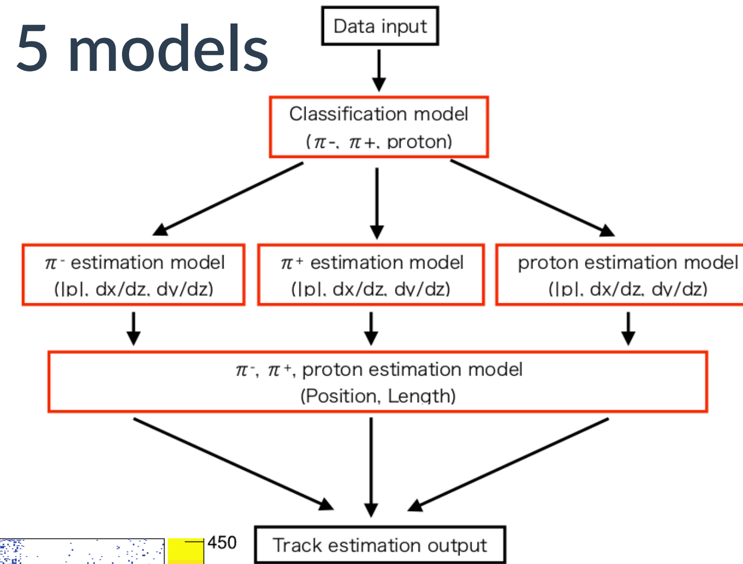


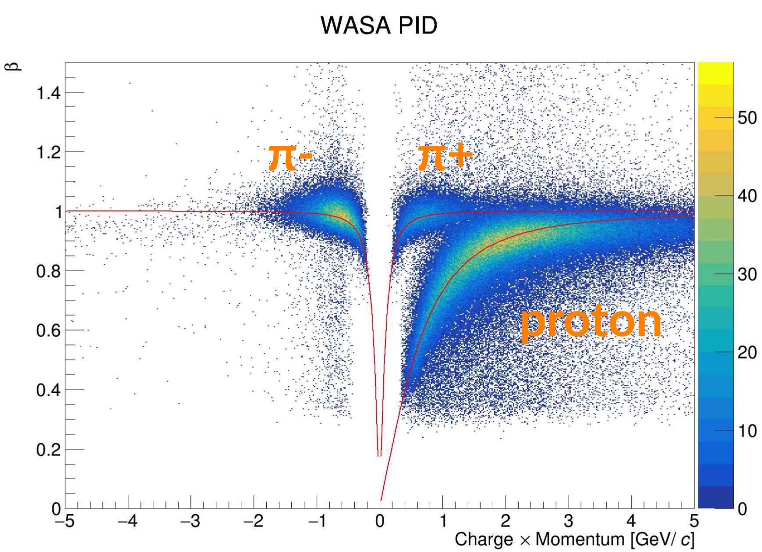
π- (perfect)	π- (valid)	Other (perfect)	Other (valid)
98.09 %	99.92 %	97.05 %	99.07 %

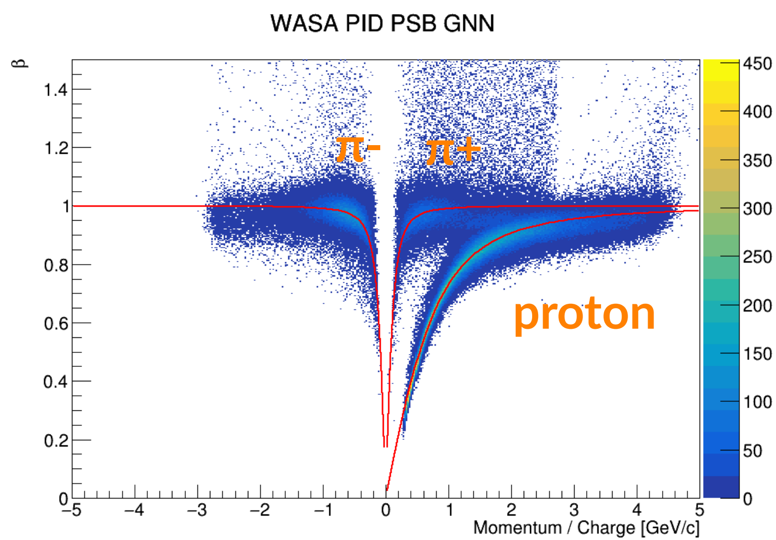
H. Ekawa et al., Eur. Phys. J. A 59 103 (2023)

Particle tracking : Graph Neural Network

- New R&D: More complex GNN models → 5 models
  - Excellent track finder → all particles
  - Good track parameters estimators → all
  - Allow Particle Identification with GNN



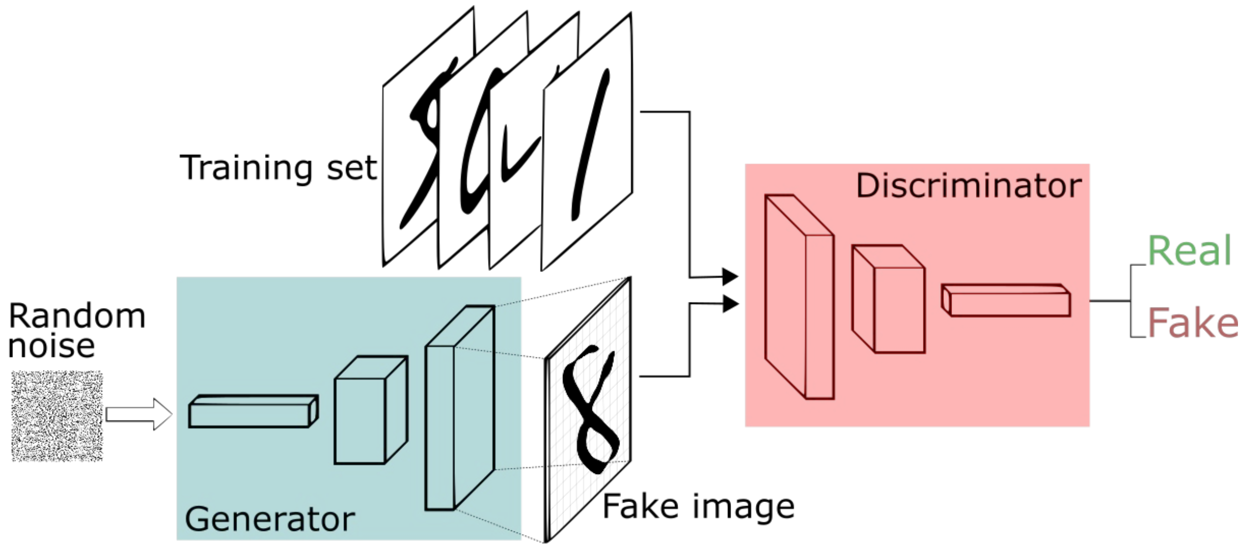




Generative AI & Object detection & Segmentation models for emulsion analysis

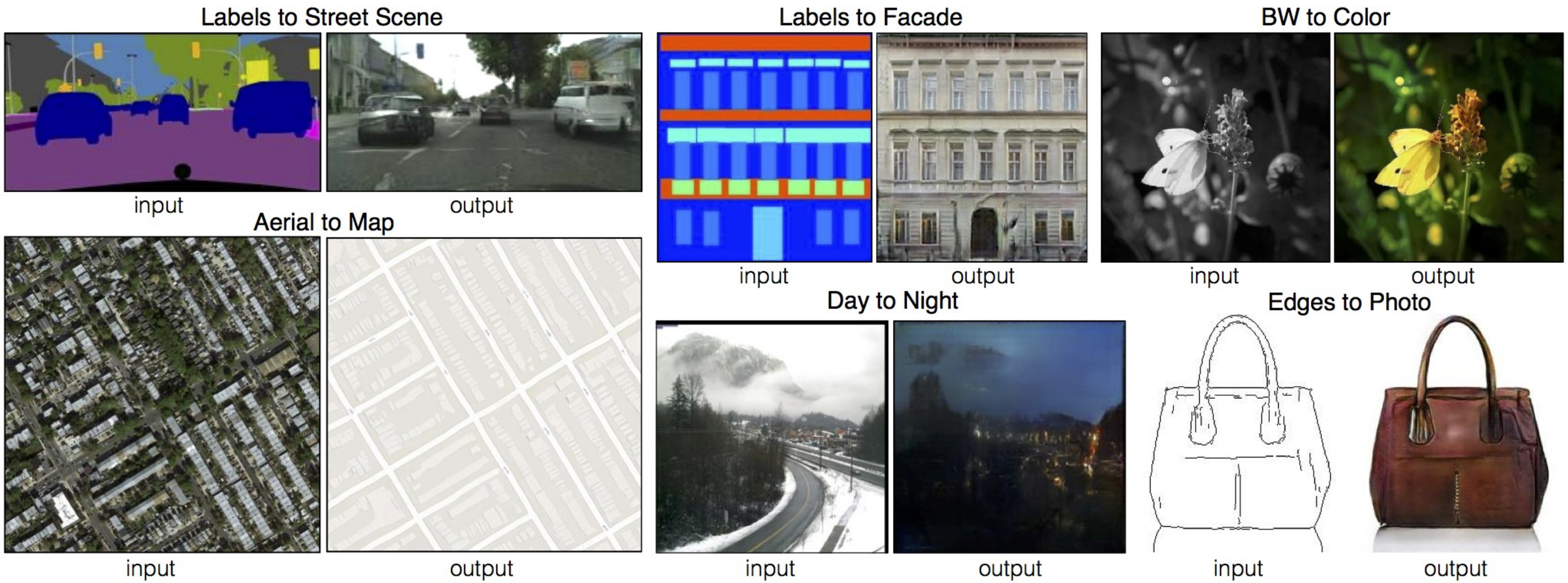
Finding hypernuclei in emulsion : MaskCNN

- Finding hypertriton :
  - Needs of training data ! But none has been found
    → generating event from simulations !
  - Problem : how to simulate nuclear emulsion ?!
- GAN : Generative adversarial networks



Finding hypernuclei in emulsion : MaskCNN

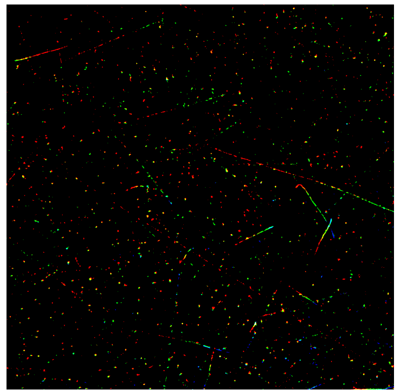
- Simulated hypertriton : GAN + Geant4
  - pix2pix (Image-to-Image Translation with Conditional Adversarial Nets)



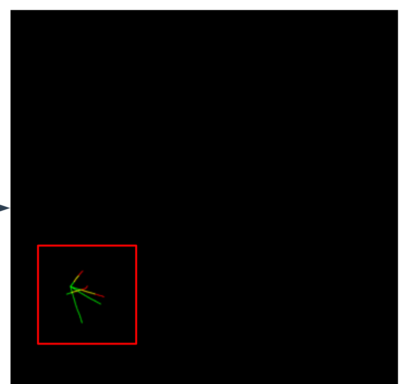
Finding hypernuclei in emulsion : MaskCNN

- Simulated emulsion :

Background (real)
→ from emulsion plate



Events (Geant4)
→ simulation of weak decay



Mix

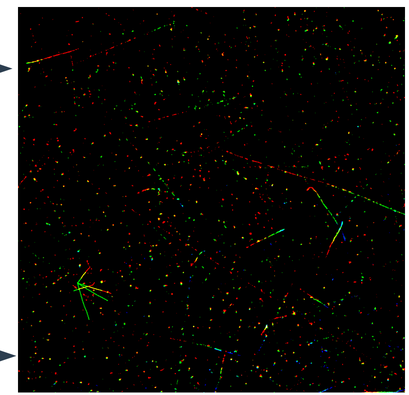
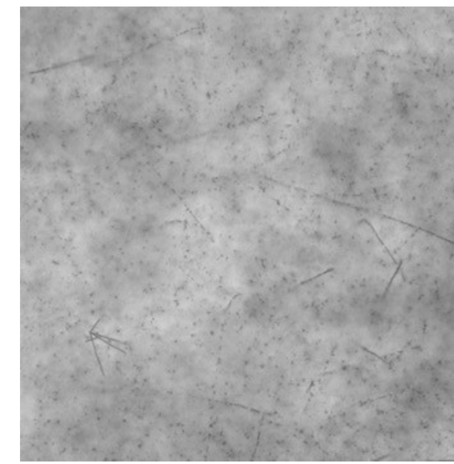


Image Converter pix2pix

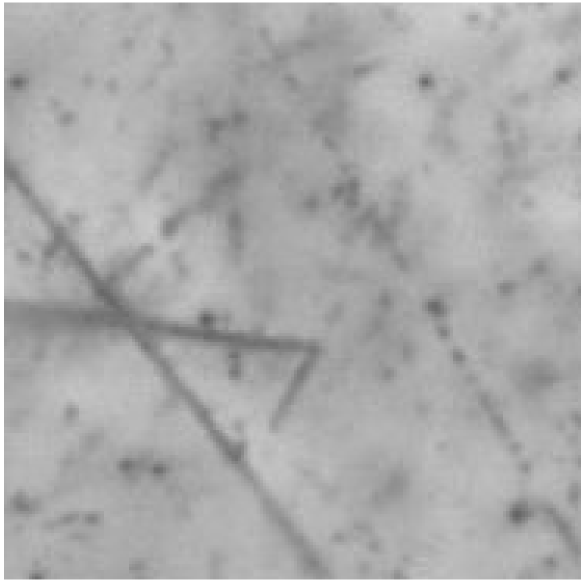
Synthesized images



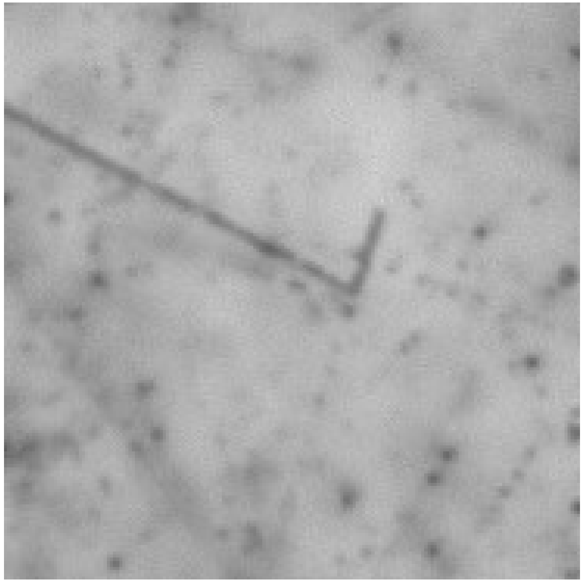
Finding hypernuclei in emulsion : MaskCNN

- Simulated event : hypertriton via GAN

A. Kasagi et al., NIMA, 1056 (2023) 168663



Simulated



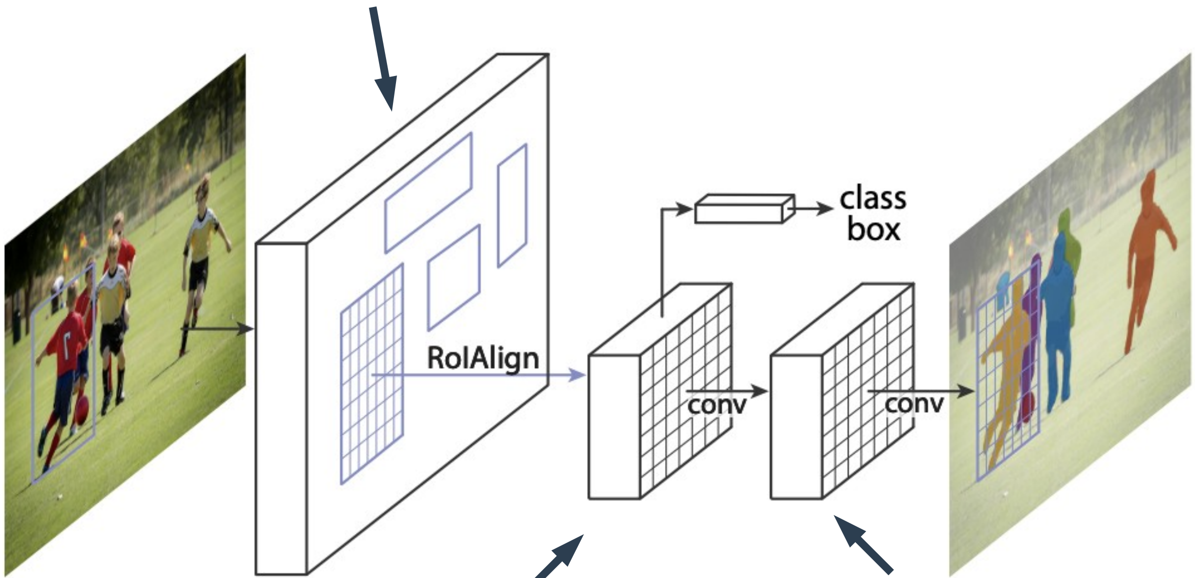
Real

- hypertriton decay at rest : $^{3}He + \pi^{-}$ back-to-back
- Q-value fixed: length of pion 28 mm of $^{3}_{\Lambda}H$ vs 42 mm for $^{4}_{\Lambda}H$

Finding hypernuclei in emulsion : MaskCNN

- Search for hypertriton-like decay:
  - Mask R-CNN : Instance Segmentation

Region Proposal Network



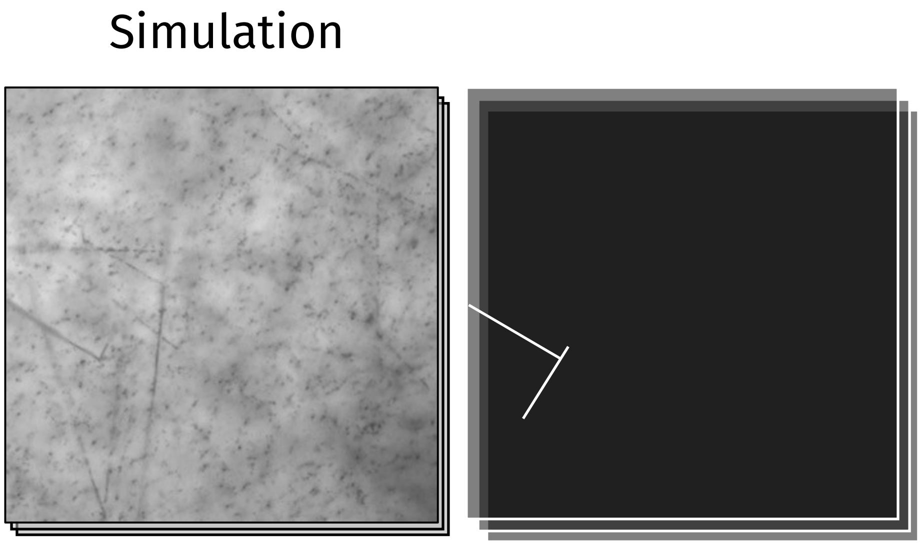
feature extraction Network

Network *head* for bounding-box recognition

Backbone architecture:
Networks inside
Ex: ResNet, ResNeXt,
Feature Pyramid Network

Finding hypernuclei in emulsion : MaskCNN

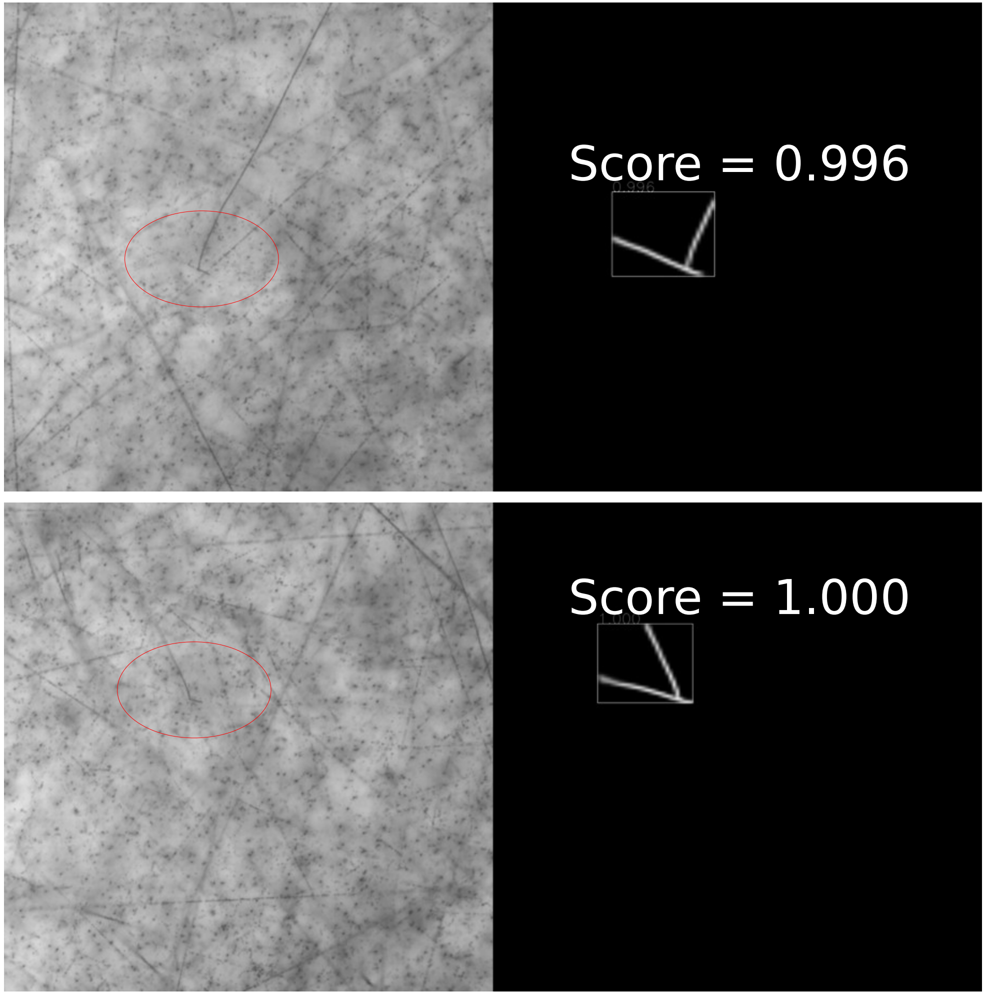
- Search for hypertriton-like decay:
  - Training on simulated and generated event
    - “Real” images of simulated emulsion
    - Masks of the instance segmentation of the decay



Finding hypernuclei in emulsion : MaskCNN

- Search for hypertriton-like decay:
  - Training on simulated and generated event → done
  - Analyze the real emulsion images

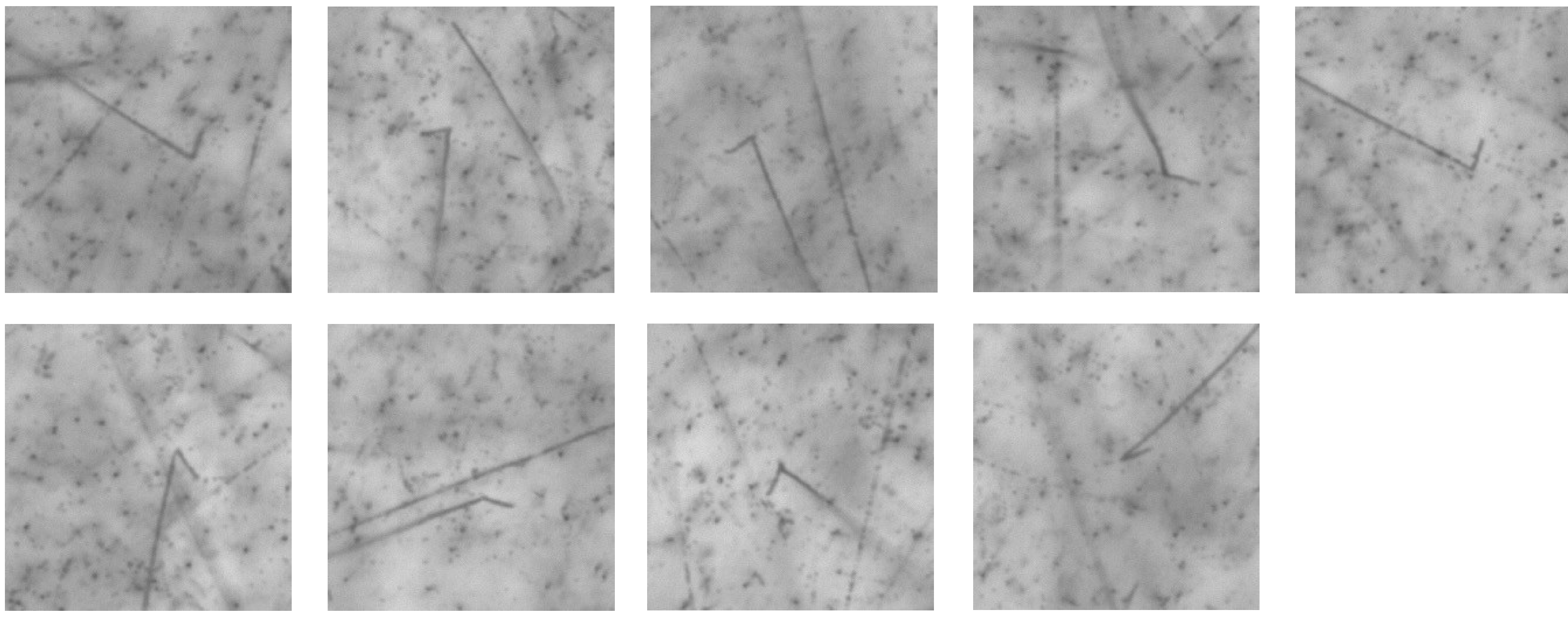
    → Give us the image and and mask – bounding box of what the algorithm found :



Finding hypernuclei in emulsion : MaskCNN

- Search for hypertriton-like decay:

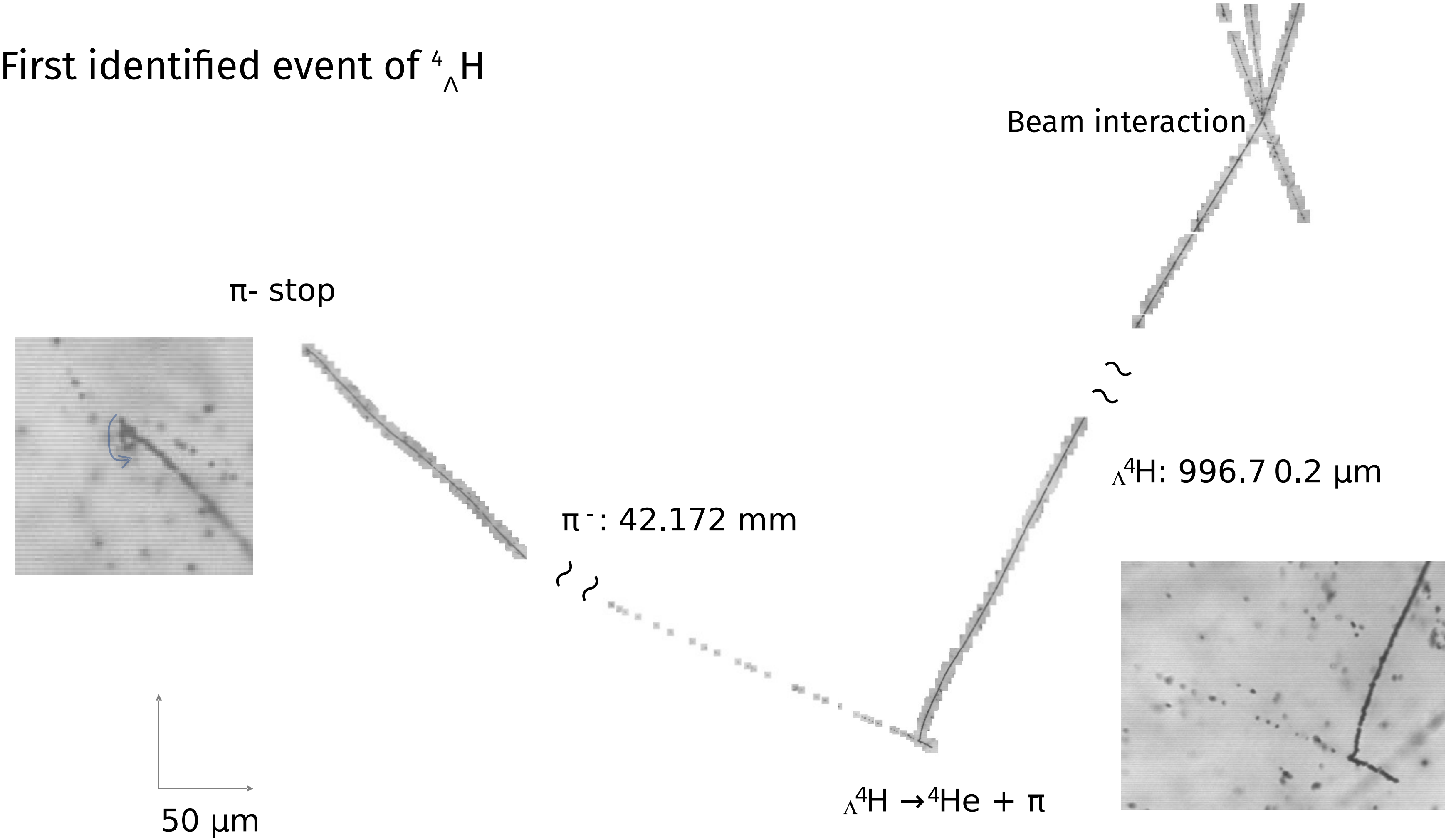
A. Kasagi et al., NIMA, 1056 (2023) 168663



10 µm

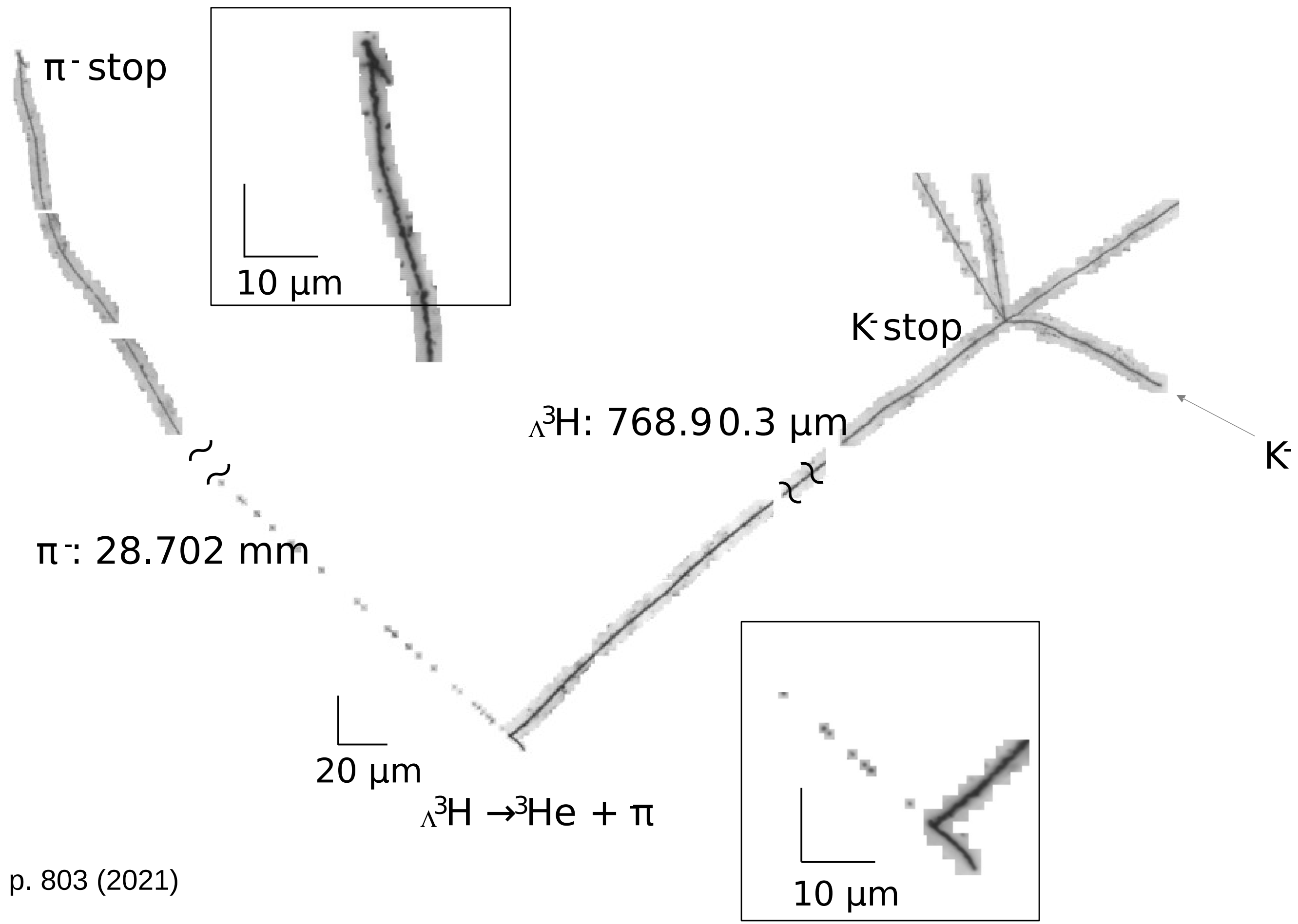
Finding hypernuclei in emulsion : MaskCNN

First identified event of ${}^{4}_{\Lambda}H$



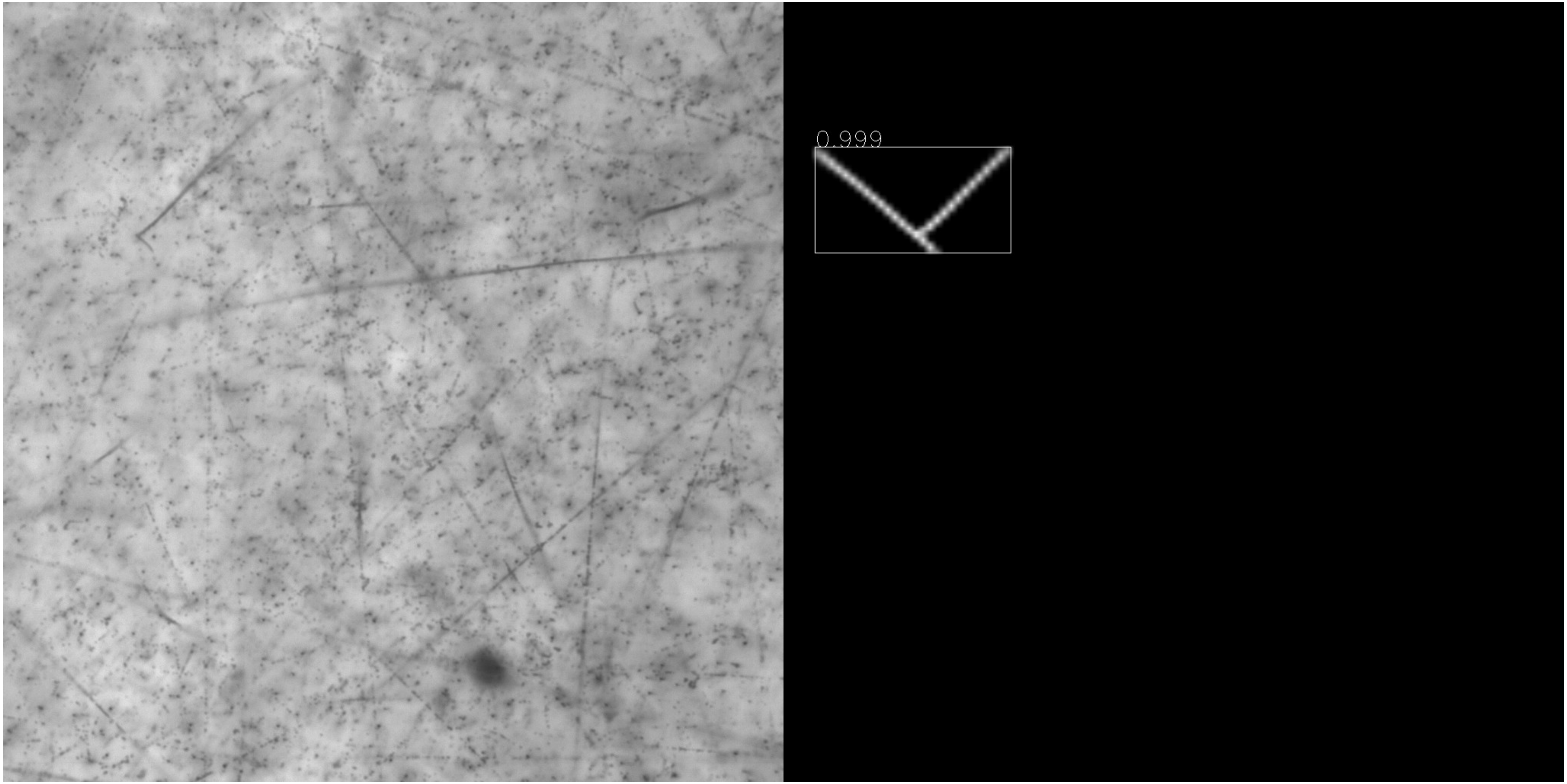
Finding hypernuclei in emulsion : MaskCNN

First $^{3}_{\Lambda}H$ identified



T. Saito et al., Nature Review Physics, 3, p. 803 (2021)

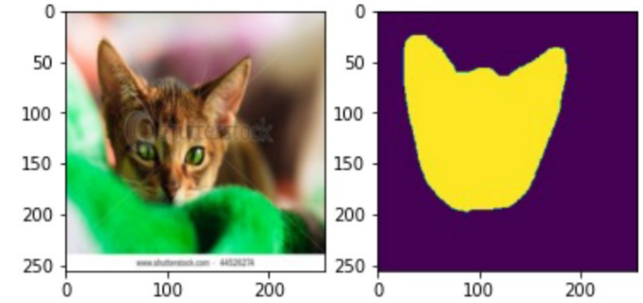
Finding hypernuclei in emulsion : MaskCNN

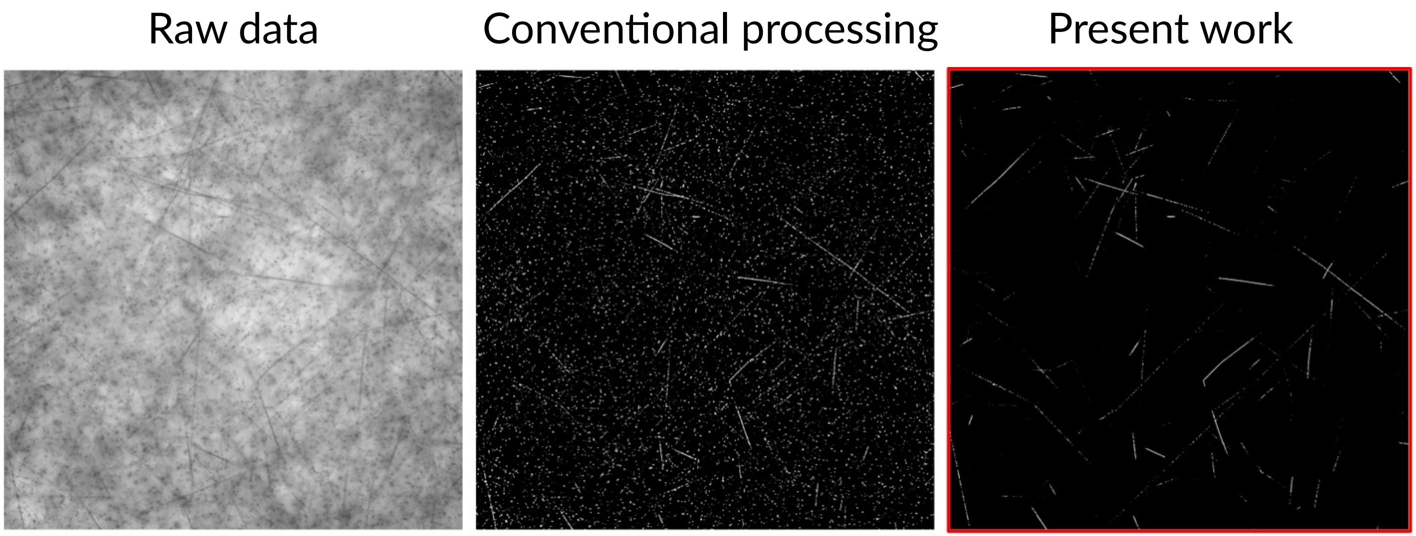


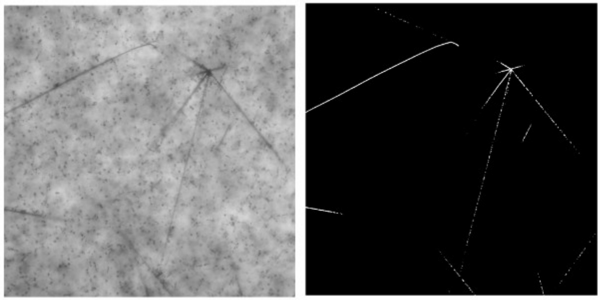
Detected by Mask R-CNN

Segmentation task to detect hit information

- Binary segmentation model:
  - Training from scratch: background or track

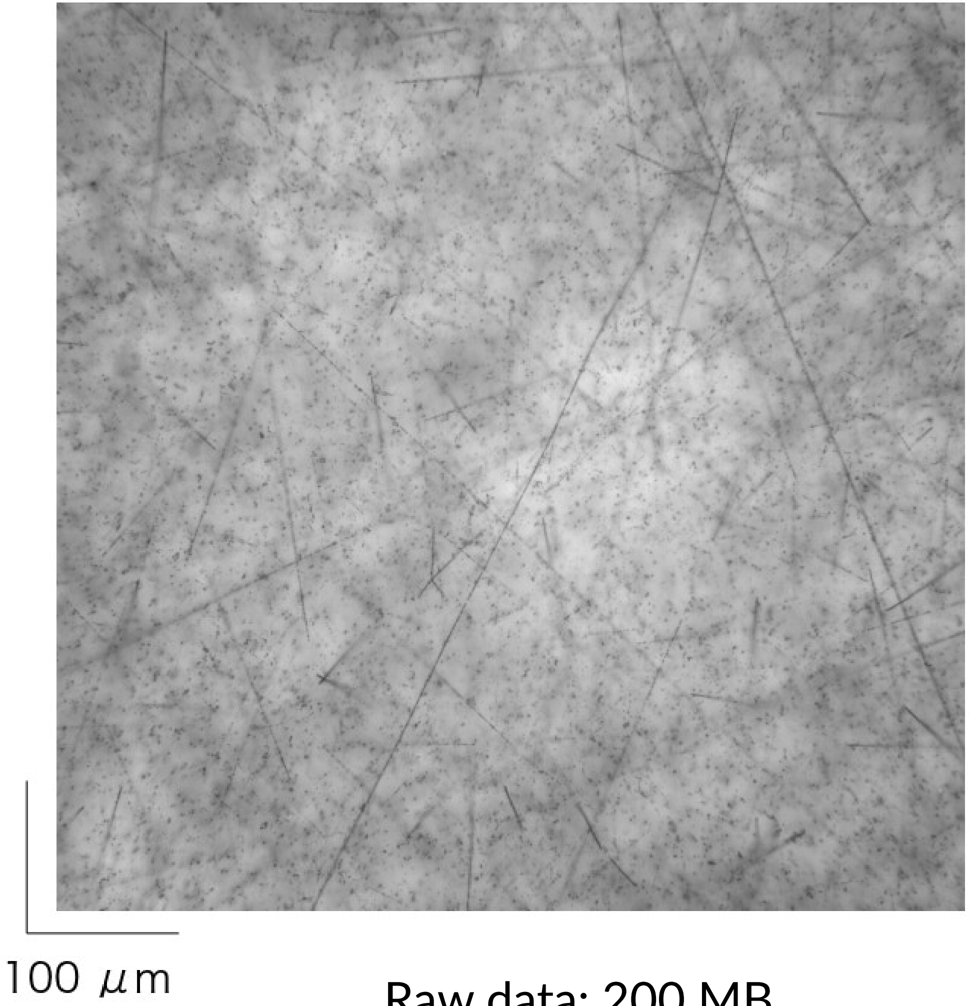




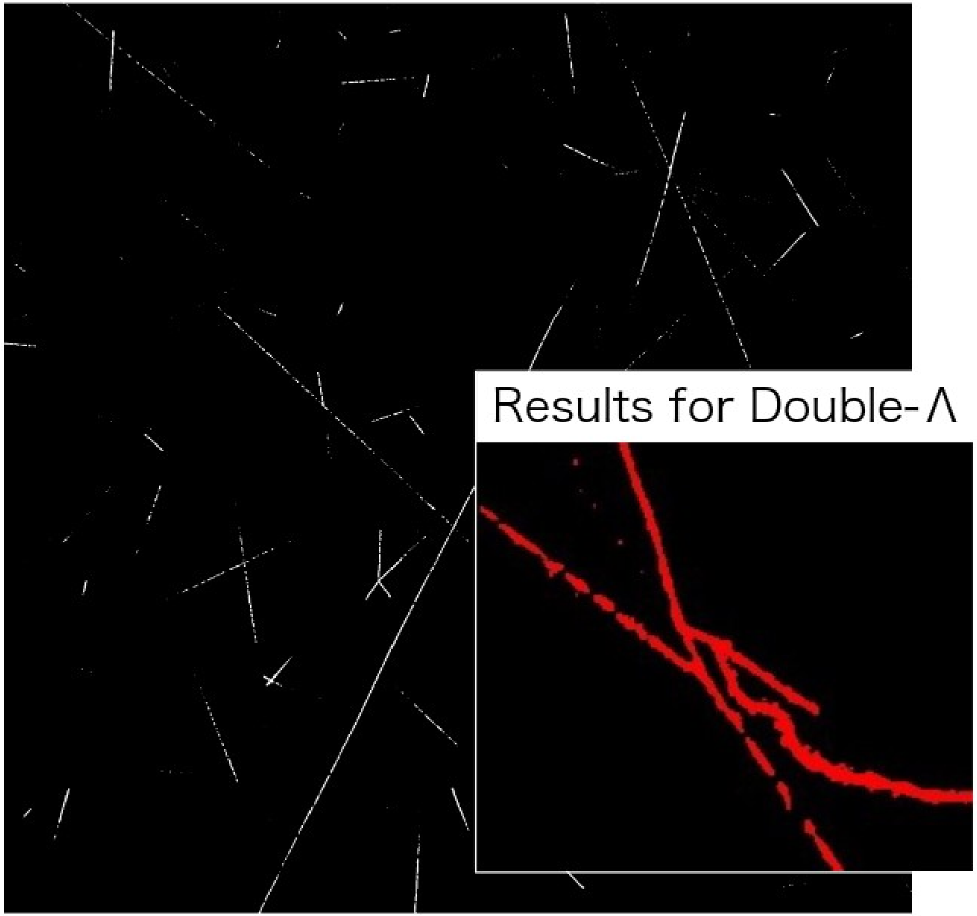


- Quantitative evaluation
- Noise reduction
- Datasize: 1/200
  → E07 image data
  150 PB → 750 TB

Segmentation task to detect hit information



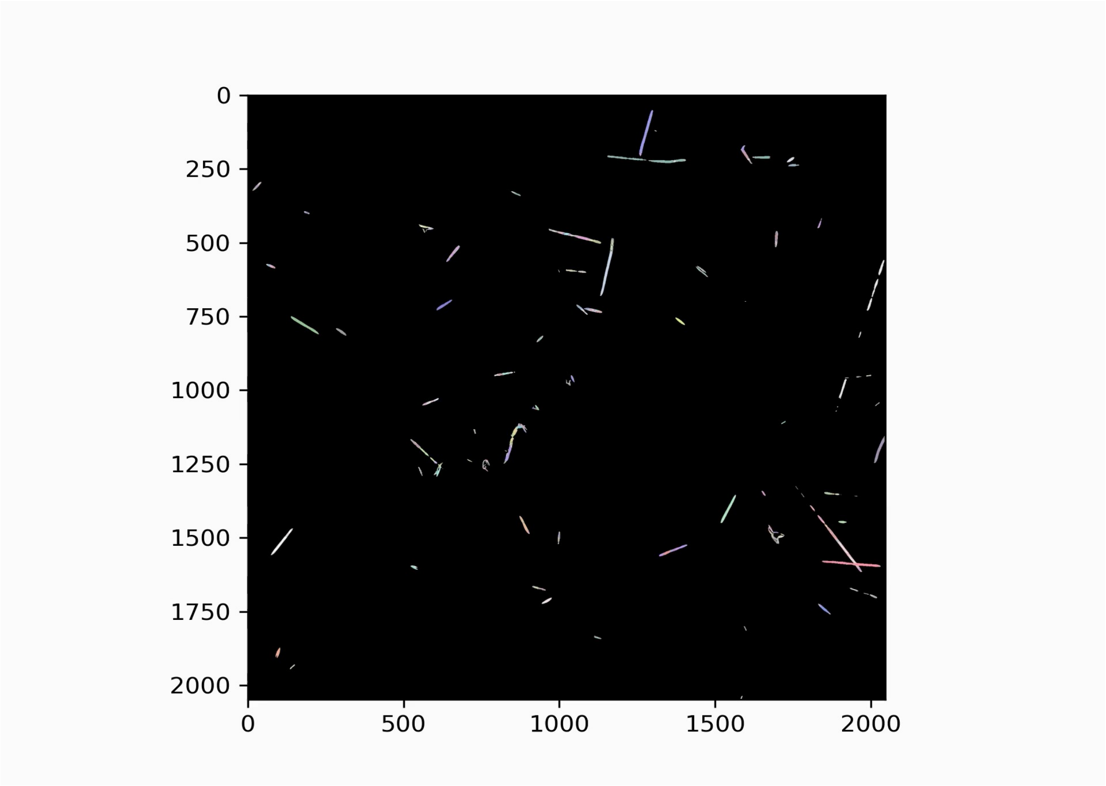
Raw data: 200 MB



Segmentation: 1MB

Segmentation task to detect hit information

- 3D track reconstruction



Any questions ?