Rejecting Electron Background using Machine Learning algorithms

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

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

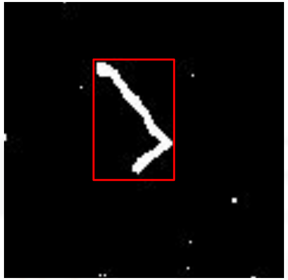
Classification

- Models: DNN, RFC,GBC
- Discriminating variables are computed
- Classification (Classify into ER and NR)

Semantic Segmentation

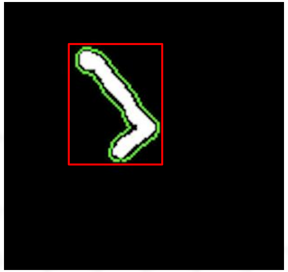
- Output of CNN, UNets, ResNet, Att-ResNet
- Classification
- Semantic Segmentation (Each pixel is classified as noise or track)

Object Detection



- 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)

Instance Segmentation



- Output of Mask-RCNN
- Classification
- Semantic Segmentation
- Object Detection
- Instance segmentation (Finds the cluster around the tracks for each object (track) detected.)

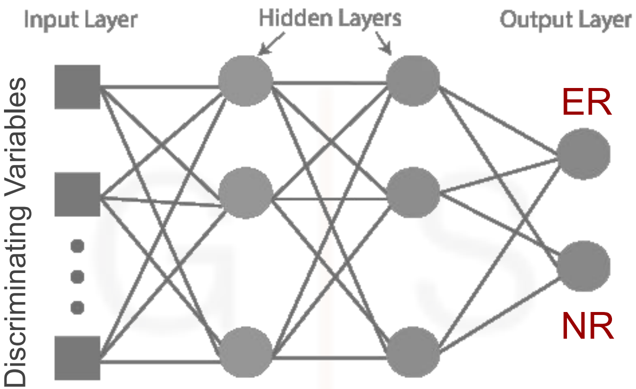
Models

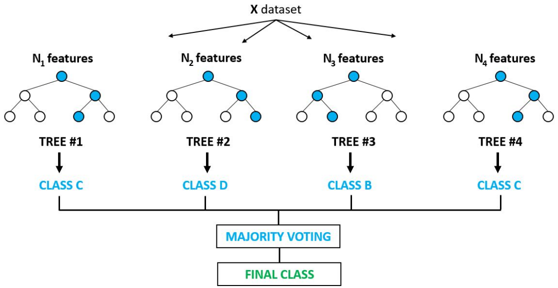
DNN, RFC, GBC

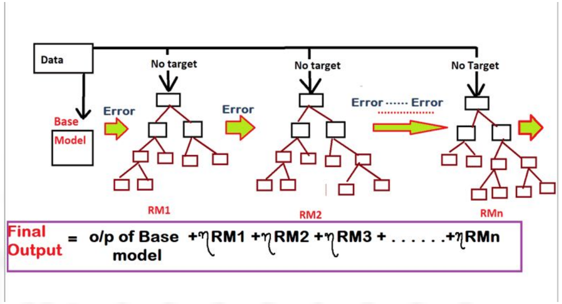
- Classification

- ❖ 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

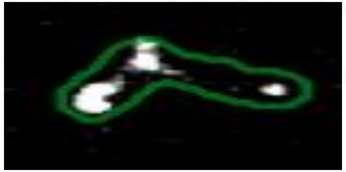
- ❖ It can build each tree independently.
- ❖ Results are combined at the end of the process.
- ❖ 400 trees

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



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

Digitization: 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.

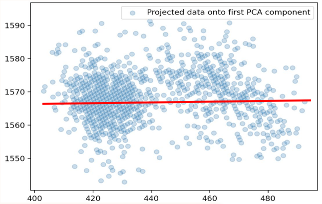
Reconstruction: Digitized tracks are reconstructed for the tracks using a iterative density based scanning algorithm called iDBSCAN.

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

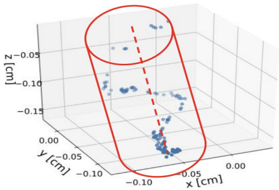
E Baracchini et. al., “Identification of low energy nuclear recoils in a gas TPC with optical readout”, arXiv:2007.12508v1

Training the Models

- Energy range used for training is 2-36 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

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

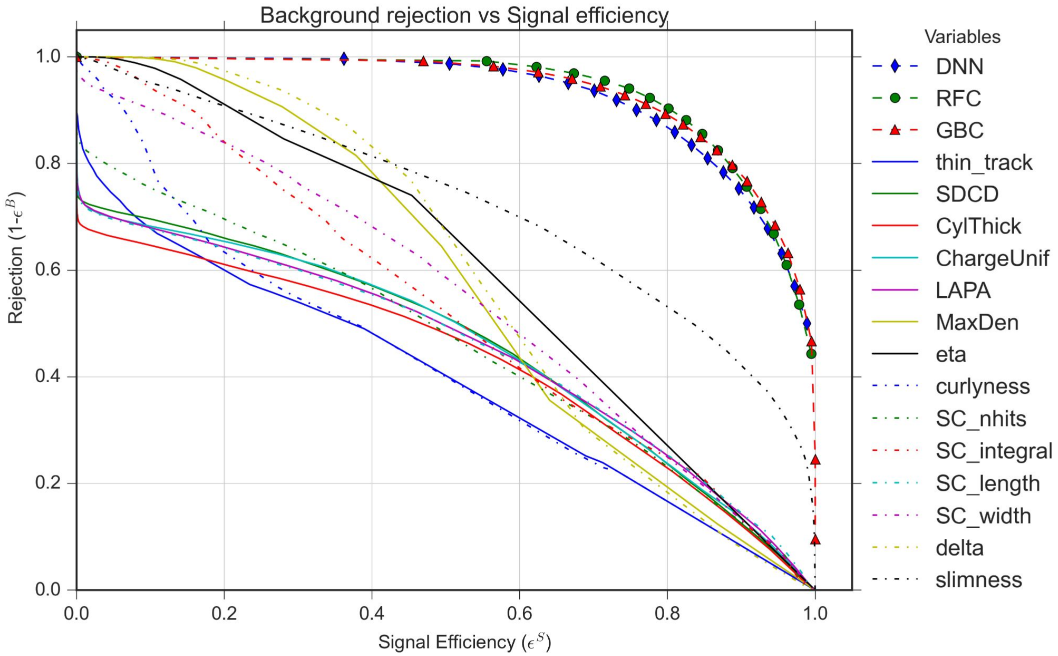
  Bkg Events (N_{bkg}) = No. of ER events from the variable passing the cut

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

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

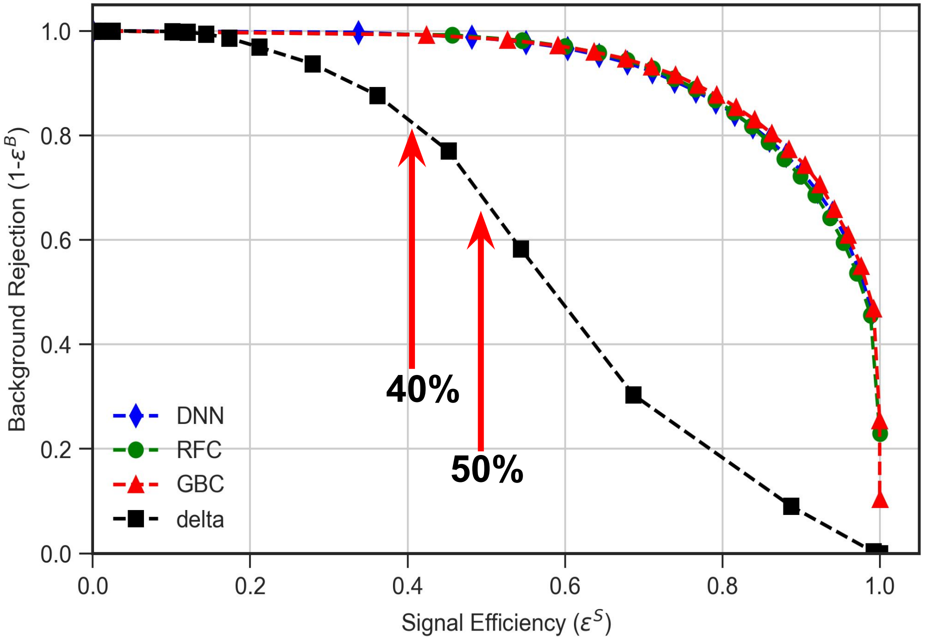
  Bkg. Rejection = $1 - B_{eff}$

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



Models	Signal Eff. $[\epsilon^S]\%$	Bkg. Rej. $[1\text{-}\epsilon^B]\%$
RFC	40	99.54
	50	98.78
GBC	40	99.38
	50	98.55
DNN	40	99.43
	50	98.50
Cut-based	40	83.13
	50	67.20

working point	Signal efficiency			Background efficiency			Bkg. Rej
	ε_S^{presel}	ε_S^{δ}	ε_S^{total}	ε_B^{presel}	ε_B^{δ}	ε_B^{total}	
WP_{50}	0.98	0.51	0.50	0.70	0.050	0.035	96.5 %
WP_{40}	0.98	0.41	0.40	0.70	0.012	0.008	99.2 %

Results are for simulated data in range 2-40 keV for NR and ER. While results published in paper mentioned below is for NR energy range of 1-100 keV discriminated against 6 keV ER.

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

Starting with Convolutional Neural Networks

Models

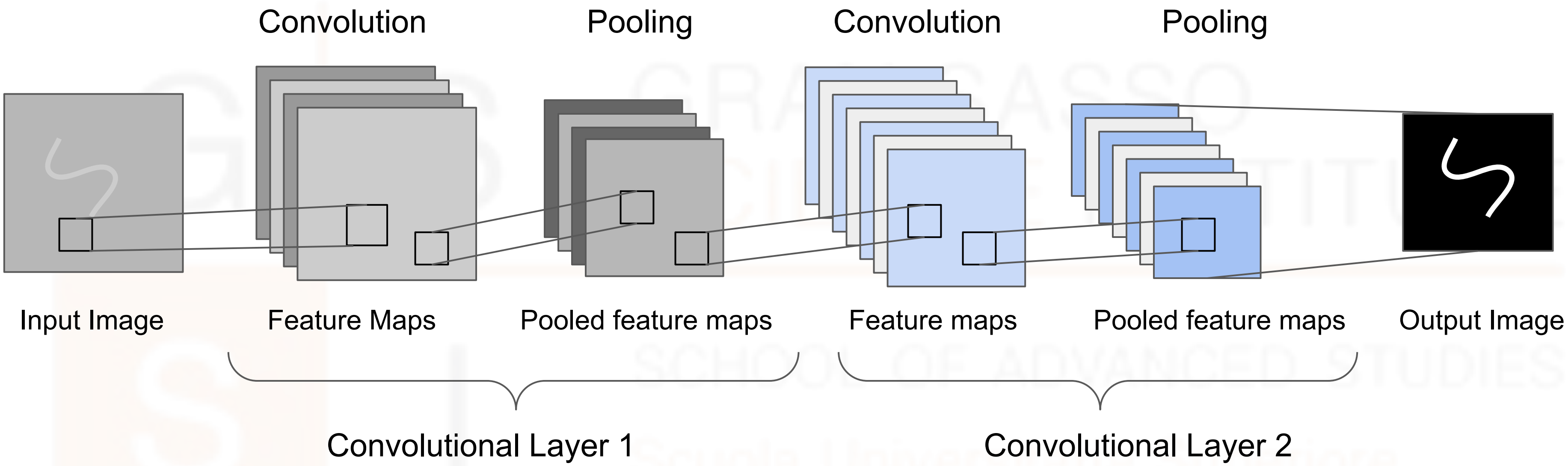
CNN

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

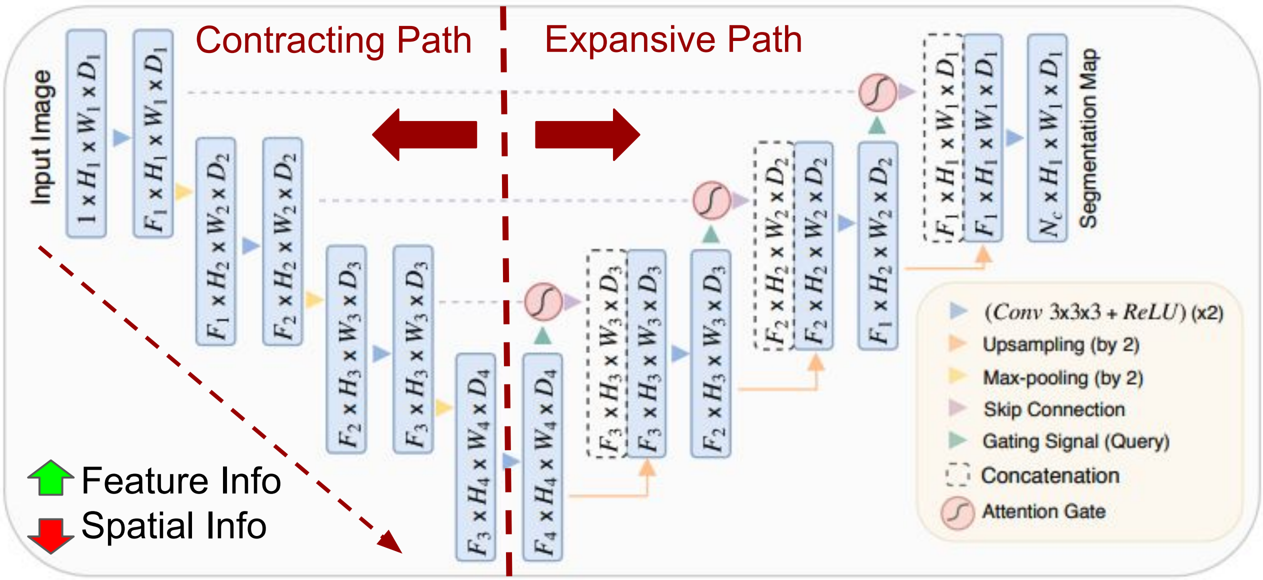
- Classification
- Semantic Segmentation
- Object Detection

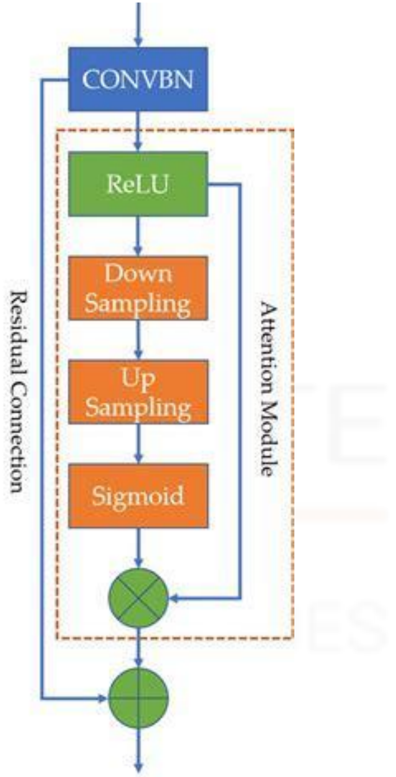
❖ UNets

❖ 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 (512x512)

b) Noise (512x512)

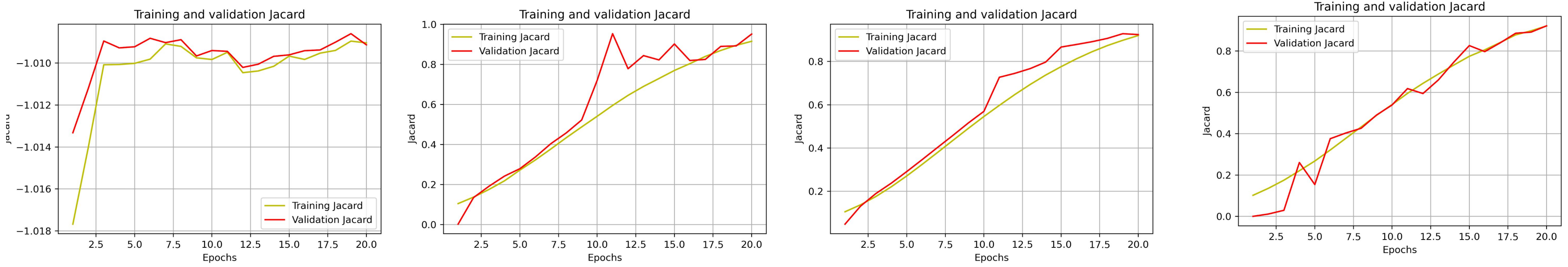
c) Noise subtracted (Blurred -> Thresholded)

d) Mask (Track 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.

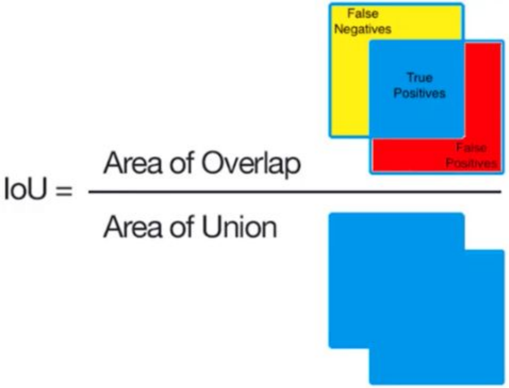
Training and Validation accuracy



(a) CNN (b) UNet (c) Att-UNet (d) Att-ResUNet

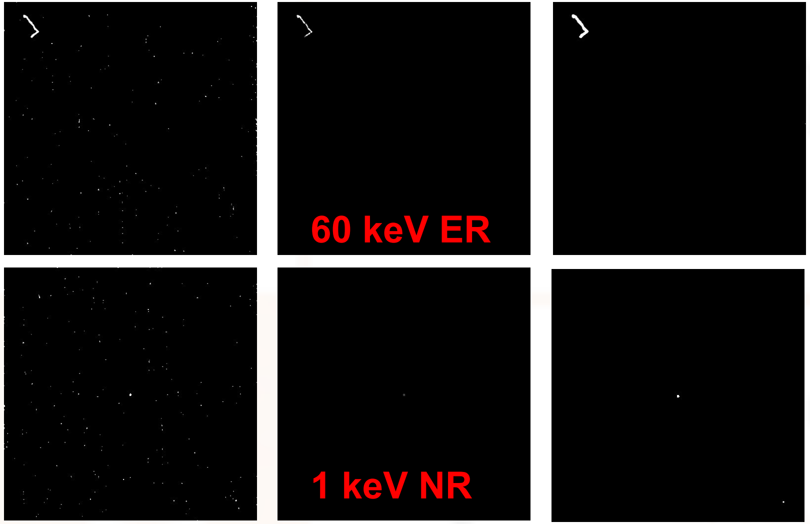
All the models were trained with 600 images and for 20 epochs.

Jaccard Coefficient = Intersection over Union

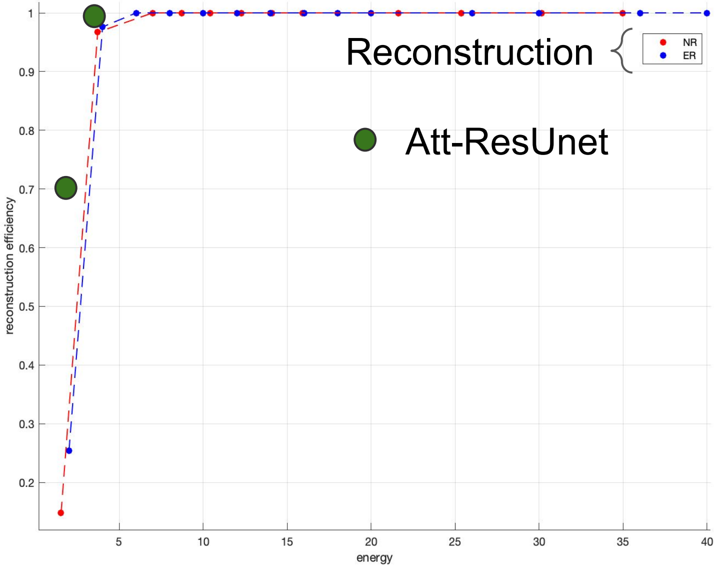


Jaccard Coefficient measures the similarity between between 2 sets of data. The closer to 1 means more similar data.

Prediction from Att-ResUNet



Input Image Target Prediction of the Network



- ❖ 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

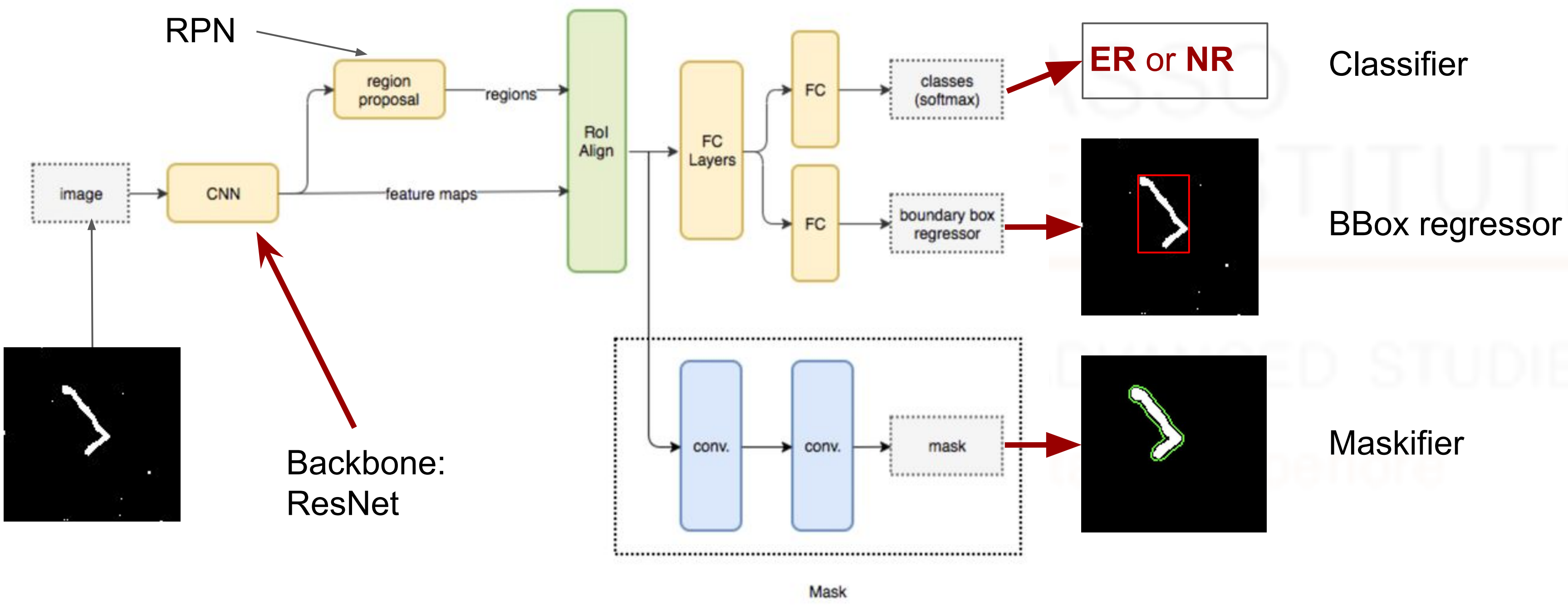
CNN

Mask-RCNN

- Classification
- Semantic Segmentation
- Object Detection
- Instance Segmentation

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

Architecture of Mask-RCNN



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

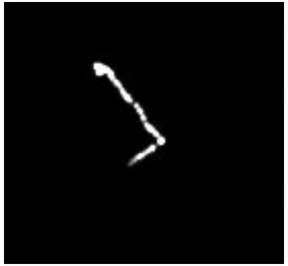
Data for Mask-RCNN

a) Digitized Image (512x512)

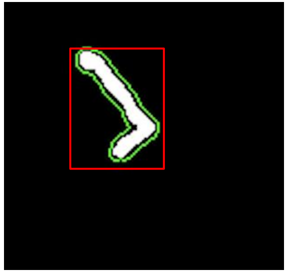
b) Noise (512x512)

c) Noise subtracted (Blurred -> Thresholded)

Input for Network



Digitized image w/o noise

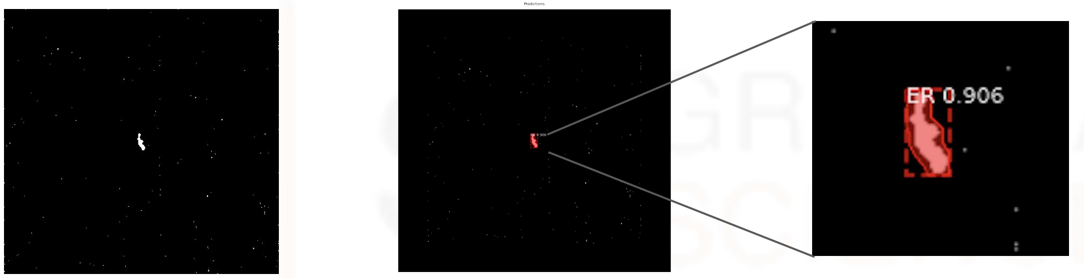


Selecting the cluster using OpenCV

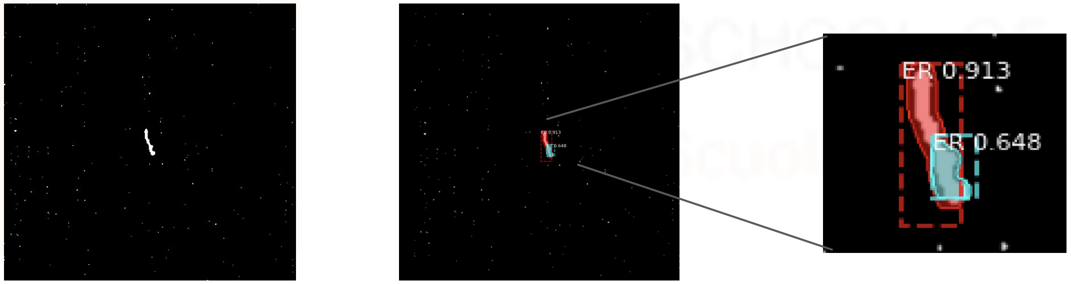
- Shape Attributes
  - Polygon
  - Coordinates of polygon
  - bbox
  - area
- Region Attributes
  - ER or NR (category)
  - Name & index

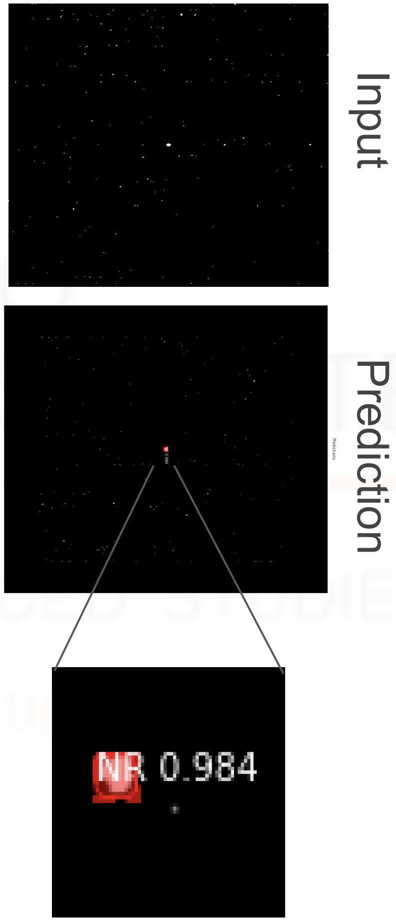
Target for Network

First results from Mask-RCNN



Input Prediction Zoomed Track





Model was trained just with 4 input images and for 20 epochs.

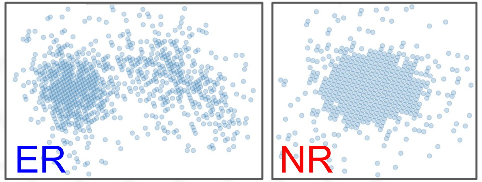
Future Work

- 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
- Test all these models on data
- Will work with the negative ions data from MANGO detector
- We are writing a paper on the results obtained with DNN and Decision Tree based models

Backup

Variables

Observables for recoil identification in gas TPCs
arXiv:2012.13649v1

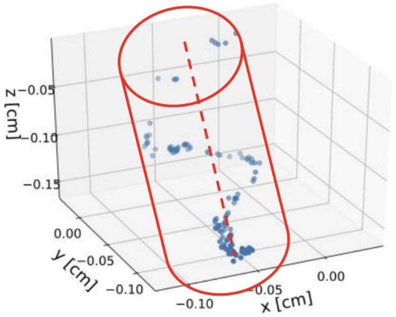


- **Standard Deviation of Charge Distribution 2D(SDCD_2D):**

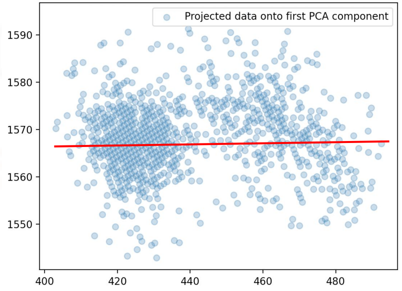
$$SDCD = \sqrt{\frac{\sum_{i=1}^{N}(\mathbf{r_i} - \bar{\mathbf{r}})^2}{N}}.$$

  - 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.



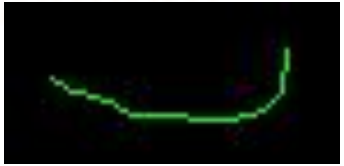
- **Length Along Principal Axis 2D (LAPA_2D):**
  - Project all the points in the charge distribution on to the principal axis.
  - LAPA is the difference between maximum and minimum projected value.
  - ER are longer compared to NRs, therefore projection is also longer.

Variables

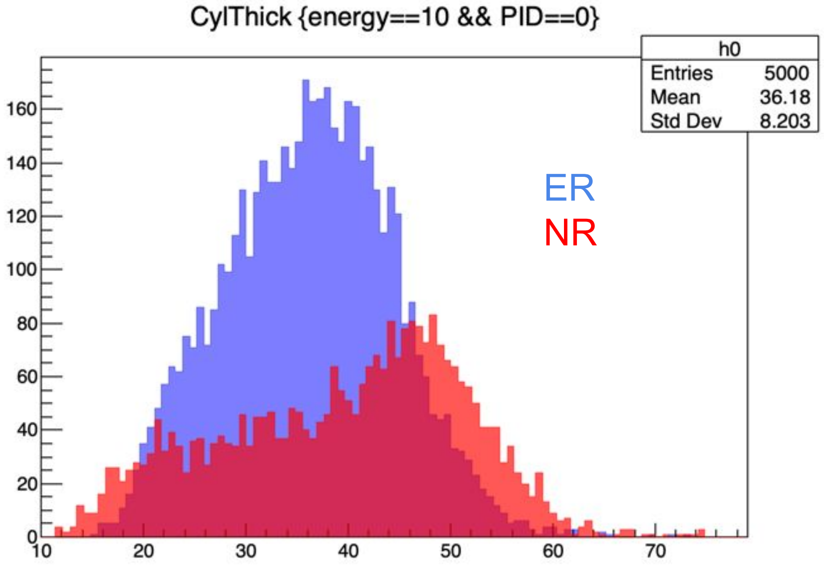
- eta:
  - MaxDen_2D divided by length (found by skeletonization)
- Light Density (delta):
  - Integral of the track divided by number of pixels in the track.
  - NR deposit higher energy over a short distance, therefore Light Density is higher for NR.
- Slimness:
  - Ratio of minor over major axis of 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

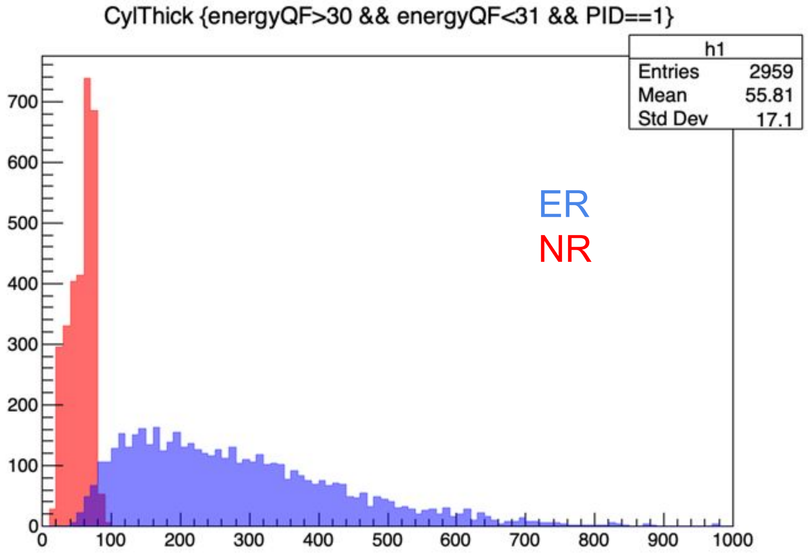
E Baracchini et. al., “Identification of low energy nuclear recoils in a gas TPC with optical readout”, arXiv:2007.12508v1



Variables with decreasing rejection at higher energy

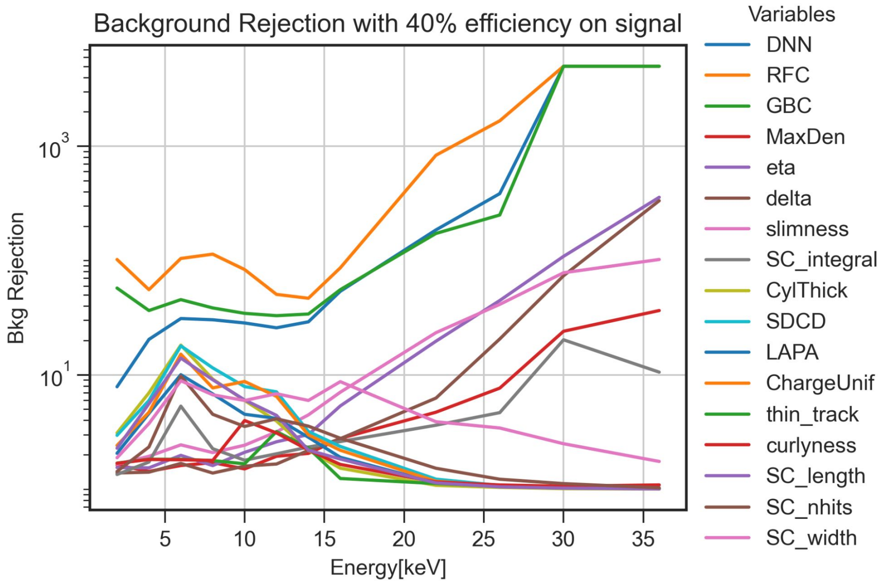


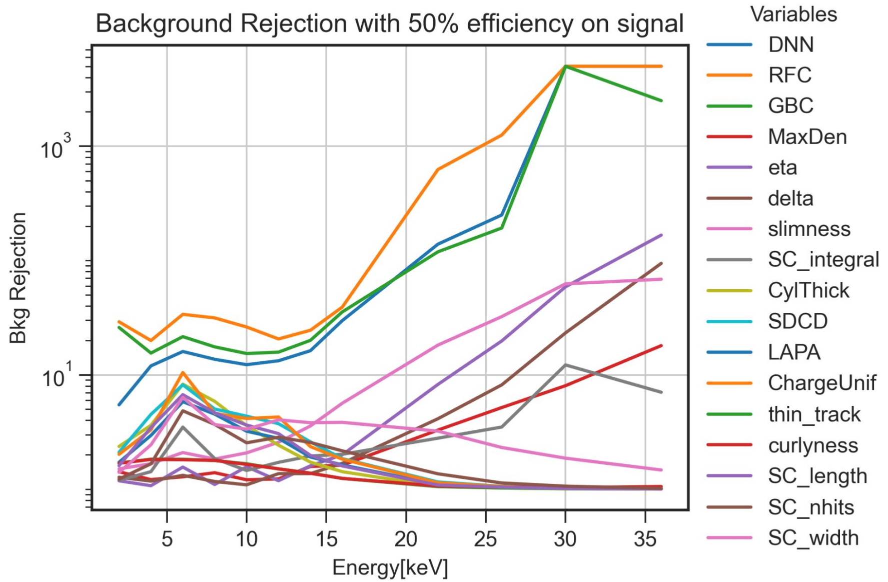
CylThick @ 10 keV



CylThick @ 30 keV

Background Rejection





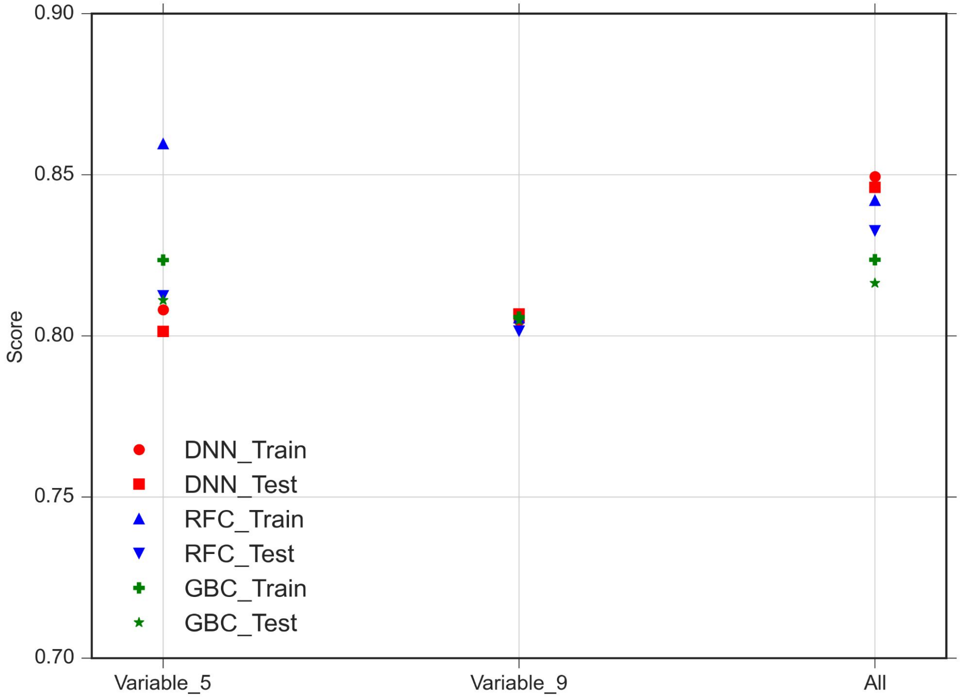
- ❖ Background Rejection is plotted with 40% and 50% signal efficiency in each energy bin.
- ❖ All the variables shown in the plot show the background rejection with classical approach.

Training and Testing Scores for all 3 models

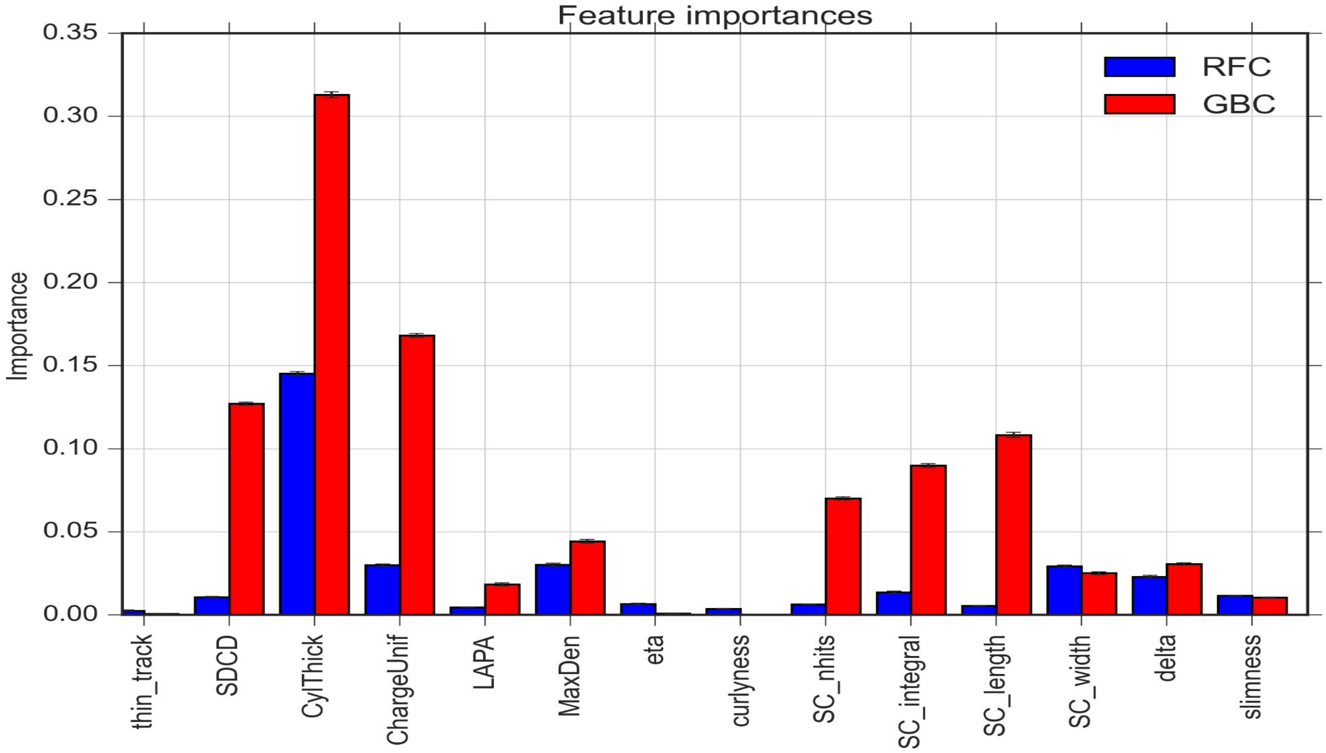
- All the 3 models were trained on all 3 different datasets namely: **All**, **Variable_5**: MaxDen, eta, delta, slimness, integral.

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

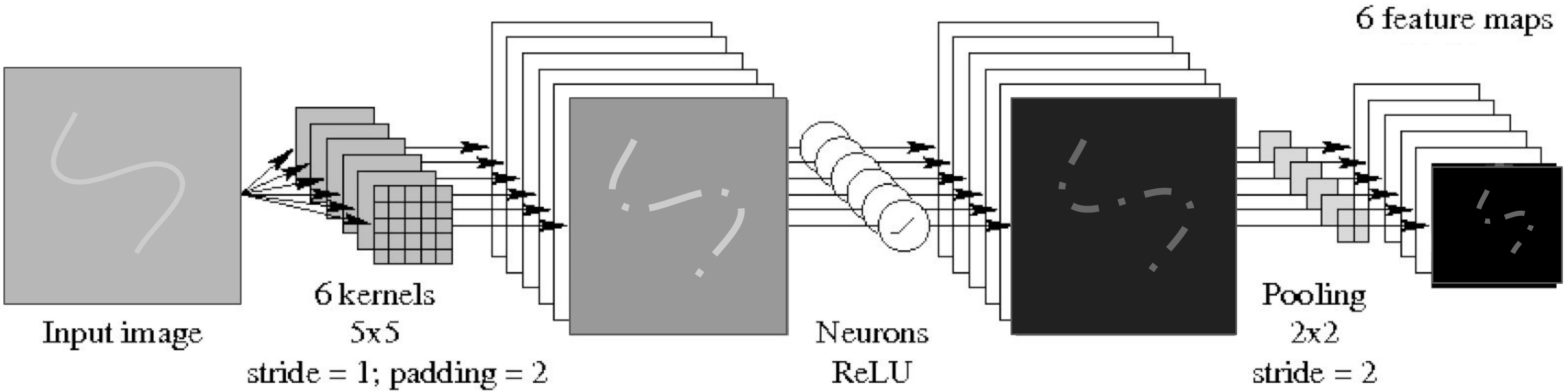
- Score of training and testing of all the models is plotted for the different datasets.



Feature Importance

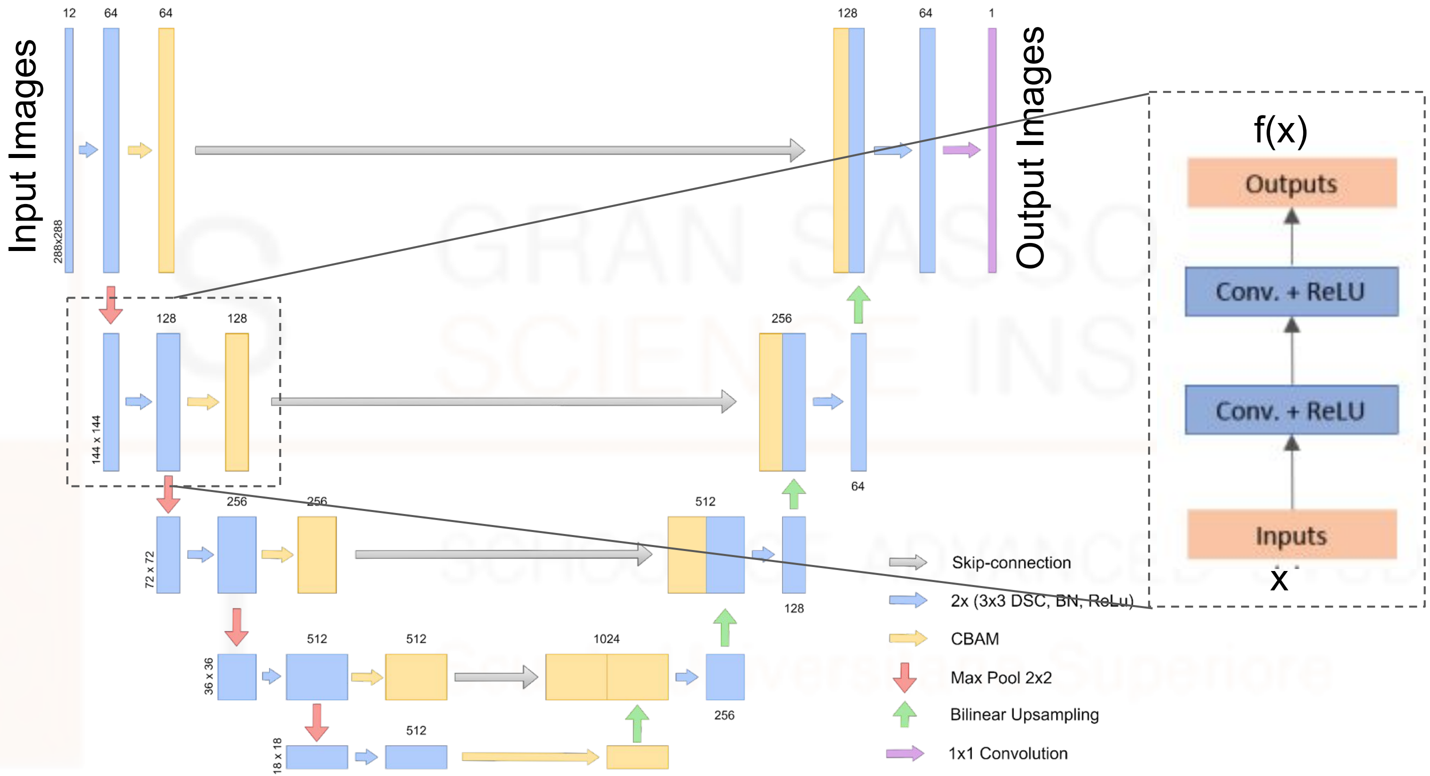


Convolutional Layer



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.

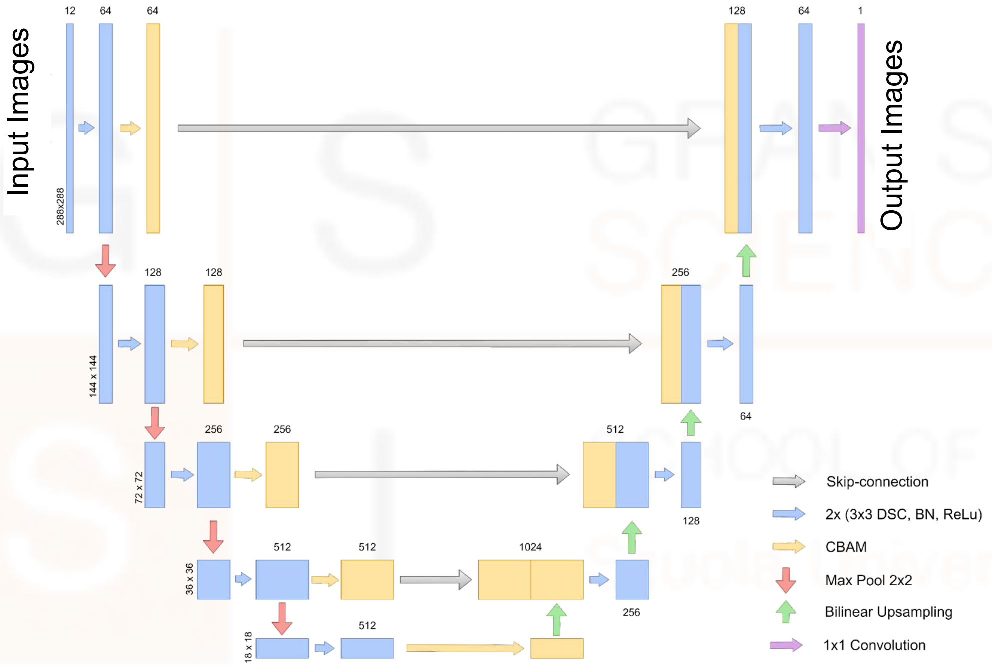
UNet

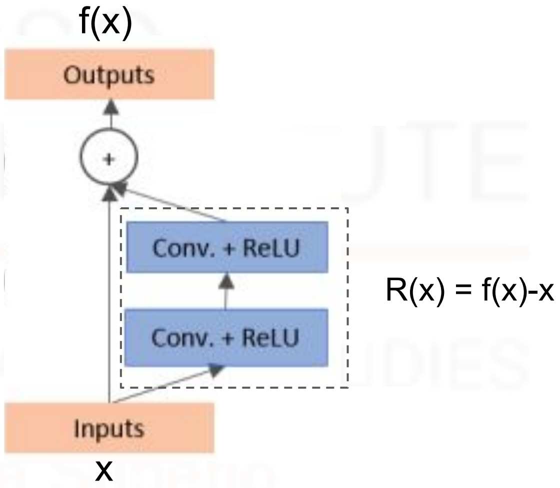


Layers are trying to learn f(x) for the given input x.

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

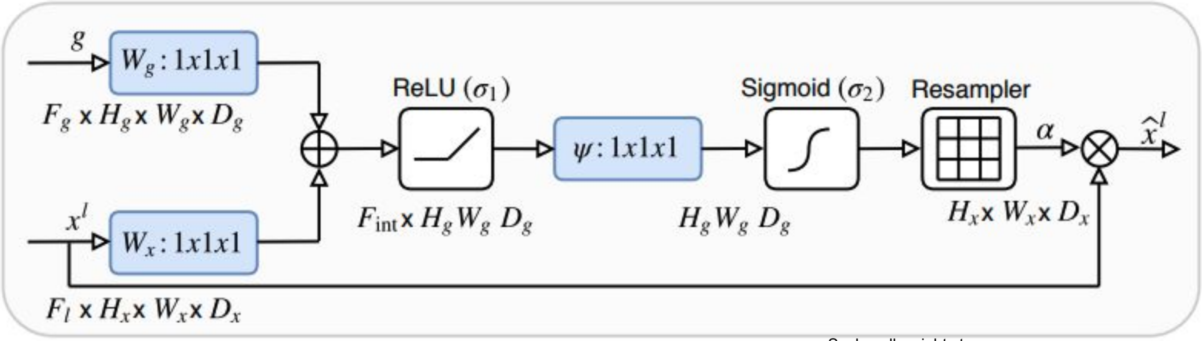
ResUNet

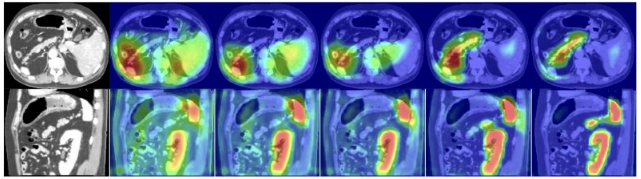




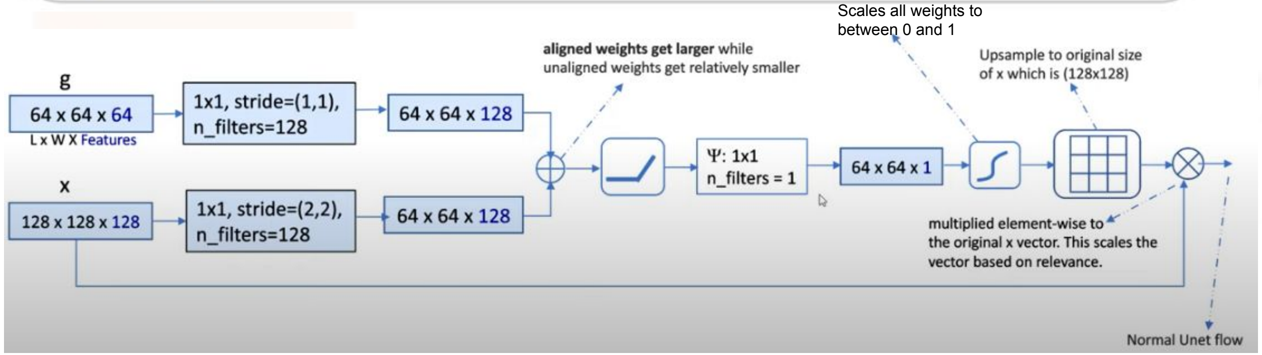
Here, Layers are trying to learn the residual unlike UNet where they try to learn f(x). ResNet helps in solving the problem of vanishing gradients and also of overfitting to an extent.

Attention Block





Attention U-Net: Learning Where to Look for the Pancreas
https://arxiv.org/pdf/1804.03999.pdf



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

Conferences

[1] 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**

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

- **Oral Presentation**

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

- **Presented a poster**

[4] 19th Rencontres du Vietnam, TIMEX- 2023, 5 -11 January 2023, Abstract accepted for an oral presentation.