Multivariate techniques for track reconstruction and particle identification

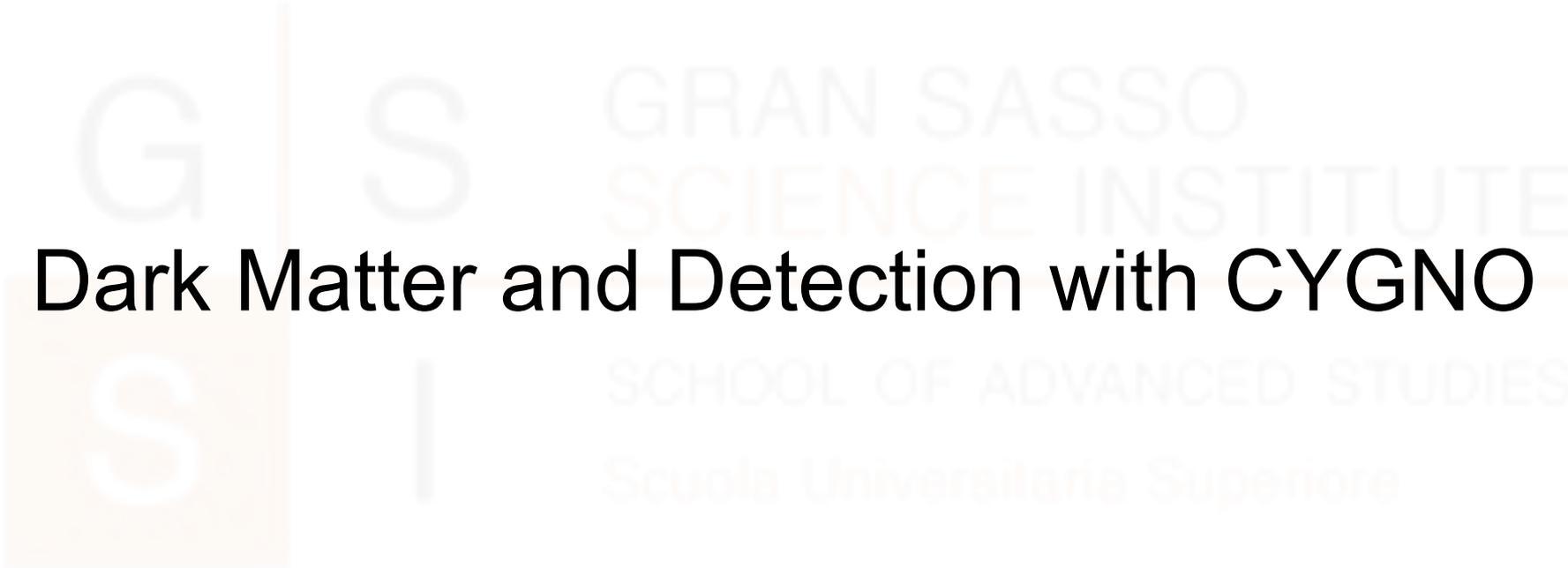**INITIUM**: **I**nnovative **N**egative **I**on **TI**me projection chamber for **U**nderground dark **M**atter searches

Candidate: Atul Prajapati   **Supervisor**: Prof. E. Baracchini

Outline

❖ Dark Matter and It's detection

❖ Low Energy X-Rays data with LIME detector

❖ Background Rejection with Deep Learning Models

❖ Starting With Convolutional Neural Networks (CNN)

❖ Future Work

Dark Matter and Detection with CYGNO

CYGNO/INITIUM Detector
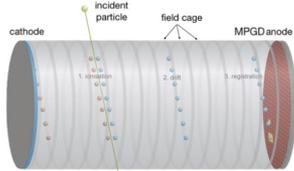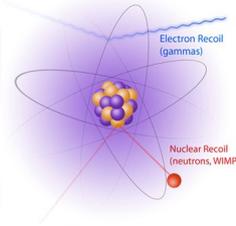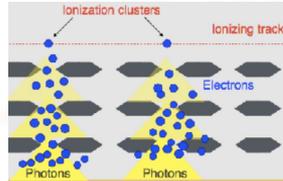
Dark Matter Signal

| Wind of DM particles	➡	Elastic scattering with ordinary matter	➡	Recoiling Nuclei

Important Sources of Background:

- ❖ Neutrinos from Sun and Atmosphere
- ❖ Cosmic Rays and Cosmogenic activation of detector materials
- ❖ Natural Radioactivity

Amplification Region + Readout (sCMOS + PMT)

- ❖ CYGNO uses $He:CF_4$ gas mixture
- ❖ 3 GEM stack is used for charge amplification
- ❖ INITIUM is a part of CYGNO project which focuses on the development of TPCs with negative ion drift using SF_6 gas

Camera

PMT
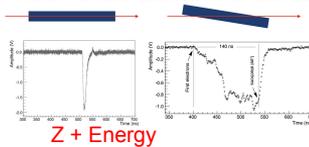
X-Y + Energy

Z + Energy

sCOMS:
- ● Single photon Sensitivity
- ● High Granularity
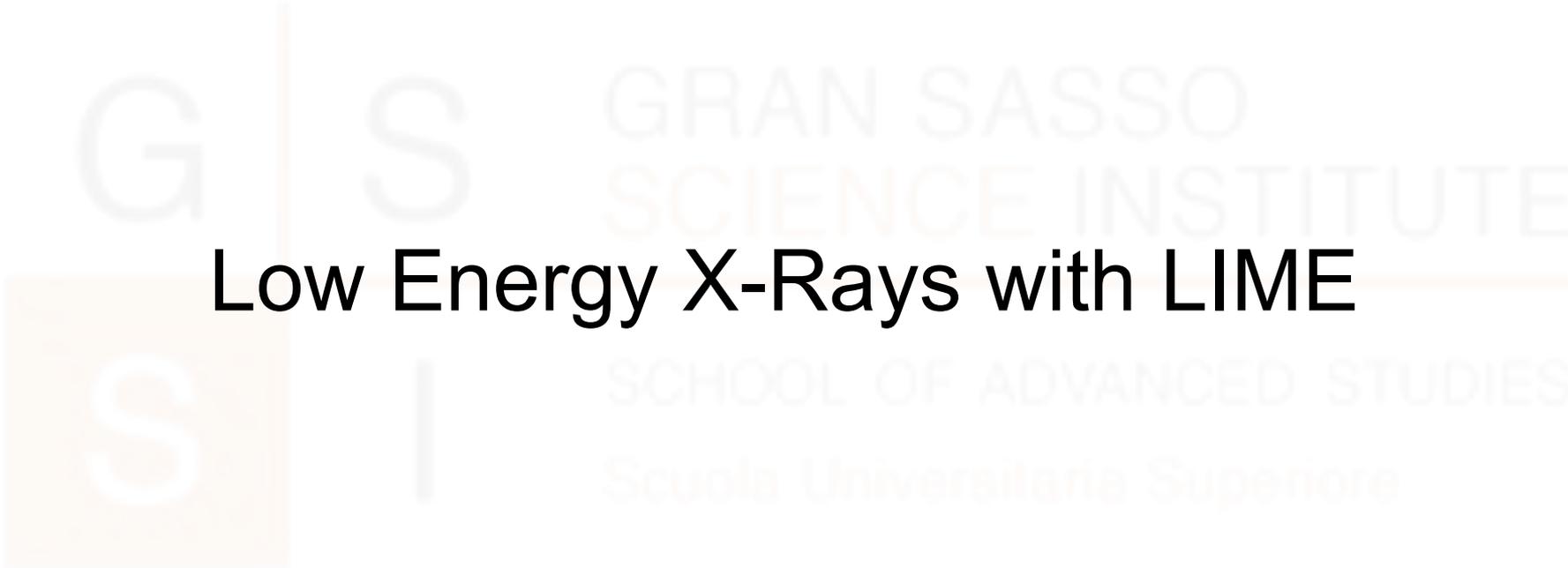- ● Large area for detection

PMT:
- ● Fast
- ● Integrated energy measurement
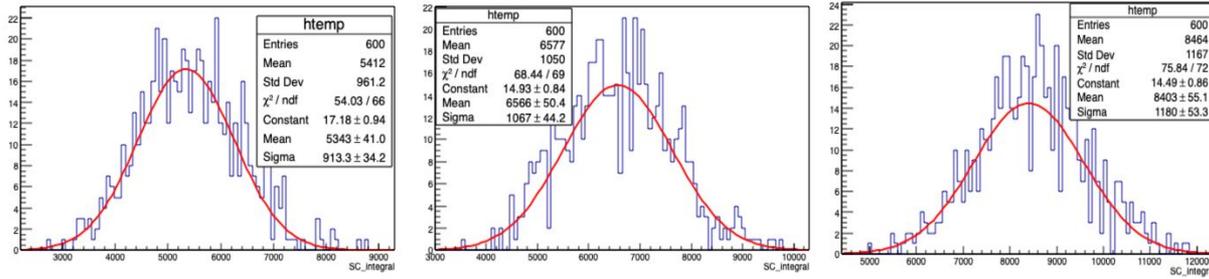
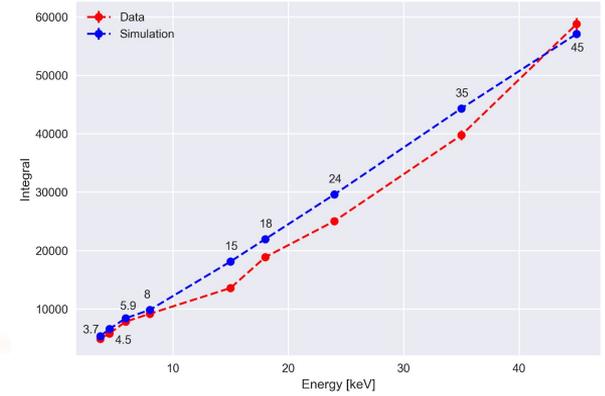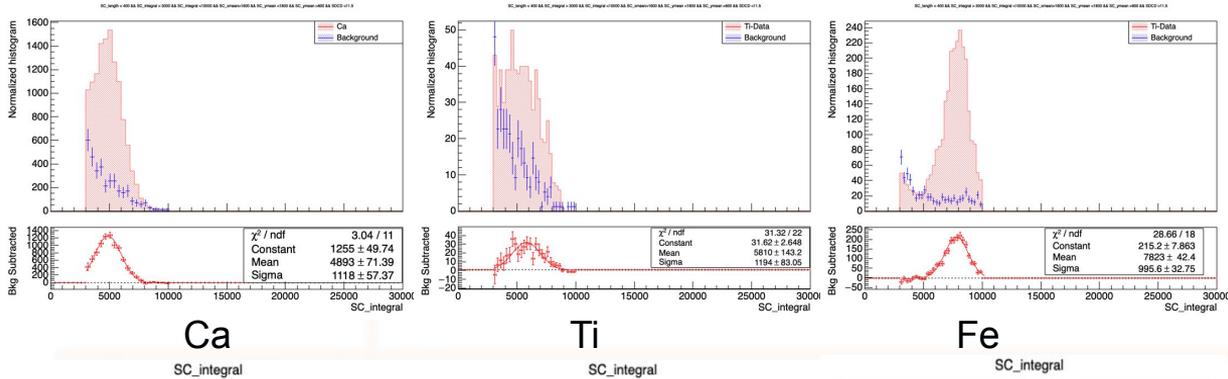**Background can be reduced by:**
- ❑ Background Control (Cleaning, Shielding,..)
- ❑ Underground Operation (Reduces muon induced neutron)
- ❑ Electron and Nuclear recoil discrimination

Interaction rate is extremely low and background rate is very high.

Low Energy X-Rays with LIME

Data/MC comparison of Ca, Ti and Fe data

Ca   Ti   Fe

- ❖ I managed to selects the events down to 3.5 keV energy.
- ❖ This work was very important to validate the data/MC agreement down to very low energy.

- ❖ I applied multiple cuts on variables to select the events.
- ❖ Integral of the data and simulated events matches well.
- ❖ Other variables were also used for data/MC comparison.
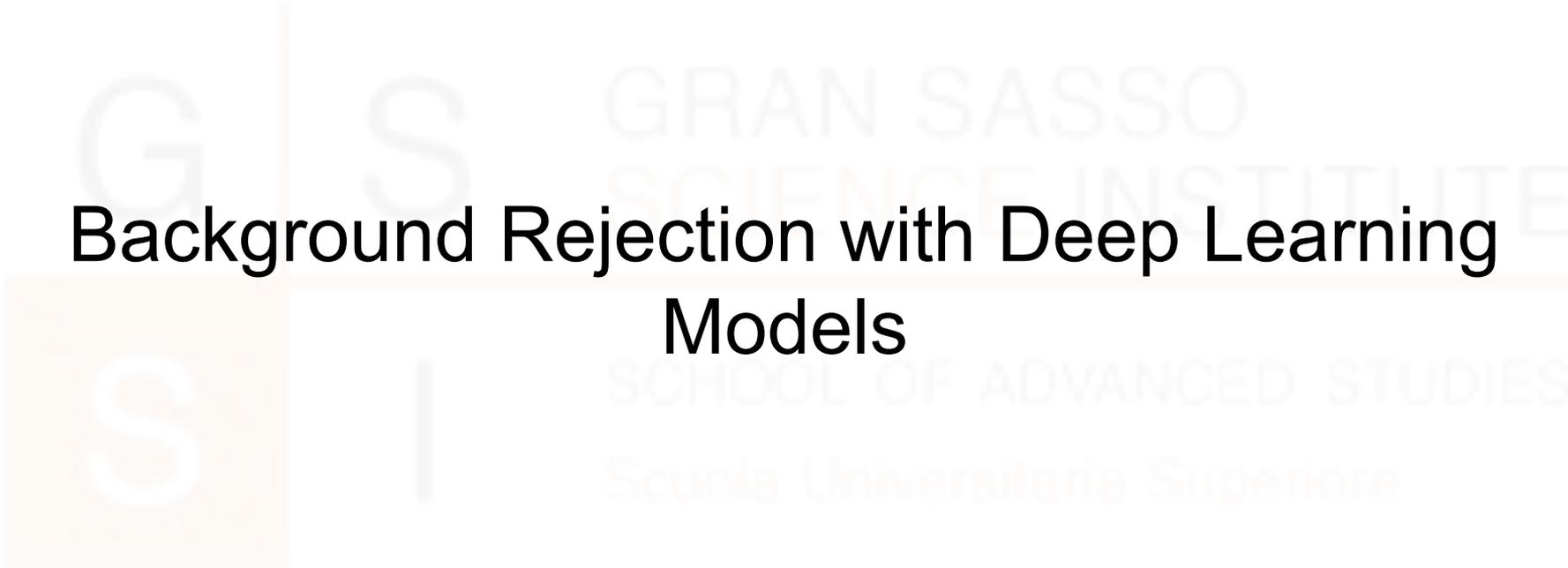
Background Rejection with Deep Learning Models

Models

DNN, RFC, GBC

- Classification

CNN

CNN

- Classification
- Semantic Segmentation

UNet, ResUNet, Att-ResUNet

- Classification
- Semantic Segmentation
- Object Detection

Mask-RCNN

- Classification
- Semantic Segmentation
- Object Detection
- Instance Segmentation
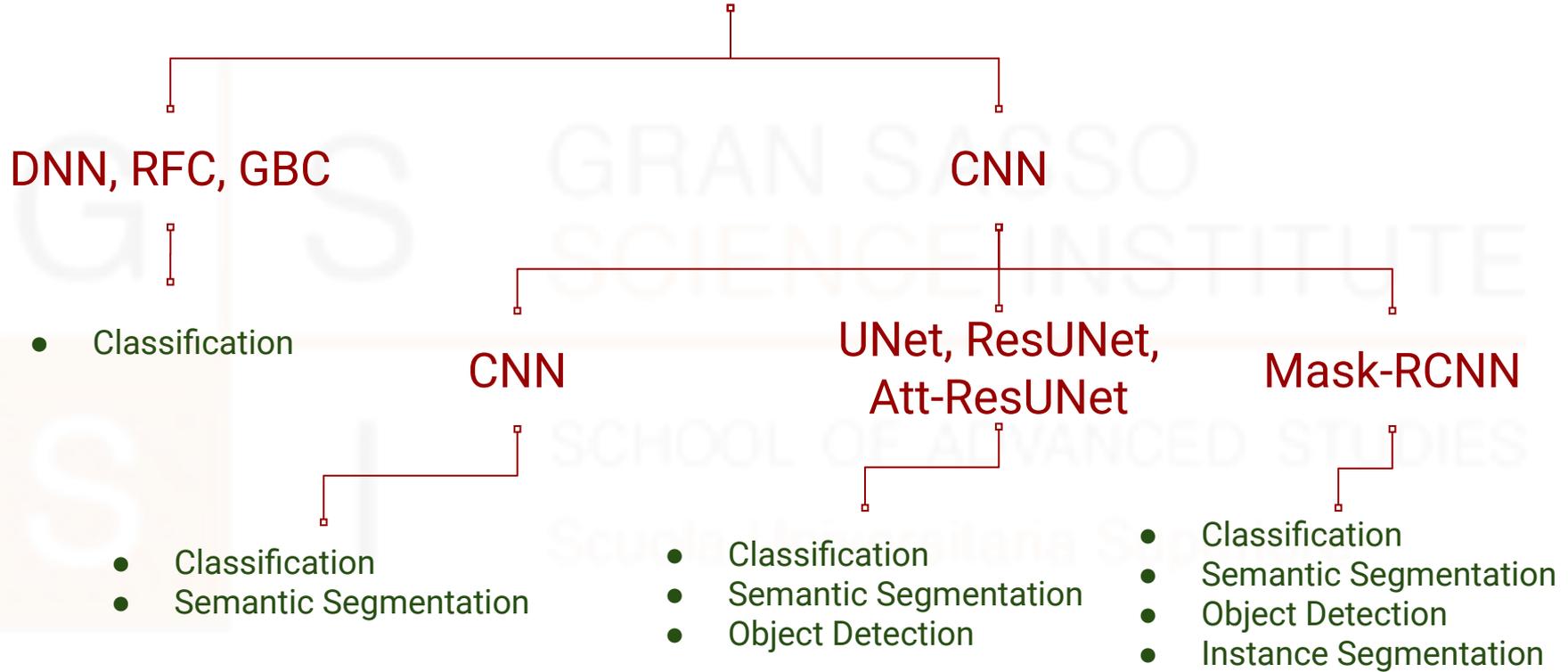
Classification	Semantic Segmentation	Object Detection	Instance Segmentation

- ❖ Models: DNN, RFC,GBC
- ❖ Discriminating variables are computed
- ❖ Classification (Classify into ER and NR)
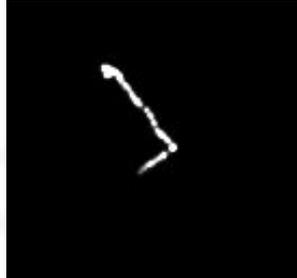
- ❖ Output of CNN, UNets, ResNet, Att-ResNet
- ❖ Classification
- ❖ Semantic Segmentation (Each pixel is classified as noise or track)
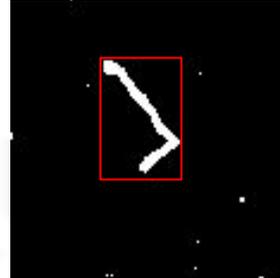
- ❖ Output of UNets, ResNet, Att-ResNet
- ❖ Classification
- ❖ Semantic Segmentation
- ❖ Object Detection (Finds a bounding box around the track and specifies if it is a ER or NR)
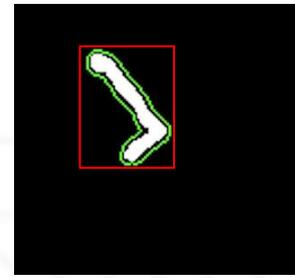
- ❖ Output of Mask-RCNN
- ❖ Classification
- ❖ Semantic Segmentation
- ❖ Object Detection
- ❖ Instance segmentation (Finds the cluster around the tracks for each object (track) detected.)

Models

DNN, RFC, GBC

- Classification

❖ Deep Neural Network

❖ Random Forest Classifier

❖ Gradient Boosted Classifier

Deep Learning Models
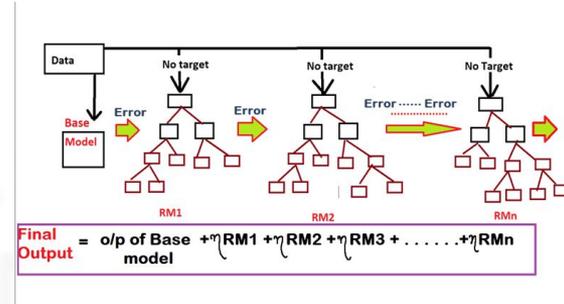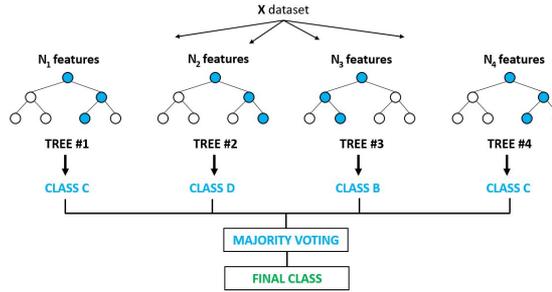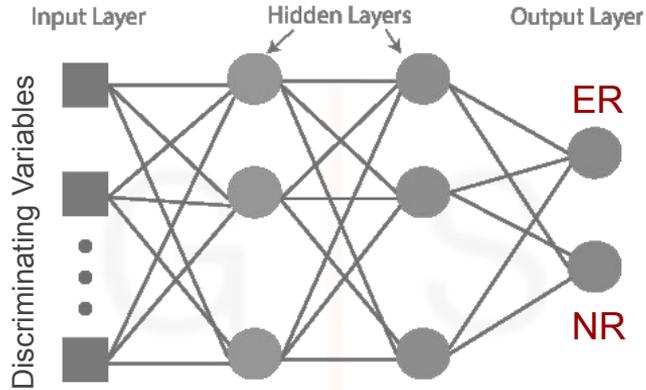
1)   Deep Neural Network

❖ Weights of the network is optimised iteratively
❖ Result is the output of the last layer.
❖ 3 hidden layers, 10 neurons in each layer

2) Random Forest Classifier

❖ It can build each tree independently.
❖ Results are combined at the end of the process.
❖ 400 trees

3) Gradient Boosted Classifier

❖ It builds one tree at a time.
❖ It combines results along the way.
❖ 400 trees

Preparing the dataset for training

ER & NR simulation	Digitization	Reconstruction	Discriminating Variables

Interaction of the particles with gas is simulated using either GEANT4 (for ER) or SRIM (for NR).

These tracks are then projected to a 2D plane and detector effects are added like diffusion, camera noise, effective ionisation, gain fluctuation and geometrical acceptance etc.

Digitized tracks are reconstructed for the tracks using a iterative density based scanning algorithm called IDBSCAN.

Reconstructed tracks are used to build several discriminating variables like skeleton, Length along principal axis, Charge uniformity, Maximum density, Slimness, Integral etc.

E Baracchini et. al., "Identification of low energy nuclear recoils in a gas TPC with optical readout", arXiv:2007.12508v1

Annual Talk   14/10/2022

Training the Models

❖ Energy range used for training is 1-36 keV for both ER and NR.

❖ 5000 events for each energy for **ER.**

❖ 3000 events for each energy for **NR.**

❖ **Variables**: thin_track, SDCD, CylThick, ChargeUnif, LAPA, MaxDen, eta, curlyness, SC_nhits, SC_integral, SC_length, SC_width, delta, slimness
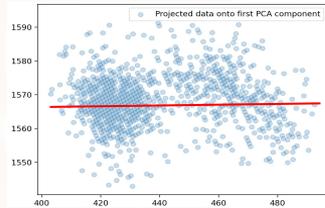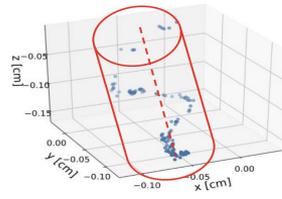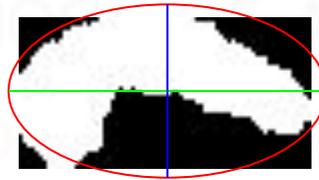
LAPA   CylThick   slimness   thin_track

Observables for recoil identification in gas TPCs: arXiv:2012.13649v1
GEM-based TPC with CCD Imaging for Directional Dark Matter Detection: arXiv:1510.02170v3

I presented the development of all the discriminating variables in previous year's presentation.

Classical Approach for Background Rejection

❖ Applying cuts on all the variables that I used for training.

Signal Events (N_{signal}) = No. of NR events from the variable passing the cut

Bkg Events (N_{bkg}) = No. of ER events from the variable passing the cut

Signal efficiency (S_{eff}) = $N_{signal}/N_{total,sig}$

Bkg. Efficiency (B_{eff}) = $N_{bkg}/N_{total,bkg}$

Bkg. Rejection = $N_{total,bkg}/N_{bkg}$

Rejection efficiency vs Signal Efficiency

Background rejection vs Signal efficiency
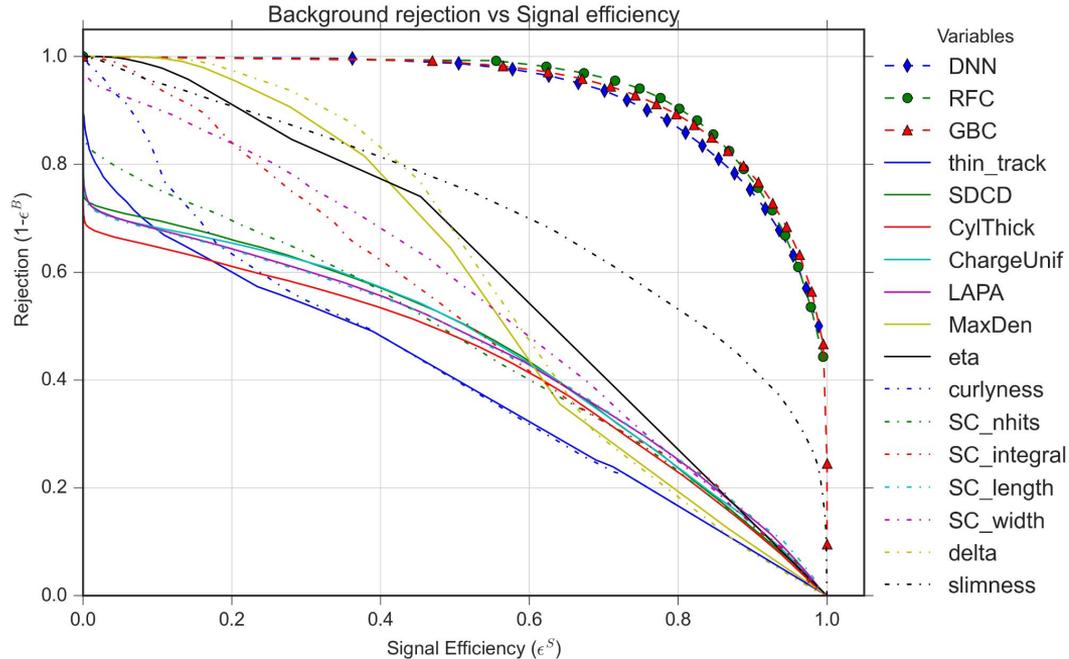
❖ All the variables shown in the plot show the rejection efficiency with classical approach.

❖ Rejection of background events was then computed at 40% and 50% signal efficiency.

Background Rejection
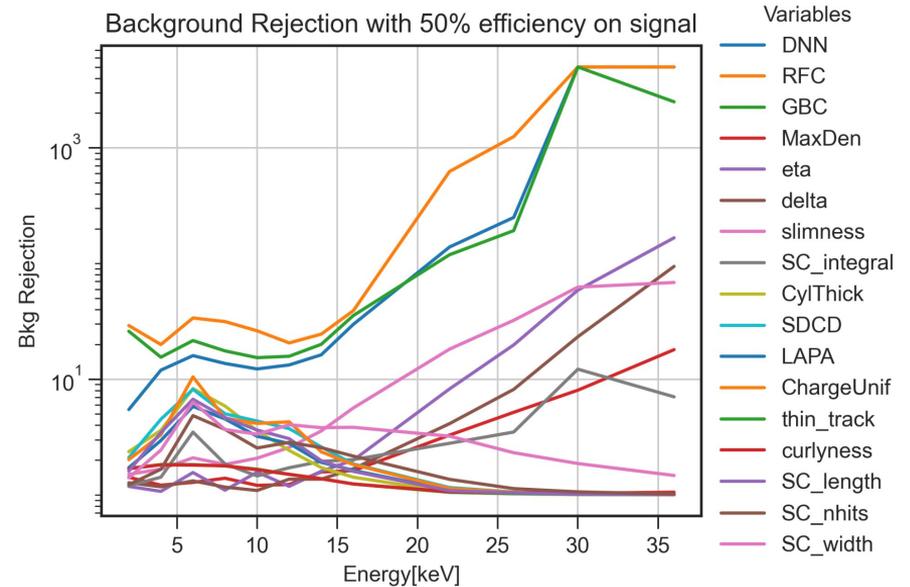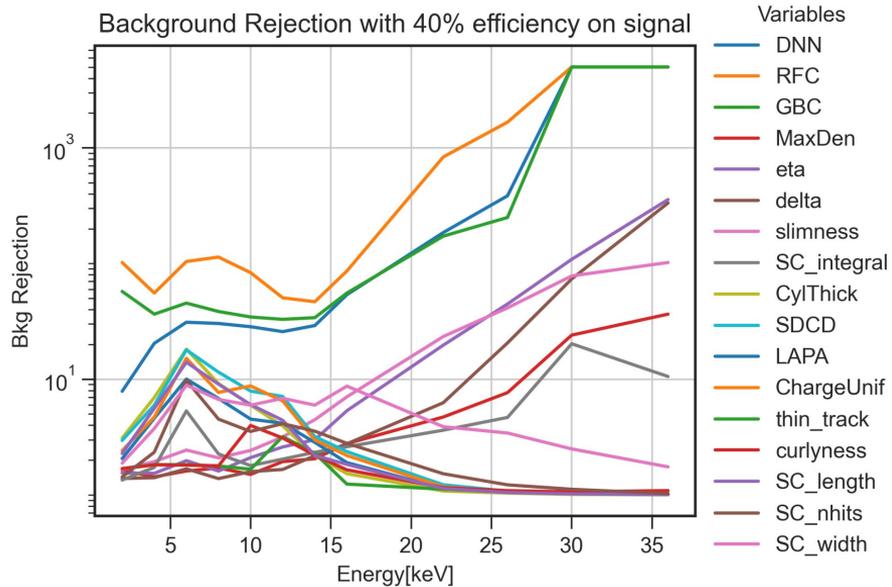
❖ Background Rejection is plotted with 40% and 50% signal efficiency in each energy bin.

❖ All the variables shown in the plot show the background rejection with classical approach.

Training and Testing Scores for all 3 models

❖ All the 3 models were trained on all 3 different datasets namely: **All**, **Variable_5**: MaxDen, eta, delta, slimness, integral.

Variable_9: thin_track, SDCD, CylThick, ChargeUnif, LAPA, curlyness, SC_nhits, SC_length, SC_width

❖ Score of training and testing of all the models is plotted for the different datasets.
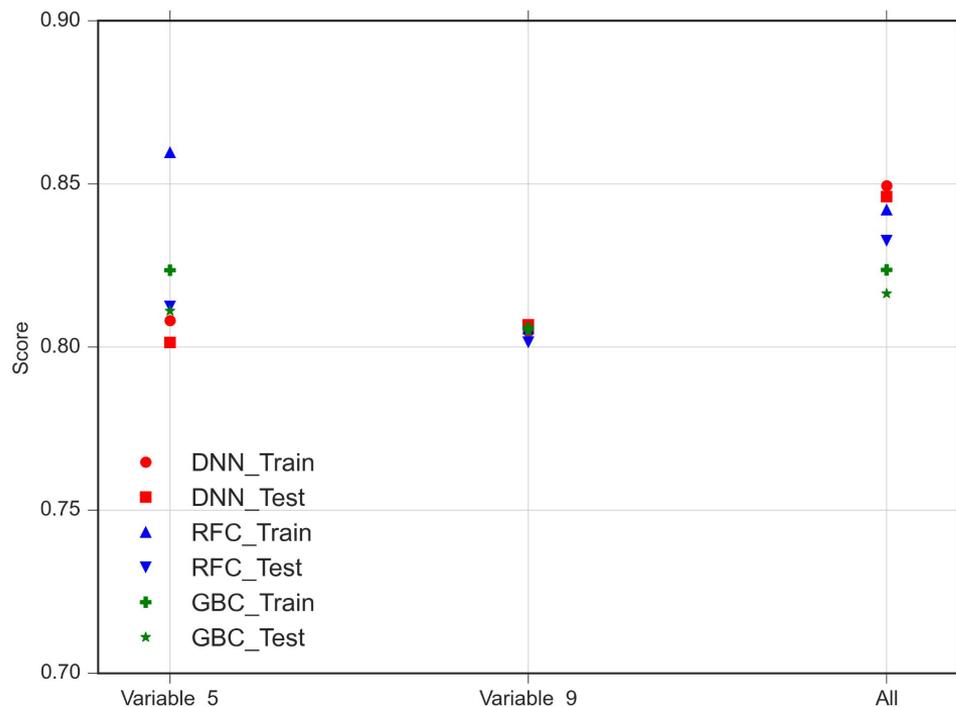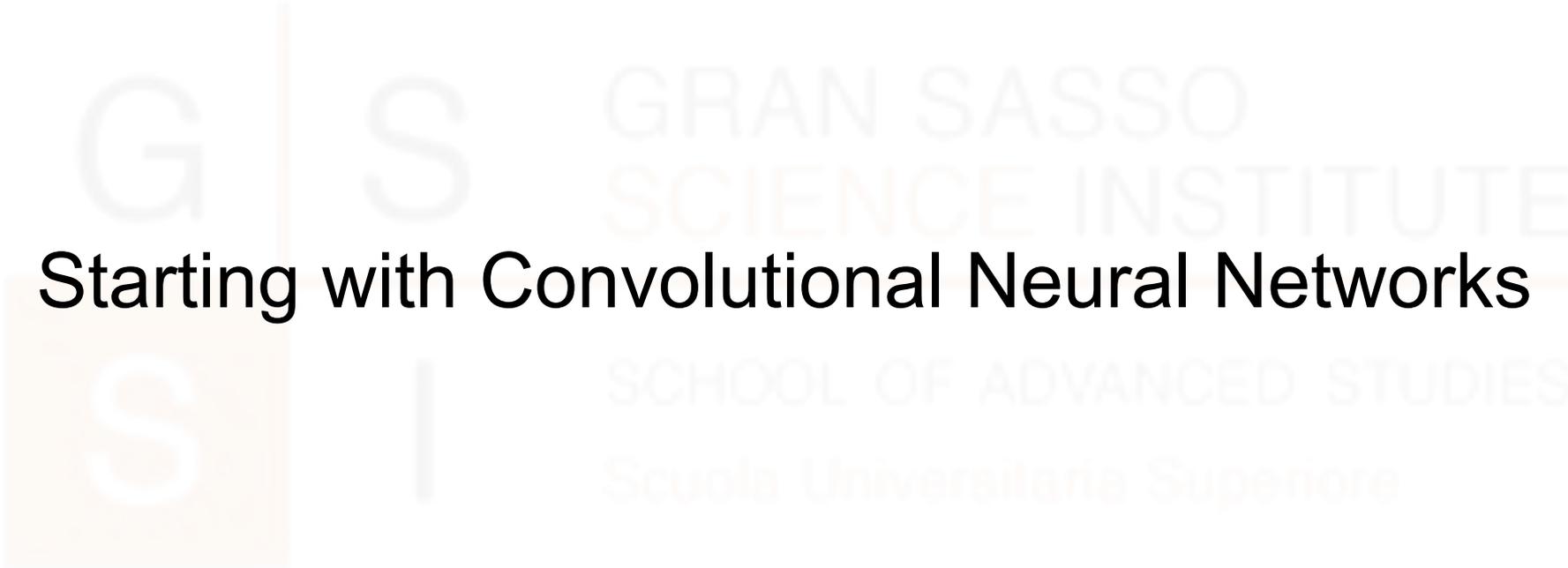
Starting with Convolutional Neural Networks

Models

CNN

CNN

- Classification
- Semantic Segmentation

❖ Convolutional Neural Networks
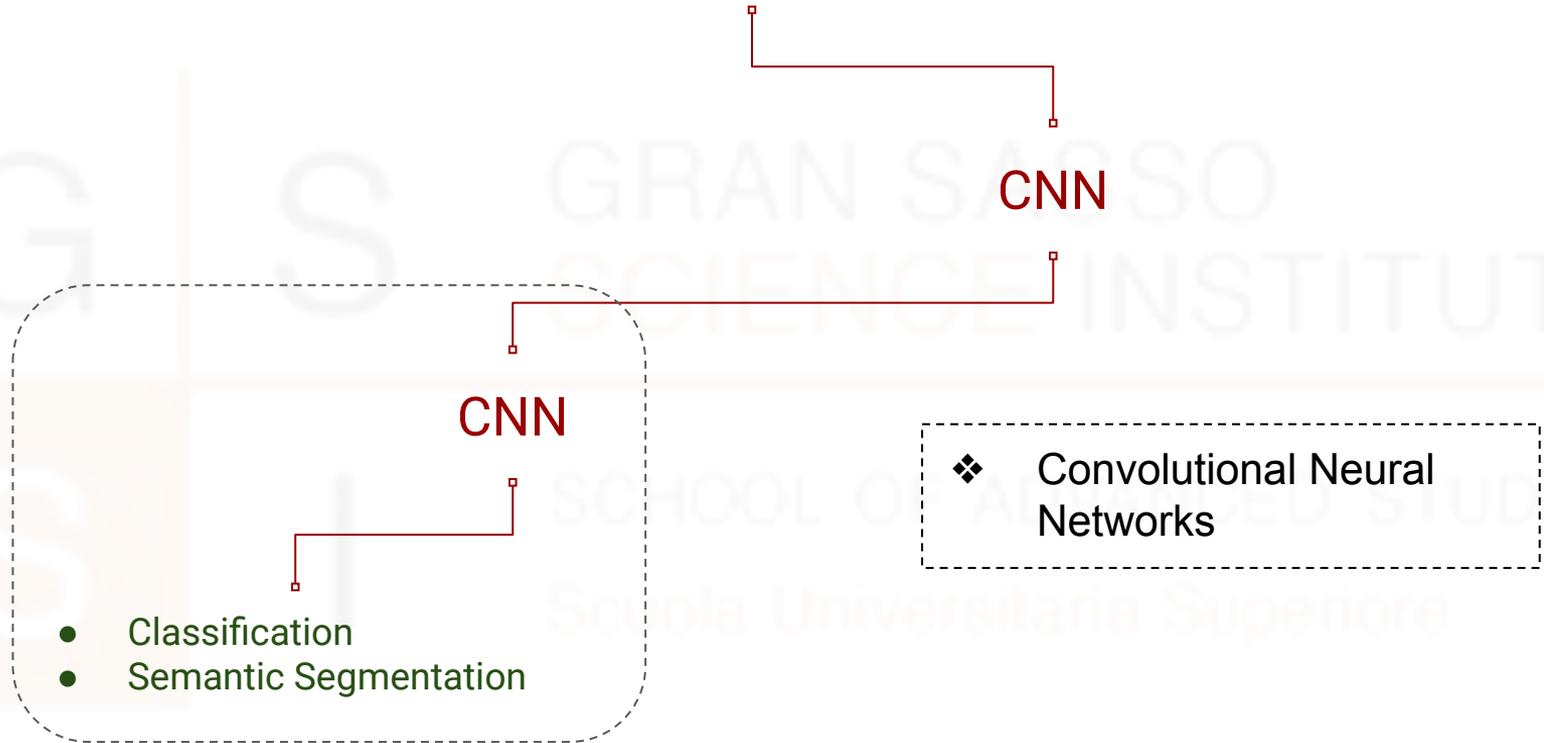
Convolutional Neural Networks
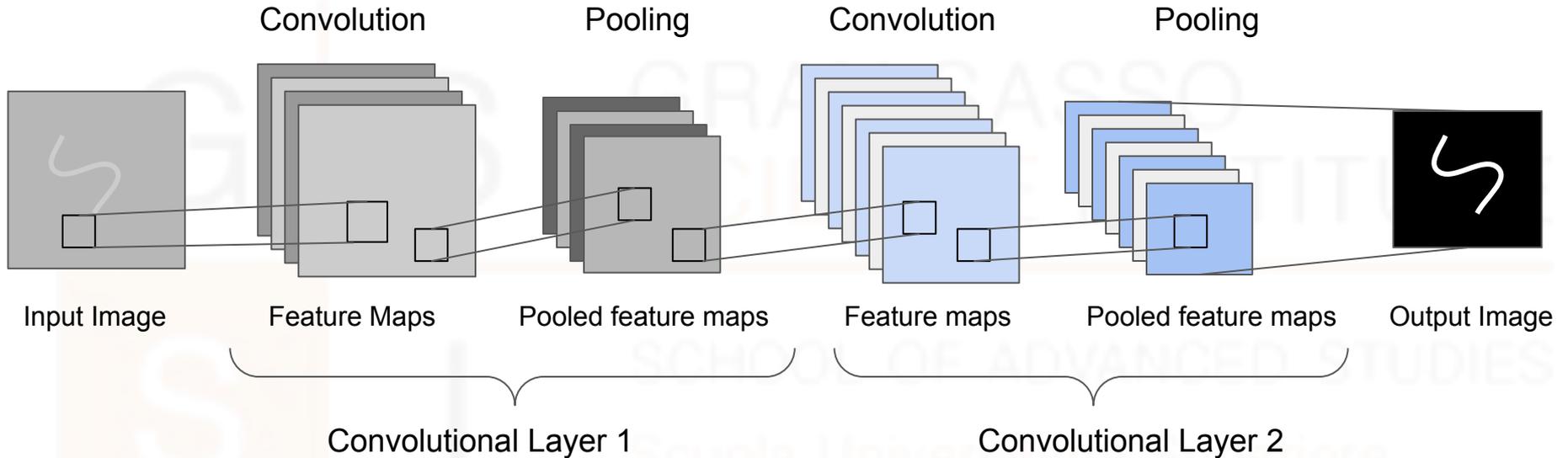
❖ A CNN (ConvNet) is a Deep Learning algorithm which can take in an input image, assign importance to various aspects/objects in the image and be able to differentiate from other.

❖ The pre-processing required in CNN is much lower as compared to other classification algorithms.

Models

CNN

UNet, ResNet,
Att-ResNet

- Classification
- Semantic Segmentation
- Object Detection

- ❖ UNets
- ❖ Residual UNets
- ❖ Attention Residual UNets

Attention Residual UNets
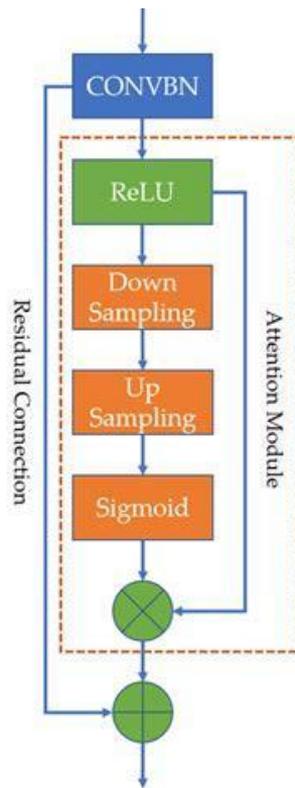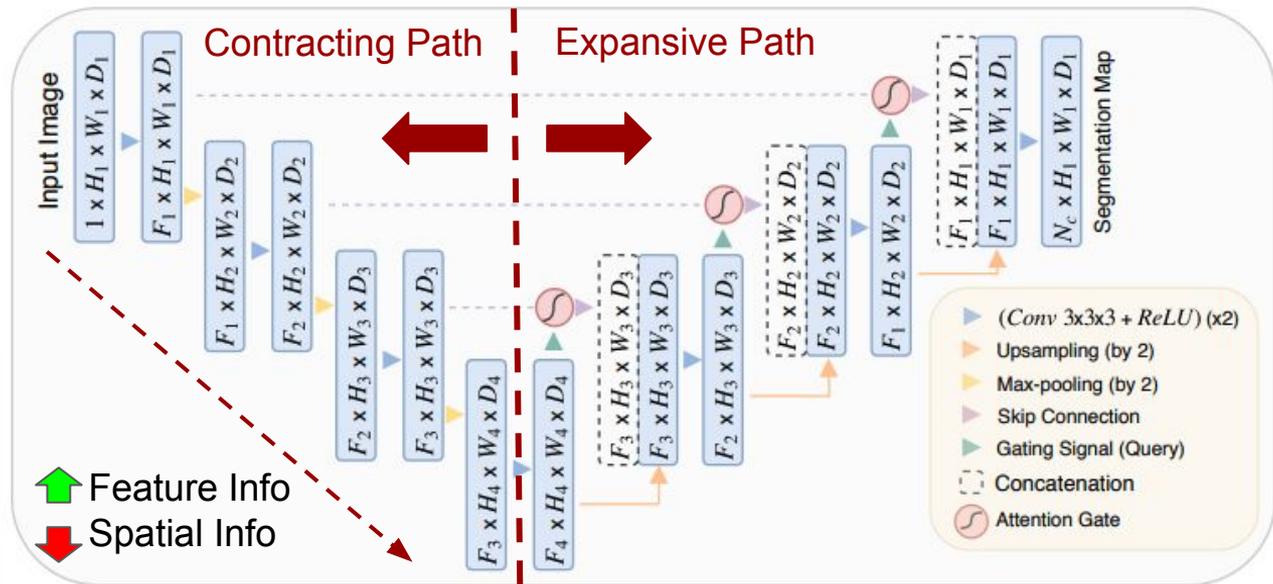
Architecture is very similar to ResNet, except there is an extra block called attention block. Attention in U-Nets is a method to highlight only the relevant activations during the training.

It reduces computation resources wasted on irrelevant activations and provides better generalization of the network.

Preparing Data for training CNNs

Input for Network   Target for Network

a) Digitized Image   b) Noise   c) Noise subtracted   d) Mask (Track
    (512x512)   (512x512)   (Blurred -> Thresholded)   w/o noise)
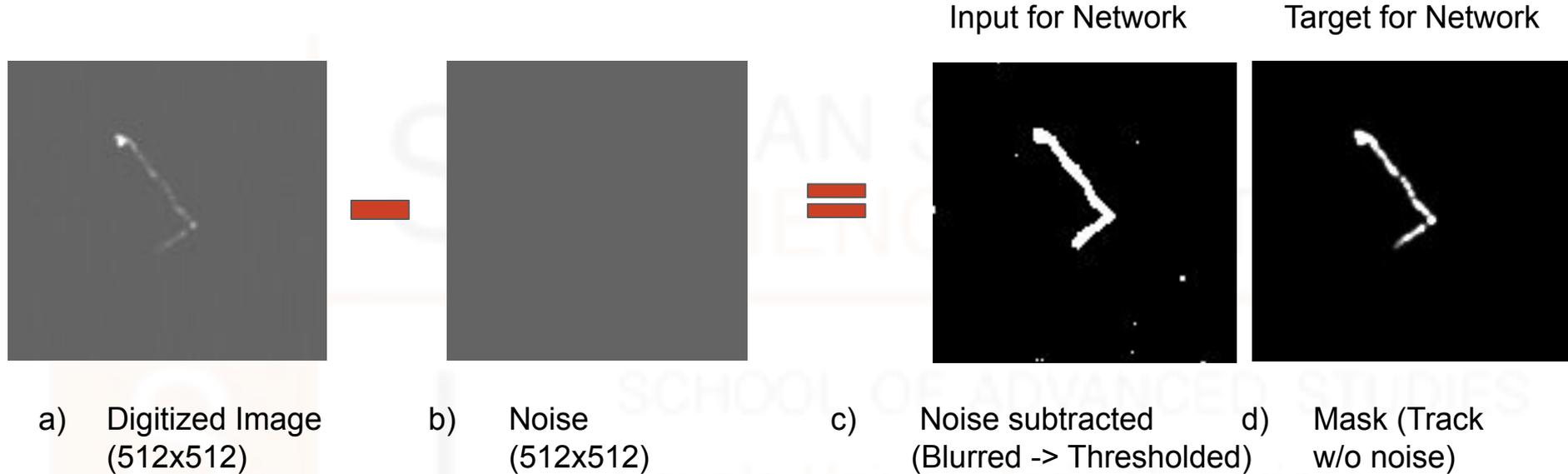
Noise map is subtracted from the digitized image. Noise subtracted image is then passed through a median filter with a kernel size of 3. Blurred image is thresholded with a threshold of 1 (pixels with intensity more than 1 becomes 255 and rest 0). These images are input for the network.

Masks are produced by digitizing the tracks without noise. Network is trained to produce images similar to masks.

Training and Validation accuracy
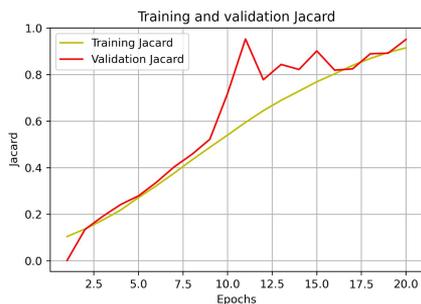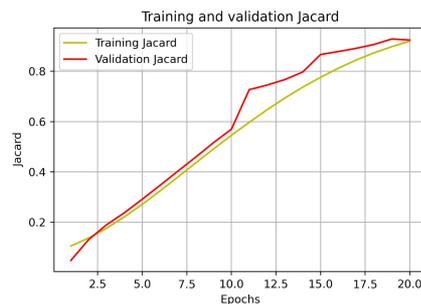
(a)   CNN   (b)   UNet   (c)   Att-UNet   (d)   Att-ResUNet
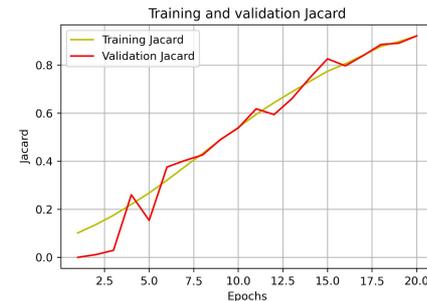
Jaccard Coefficient = Intersection over Union

$$IoU = \frac{Area\ of\ Overlap}{Area\ of\ Union}$$

Jaccard Coefficient measures the similarity between between 2 sets of data. The closer to 1 means more similar data.
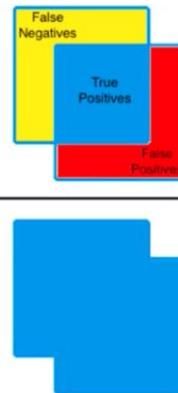
All the models were trained with 600 images and for 20 epochs.

Prediction from Att-ResUNet

60 keV ER

1 keV NR

Input Image   Target   Prediction of the Network

Reconstruction

Att-ResUnet
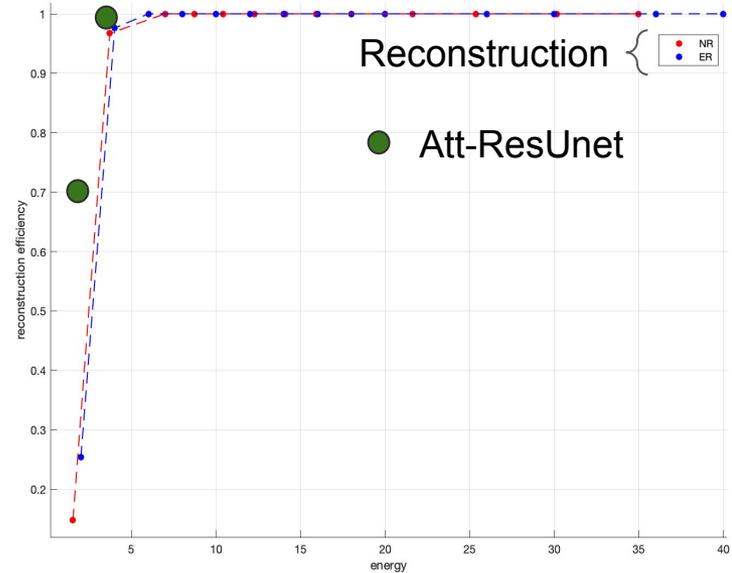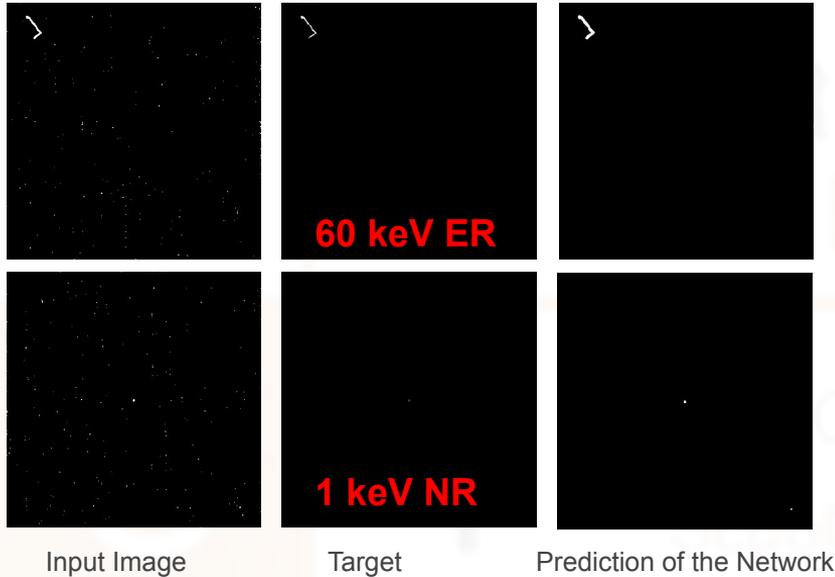
- ❖ Predicted images were used to find the cluster around the track using OpenCV.
- ❖ Reconstruction efficiency at 1 keV of NR is ~ 70% which is ~10% with usual reconstruction algorithm and at 3 keV NR is 100% and with usual reconstruction algorithm it is ~97%.

Models

CNN

Mask-RCNN

❖ Mask - Region based Convolutional Neural Networks (Mask-RCNN)

- Classification
- Semantic Segmentation
- Object Detection
- Instance Segmentation

Architecture of Mask-RCNN

RPN

region proposal

regions

CNN

feature maps

image

RoI Align

FC Layers

FC

classes (softmax)

ER or **NR** Classifier

FC

boundary box regressor

BBox regressor

conv. conv. mask

Maskifier

Mask

Backbone: ResNet

Mask-RCNN Paper: https://doi.org/10.48550/arXiv.1703.06870

Data for Mask-RCNN

a)　Digitized Image
(512x512)

b) Noise
(512x512)

c) Noise subtracted
(Blurred -> Thresholded)

Input for Network

Digitized image
w/o noise

Selecting the
cluster using OpenCV

❖ **Shape Attributes**
  ➢ Polygon
  ➢ Coordinates of polygon
  ➢ bbox
  ➢ area
❖ **Region Attributes**
  ➢ ER or NR (category)
  ➢ Name & index

Target for Network

First results from Mask-RCNN
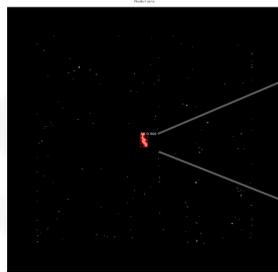
Input   Prediction   Zoomed Track
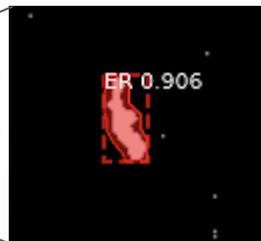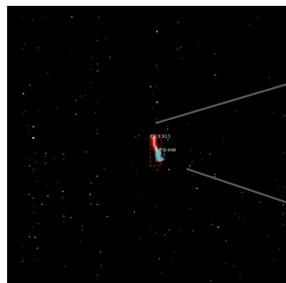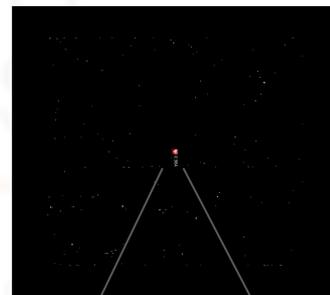
Model was trained just with 4 input images and for 20 epochs.

Future Work

❖ We are working on simulating larger samples for training the networks to test the rejection capabilities of the network

❖ To train Mask-RCNN with larger sample of simulated data

❖ Working with Gustavo V. Lopez from CS department at GSSI to build a network in which we can integrate the sCOMS data and PMT data

❖ Test all these models on data

❖ Will work with the negative ions data from MANGO detector

References

1. G. Bertone et. al., "Particle Dark Matter: Evidence, Candidates and constraints",
   arXiv:hep-ph/0404175
2. V.C. Antochi et. al., "Combined readout of a triple-GEM detector", JINST13 (2018) 05, P05001
3. Physics Reports Volume 627, 20 April 2016, Pages 1–49
4. E. Baracchini et al.: Measur.Sci.Tech.32(2021)2,025902
5. C. J. Martoff et. al., "Suppressing drift chamber diffusion without magnetic field", Nucl. Instr. and Meth. in Phys. Res. A, 440 (2000), p. 355.
6. E Baracchini et. al., "Identification of low energy nuclear recoils in a gas TPC with optical readout",
   arXiv:2007.12508v1
7. M. Ghrear et. al., "Observables for Recoil Identification in High-Definition Gas Time Projection Chambers",
   arXiv:2012.13649v1
8. N. S. Phan et. al., "GEM-based TPC with CCD Imaging for Directional Dark Matter Detection",
   arXiv:1510.02170v3

Schools/Conferences

[1] School on Underground Physics, SOUP 2021, 28 June - 2 July (Online)

[2] 4th International School on Deep Learning, DeepLearn 2021 Summer, Las Palmas de Gran Canaria, Spain, 26 July - 30 July (In person)

[3] Quantum Sensors for Fundamental Physics, QSFP 2021, 6 September -17 September (Online)

[4] Advanced Computing and Analysis Techniques in Physics Research (ACAT) 2021, 29 November - 3 December 2021, (Online)

- **Presented a Poster and won the best poster award**

[5] CYGNO Collaboration meeting, 21-22 Dec 2021 at GSSI. (Online)

- **Oral Presentation**

[6] SiPM Radiation: Quantifying Light for Nuclear, Space and Medical Instruments under Harsh Radiation Conditions, 25-29 April 2022 (Online)

[7] INFN School of Statistics 2022, Paestum, 15-20 May 2022. (In person)

[8] International conference on Machine Learning for Astrophysics - ML4Astro, Catania, 30 May - 1 June 2022. (In person)

- **Presented a poster**

[9] PyHEP 2022 Workshop, 12- 16 Sept 2022. (Online)

[10] Micro Pattern Gaseous Detectors 2022, 11 -16 December 2022, Submitted an abstract for oral presentation.

Publications

[1] F. D. Amaro, E. Baracchini et. al., "*Directional Dark Matter Searches with CYGNO"*, MDPI, July 2021, DOI: 10.3390/particles4030029

[2] F. D. Amaro, E. Baracchini et. al., "*The CYGNO Experiment",* MDPI, February 2022, arXiv:2202.05480

[3] F. D. Amaro et al., "*Study of Performance of Different Photodetectors and Electrical Signal for Fast Detection in Fifty Litres CYGNO Prototype (LIME)*," IEEE, 2021, pp. 1-5, doi:10.1109/NSS/MIC44867.2021.9875803

[4] F. D. Amaro et. al., "*Performances of a 3D optical readout TPC for the CYGNO experiment*", 2022, DOI: https://doi.org/10.22323/1.398.0799

[5] C.A.J. O'Hare, D. Loomba et. al., *Recoil imaging for directional detection of dark matter, neutrinos, and physics beyond the Standard Model,* July 2022, arXiv:2203.05914v3 (Submitted)

[6] **A. Prajapati** et. al., *Reconstruction and Particle Identification with CYGNO Experiment*, Springer Nature, Astrophysics and Space Science Proceedings, August 2022 (Submitted)
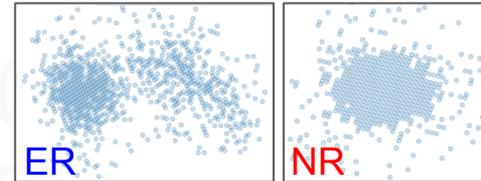
Backup

Variables

❖ Standard Deviation of Charge Distribution 2D(SDCD_2D):

$$SDCD = \sqrt{\frac{\sum_{i=1}^{N}(\mathbf{r_i} - \bar{\mathbf{r}})^2}{N}}.$$

ER   NR

➢ Electron recoils (ER) are longer, so the spread of charge is higher for ER when compared to Nuclear recoils (NR).

❖ Charge Uniformity 2D (ChargeUnif_2D):
➢ For each point within the charge distribution, find the average distance to all other points.
➢ ChargeUnif_2D is standard deviation of values computed in step 1.
➢ Electron recoils tend to have charge distribution which is dense in some areas and sparse in other areas, while nuclear recoils are generally uniform.

❖ Maximum Density 2D (MaxDen_2D):
➢ MaxDen is the value of most intense pixel from the image after rebinning it by a factor 2.
➢ Electrons lose their energy at a slower rate than nuclei, this suggests that electron recoils are travel greater distance between interactions resulting in more sparse energy distribution.
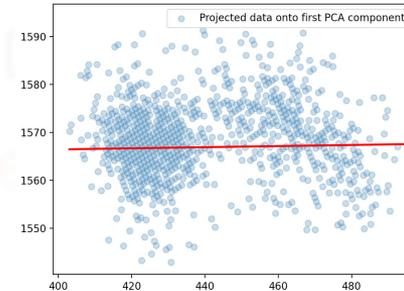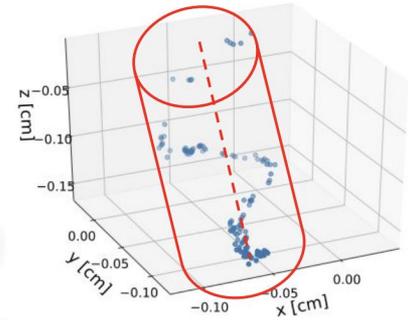
Variables

❖ Cylindrical Thickness 2D (CylThick_2D):
- ➢ For each charge, calculate the squared distance from the principal axis.
- ➢ CylThick is the sum of all squared distances.
- ➢ It is a measure of how much a recoil track deviates from the trajectory approximated by the principal axis.
- ➢ Electrons experience far more scattering compared to nuclei, so principal axis approximates NR's trajectory much more accurately than it does for ER.

❖ Length Along Principal Axis 2D (LAPA_2D):
- ➢ Project all the points in the charge distribution on to the principal axis.
- ➢ LAPA is the difference between maximum and minimum projected value.
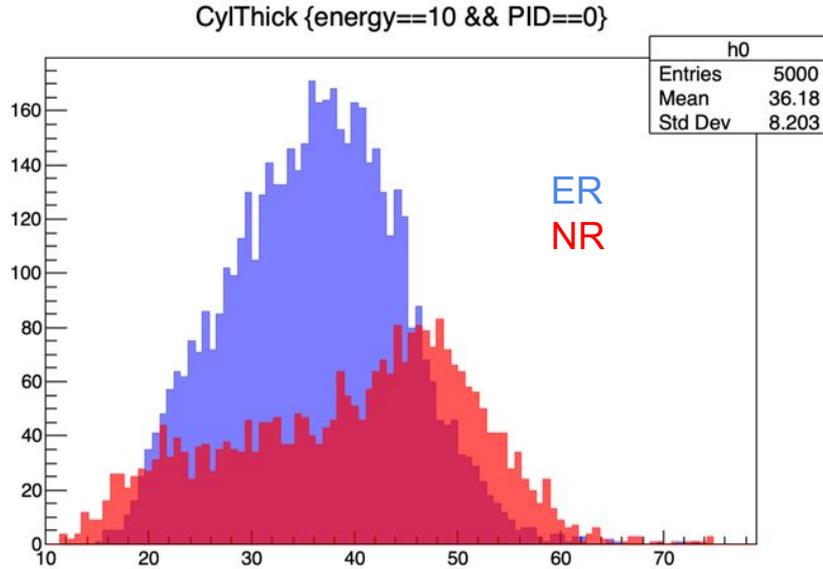- ➢ ER are longer compared to NRs, therefore projection is also longer.

Variables

❖ **eta:**
  ➢ MaxDen_2D divided by length (found by skeletonization)

❖ **Light Density (delta):**
  ➢ Integral of the track divided by number of pixels in the track.
  ➢ NR deposit higher energy over a short distance, therefore Light Density is higher for NR.

❖ **Slimness:**
  ➢ Ratio of minor over major axis of the ellipse which bounds the track.
  ➢ Electrons recoils suffer more scattering, so minor axis of the bounding ellipse is bigger when compared to NR which are generally straight.

❖ **Skeleton length (thin_track):**
  ➢ Length in pixels found by thinning procedure.

GEM-based TPC with CCD Imaging for Directional Dark Matter Detection
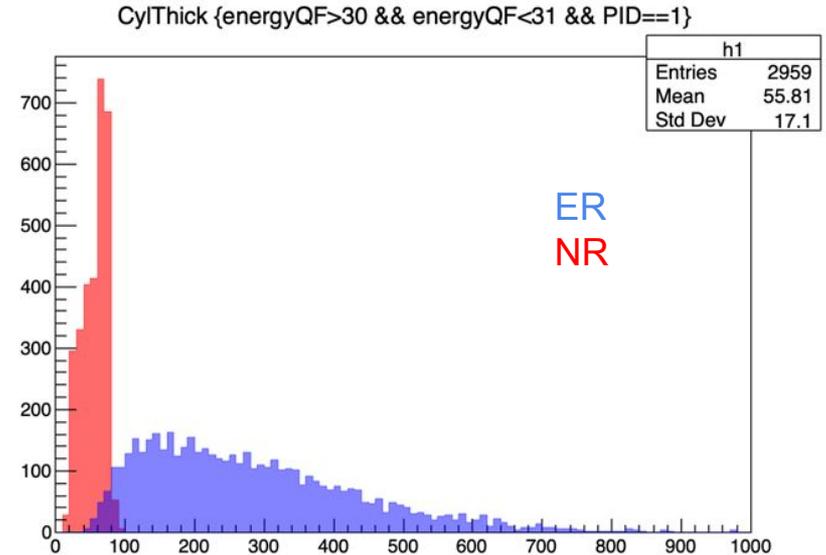arXiv:1510.02170v3

E Baracchini et. al., "Identification of low energy nuclear recoils in a gas TPC with optical readout", arXiv:2007.12508v1
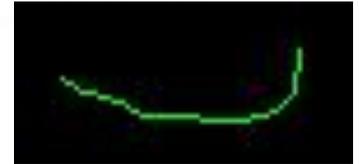
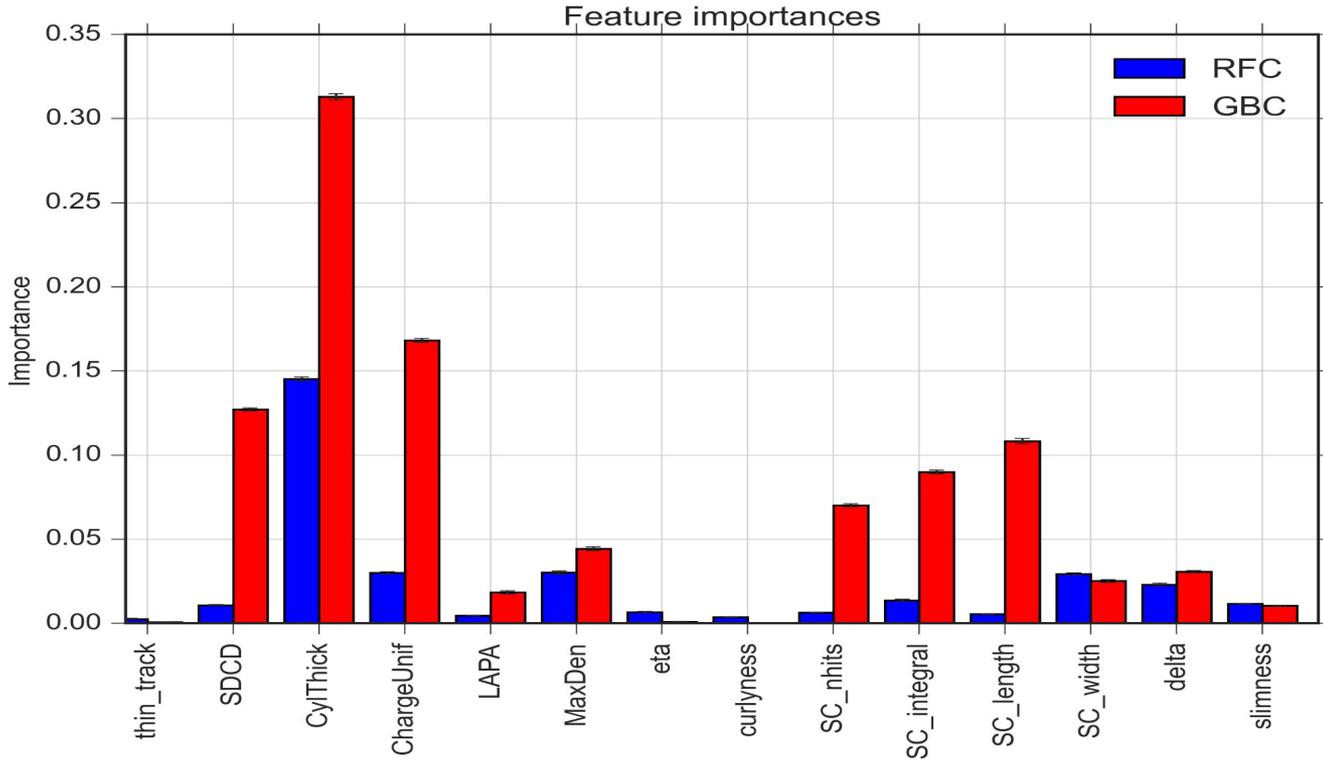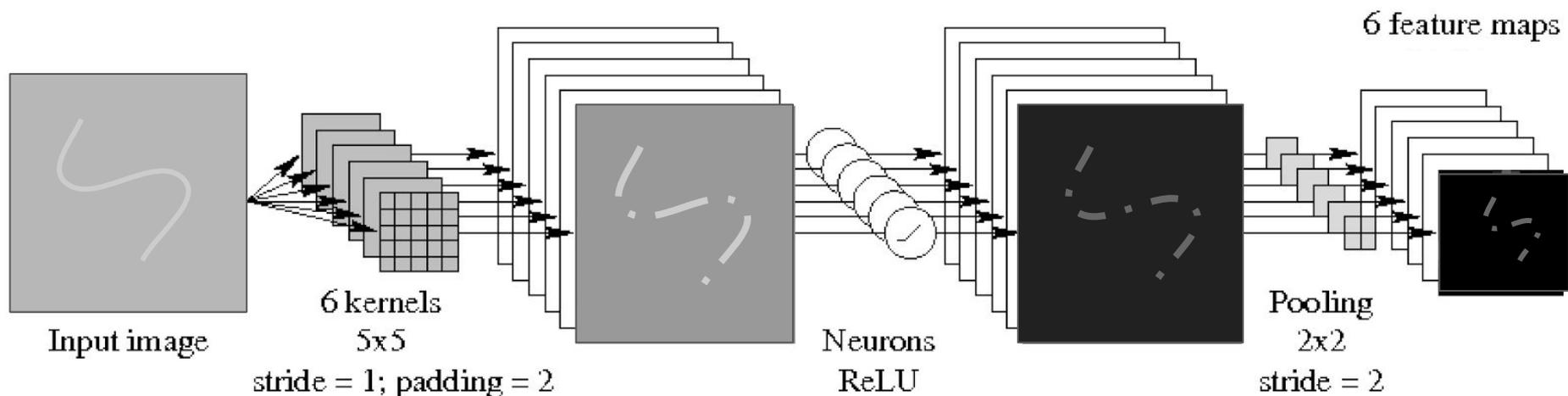# Variables with decreasing rejection at higher energy

CylThick @ 10 keV

CylThick @ 30 keV

Feature Importance

Convolutional Layer

6 feature maps

Input image — 6 kernels 5x5 stride = 1; padding = 2 — Neurons ReLU — Pooling 2x2 stride = 2

**Feature Maps**: Features maps are generated by applying filters (kernels) to the input image. Filters try to gain some understanding of what features our CNN detects.
Activation Function: Activation functions decide if the neuron would fire or not.
Pooling: Pooling reduces the number of parameters and computation in the network, controlling overfitting by progressively reducing the spatial size of the network.

UNet
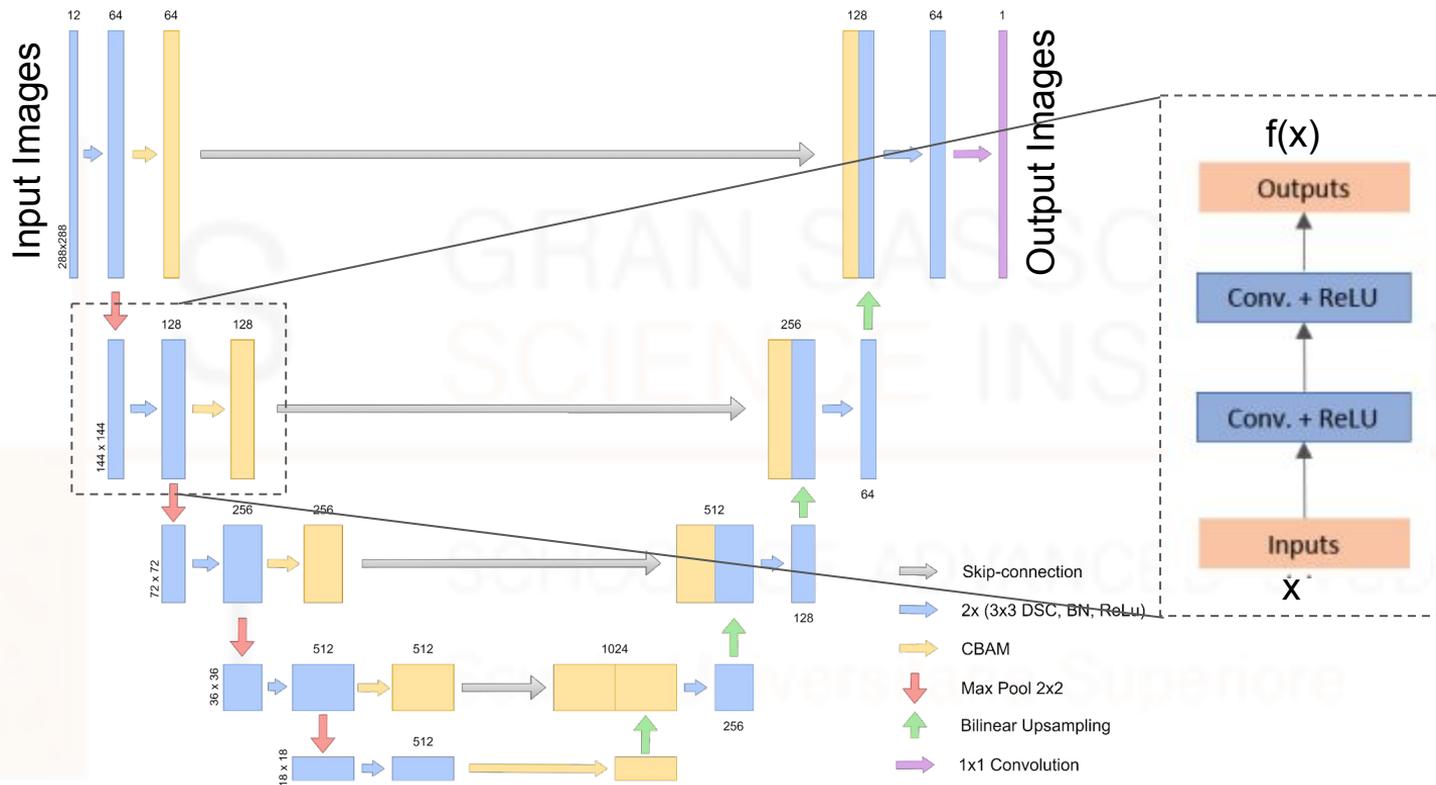
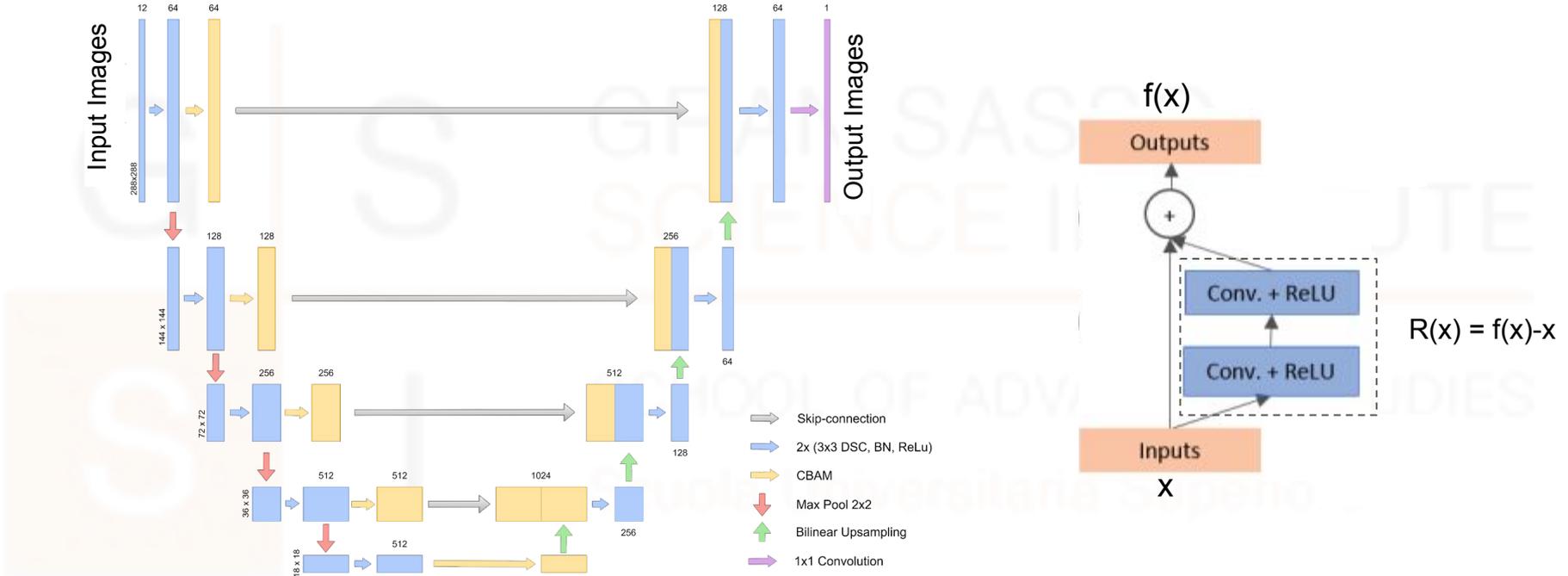Layers are trying to learn f(x) for the given input x.

Recurrent Residual Convolutional Neural Network based on U-Net (R2U-Net) for Medical Image Segmentation
https://arxiv.org/pdf/1802.06955.pdf

ResUNet
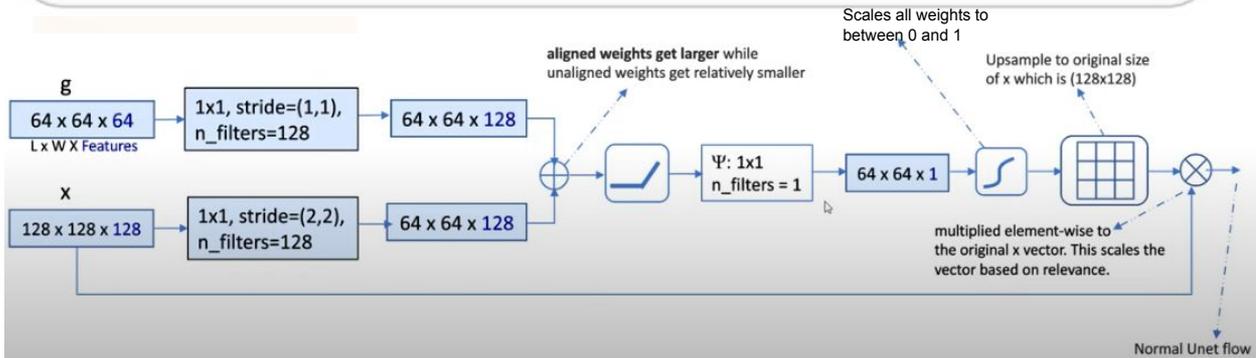

Here, Layers are trying to learn the residual unlike UNet where they try to learn f(x). ResNet helps in solving the problem of vanishing gradients and also of overfitting to an extent.

Attention Block

$$F_g \times H_g \times W_g \times D_g$$

$$W_g : 1x1x1$$

ReLU (σ_1)

$$F_{int} \times H_g W_g \, D_g$$

$$\psi : 1x1x1$$

Sigmoid (σ_2)

$$H_g W_g \, D_g$$

Resampler

$$H_x \times W_x \times D_x$$

$$\alpha \qquad \hat{x}^l$$

$$x^l$$

$$W_x : 1x1x1$$

$$F_l \times H_x \times W_x \times D_x$$
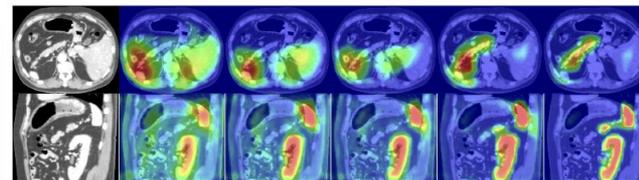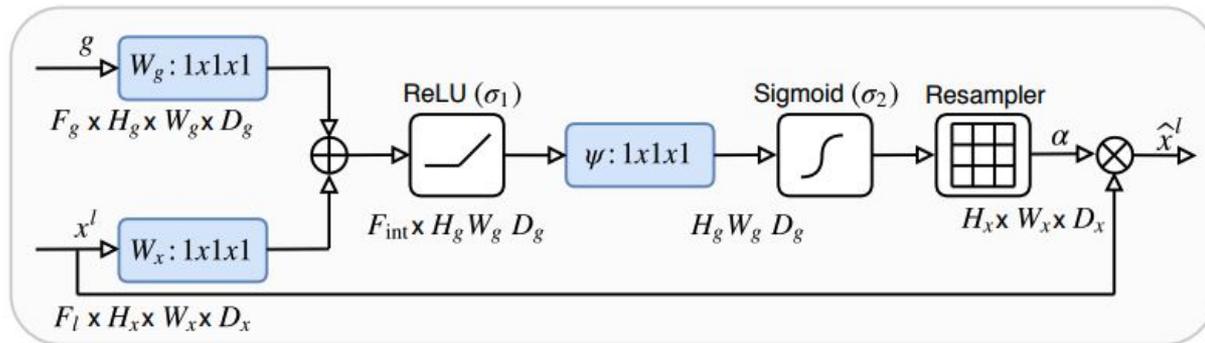
Attention U-Net: Learning Where to Look for the Pancreas
https://arxiv.org/pdf/1804.03999.pdf

g
64 x 64 x 64
L x W X Features

1x1, stride=(1,1), n_filters=128

64 x 64 x 128

aligned weights get larger while unaligned weights get relatively smaller

Scales all weights to between 0 and 1

Upsample to original size of x which is (128x128)

x
128 x 128 x 128

1x1, stride=(2,2), n_filters=128

64 x 64 x 128

Ψ: 1x1 n_filters = 1

64 x 64 x 1

multiplied element-wise to the original x vector. This scales the vector based on relevance.

Normal Unet flow

It reduces computation resources wasted on irrelevant activations and provides better generalization of the network.