ER/NR discrimination

Preliminary results with Convolutional Neural Networks

Luan G. M. de Carvalho

*with **Rafael Antunes Nóbrega (UFJF)***

luan.gomes@estudante.ufjf.br

Summary

1. Introduction
2. Datasets
3. Data pre processing
4. Architecture for the model
5. Training the model
6. Results
7. Conclusions

Introduction

From my last presentation...

On going

- Development of algorithms using Feature-based Machine Learning:
  - Non-linear Classifiers: Random Forest, Gradient Boosting, etc
- Development of algorithms using Deep Learning:
  - Convolutional Neural Network
- Evaluate performance and compare the results

Some results already obtained in the past

Currently starting

Future Plans

- Use the new simulated data as dataset for the models
  - New experiment configuration
    - Low gain
    - Quest camera

Introduction

Atul in his thesis proposed the Mask-RCNN architecture

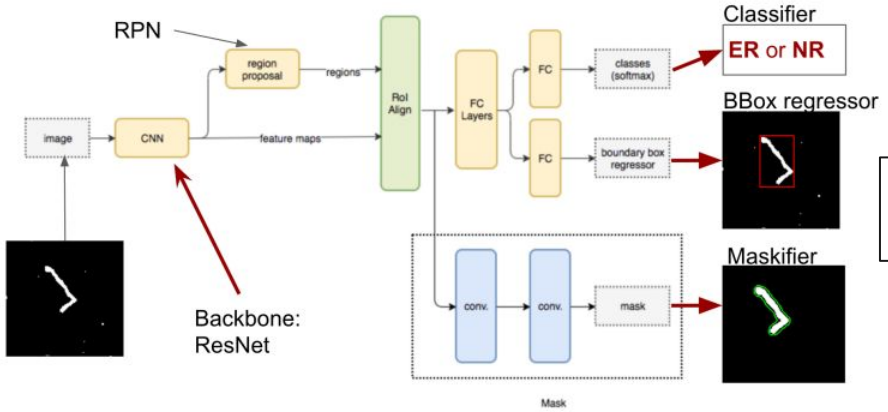


Figure B.1: Architecture of Mask-RCNN model. It consists of several sub-networks like a CNN based backbone network (ResNet here), RPN (region proposal network), classifier, maskifier and bounding box regressor.

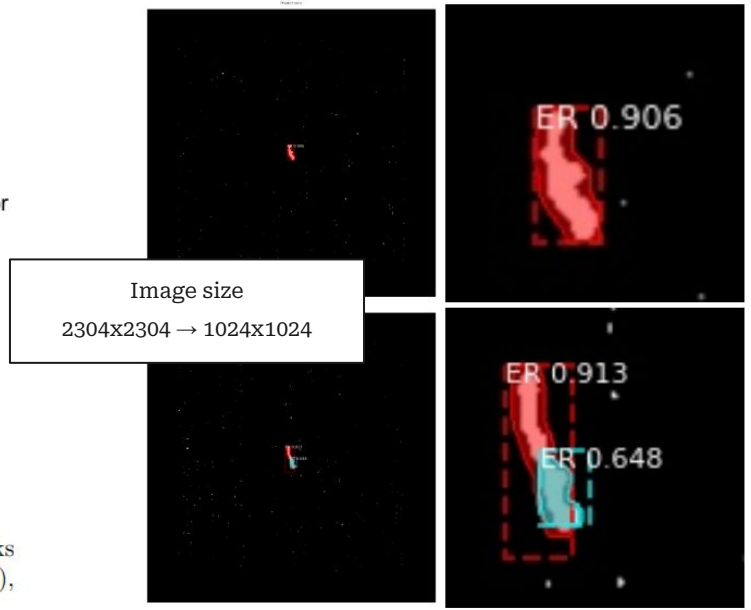


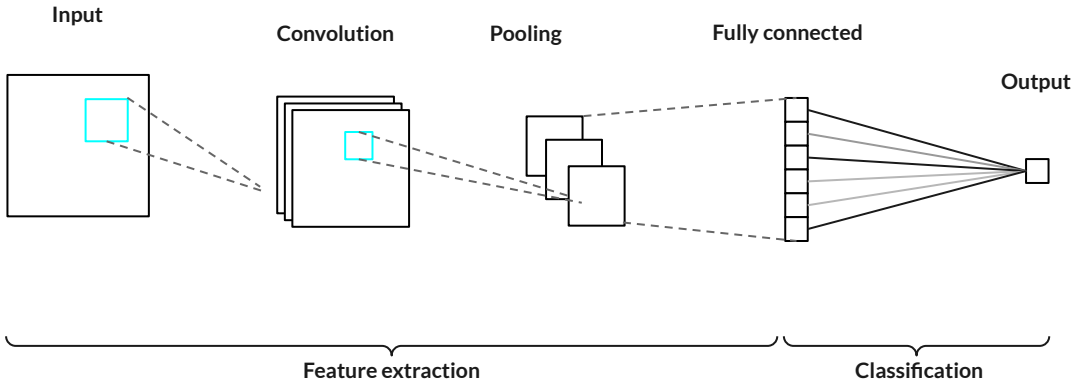
Image size

2304x2304 → 1024x1024

Figure 8.6: Prediction on 60 keV ERs. Plots on the left shows the prediction on the original image of size 1024×1024 pixels, whereas plots on the right show the zoomed in versions of the same tracks.

Introduction

Let’s start with a basic CNN architecture

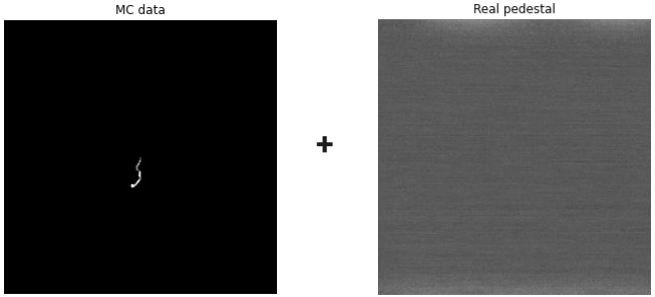


Datasets

Datasets are basically formed by:

MC data no noise + Real pedestal

Example of event



Old simulated data

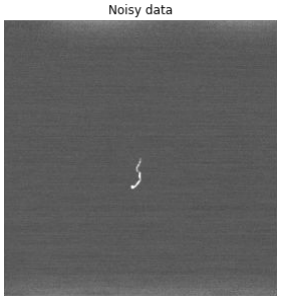
- **ER** 1 keV
- **ER** 3 keV
- **ER** 6 keV
- **ER** 10 keV
- **ER** 30 keV
- **ER** 60 keV
- **NR** 1 keV
- **NR** 3 keV
- **NR** 6 keV
- **NR** 10 keV
- **NR** 30 keV
- **NR** 60 keV

Dataset split

60% training (7200)

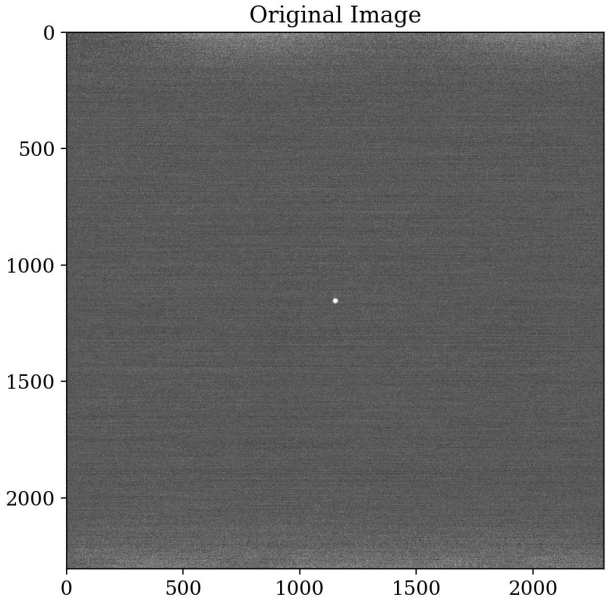
20% validation (2400)

20% test (2400)



Data pre processing

Image size

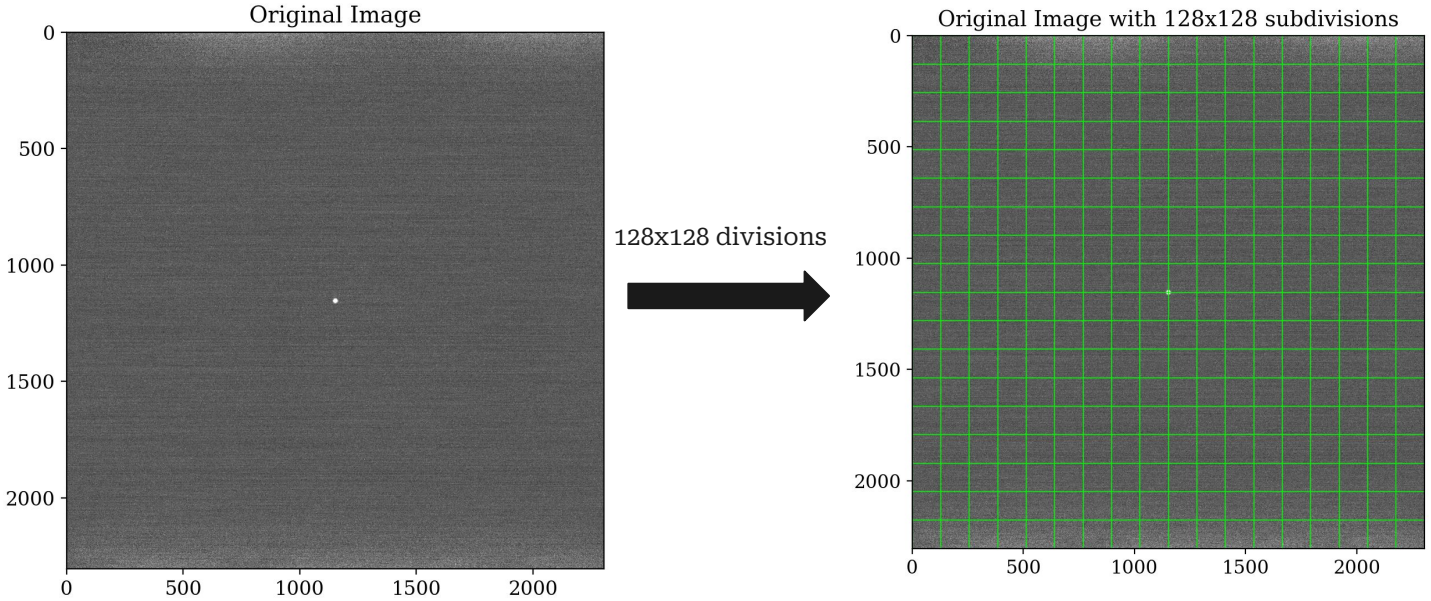


2304x2304

↳ unfeasible for CNN input

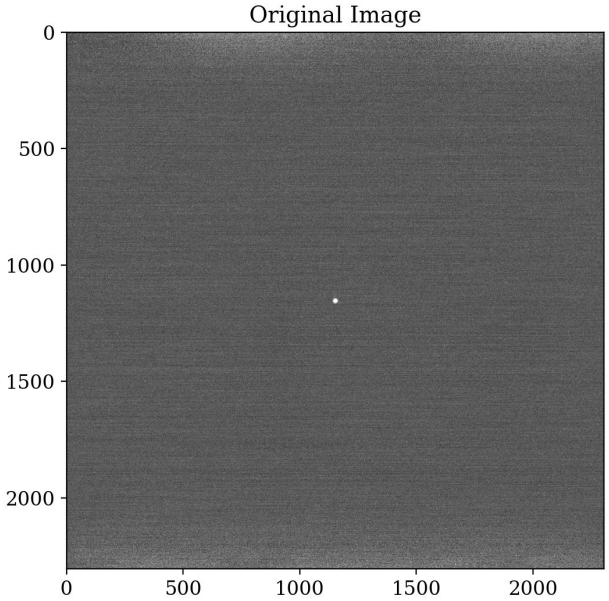
Data pre processing

Image size

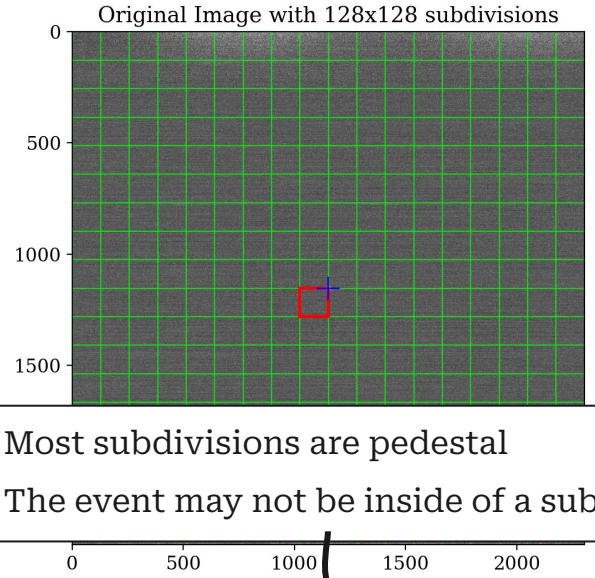


Data pre processing

Image size



128x128 divisions

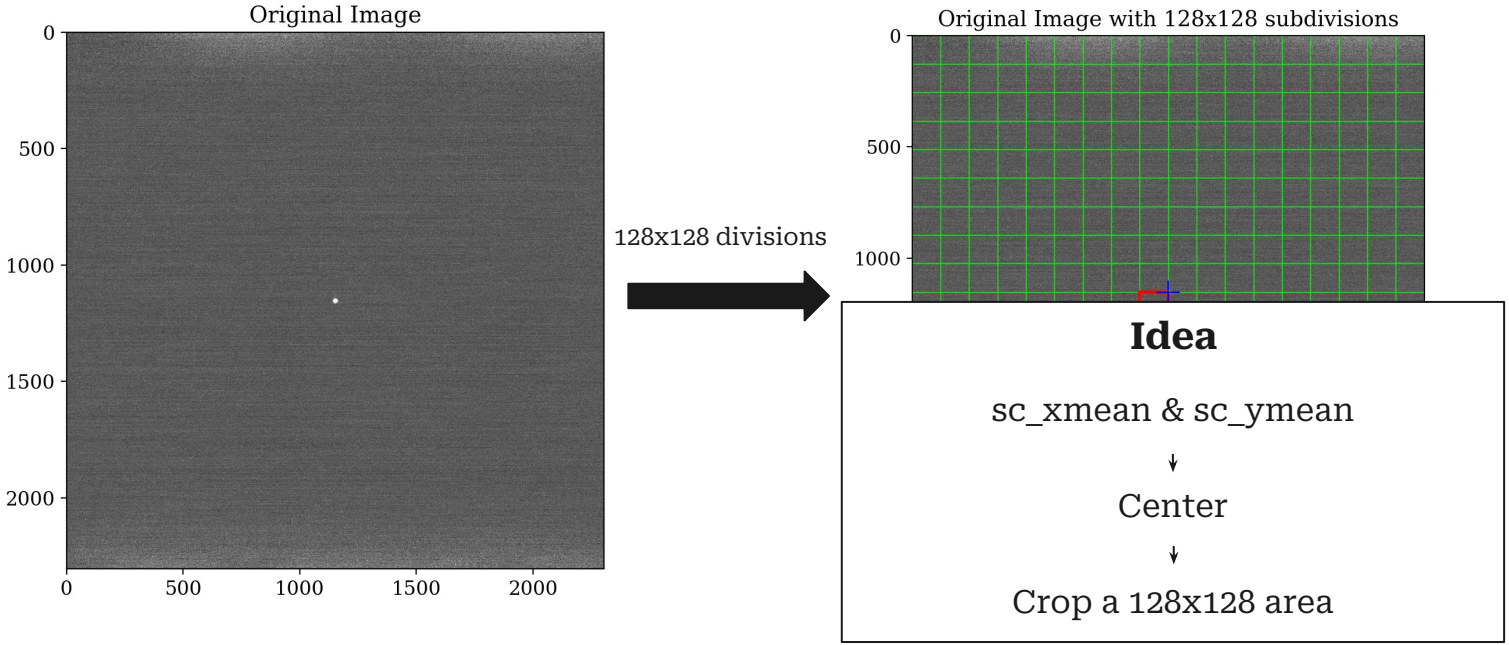


- Most subdivisions are pedestal
- The event may not be inside of a subdivision

Overlap subdivisions

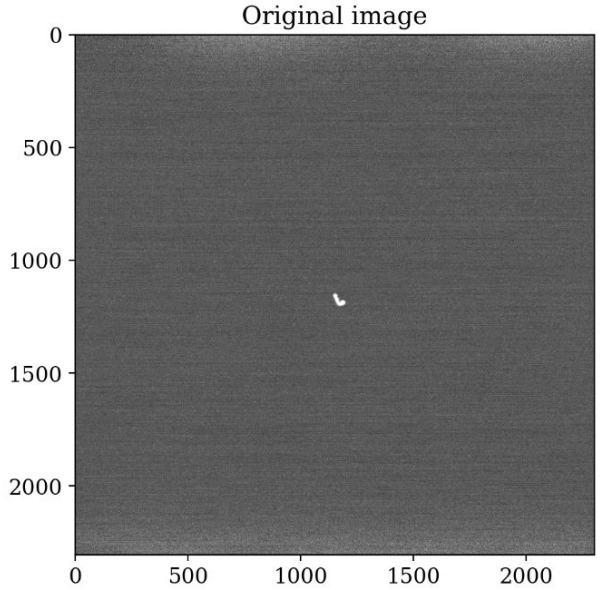
Data pre processing

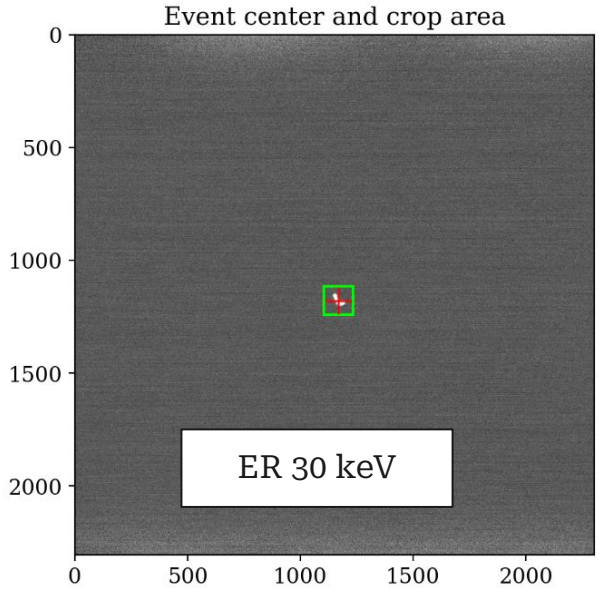
Image size

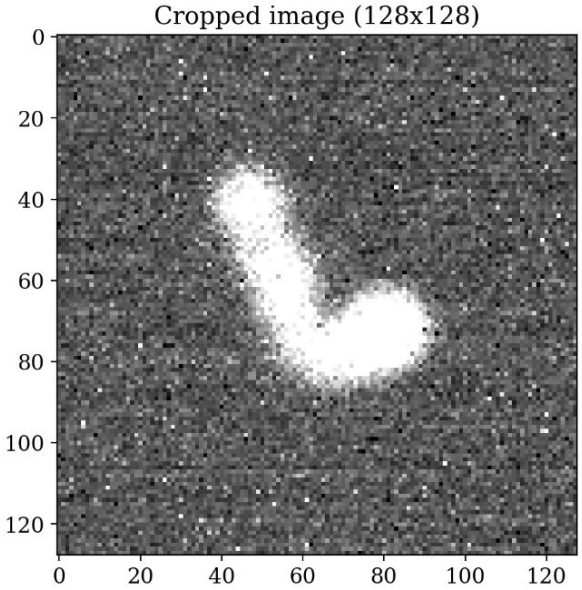


Data pre processing

Image size

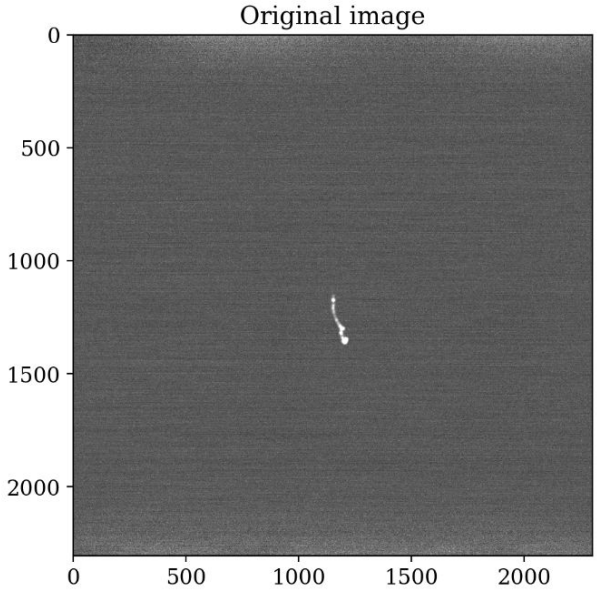


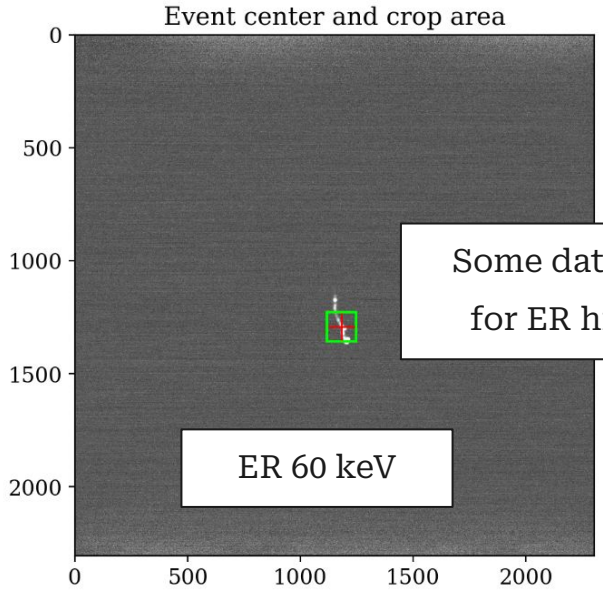




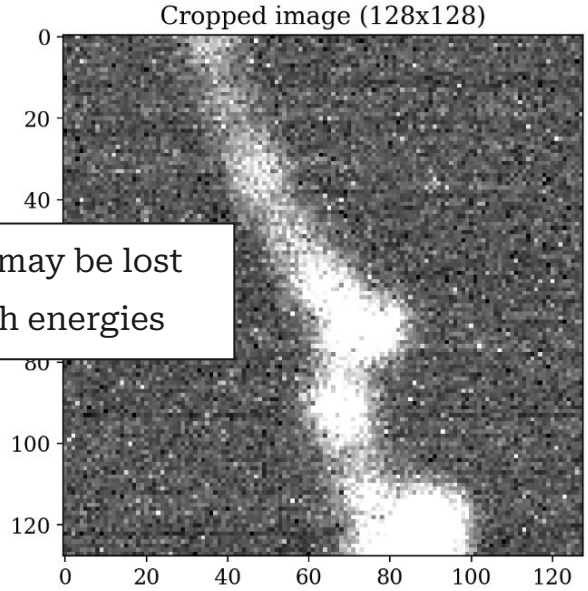
Data pre processing

Image size



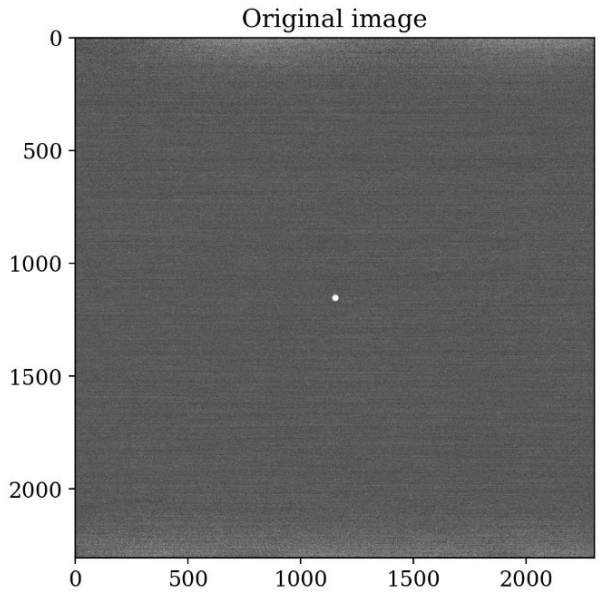


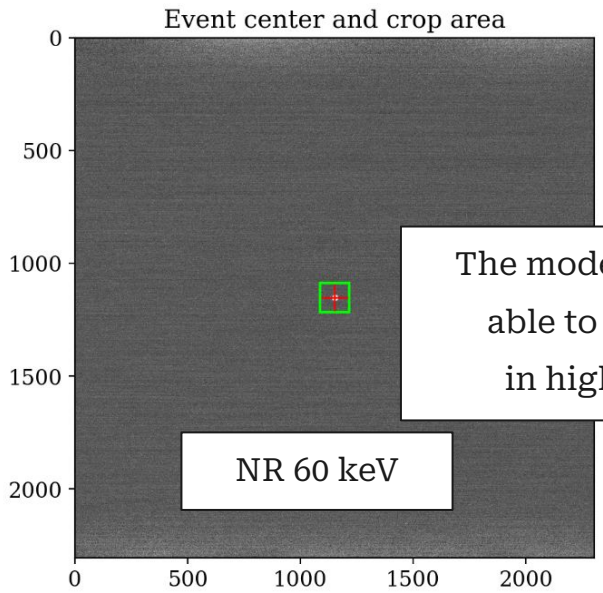
Some data may be lost
for ER high energies

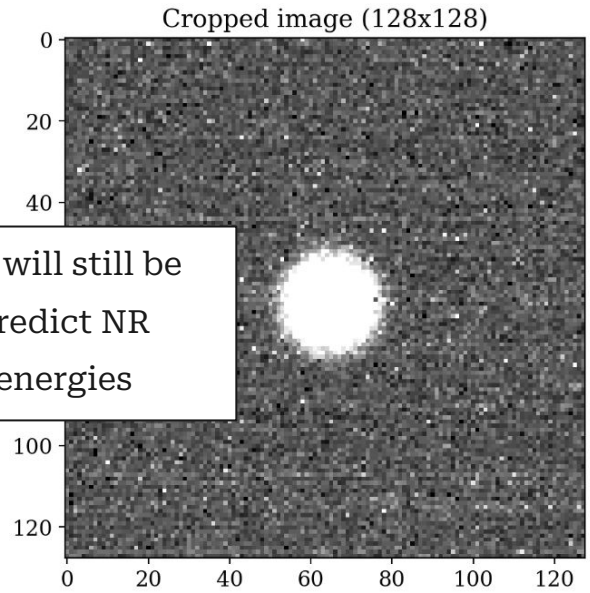


Data pre processing

Image size

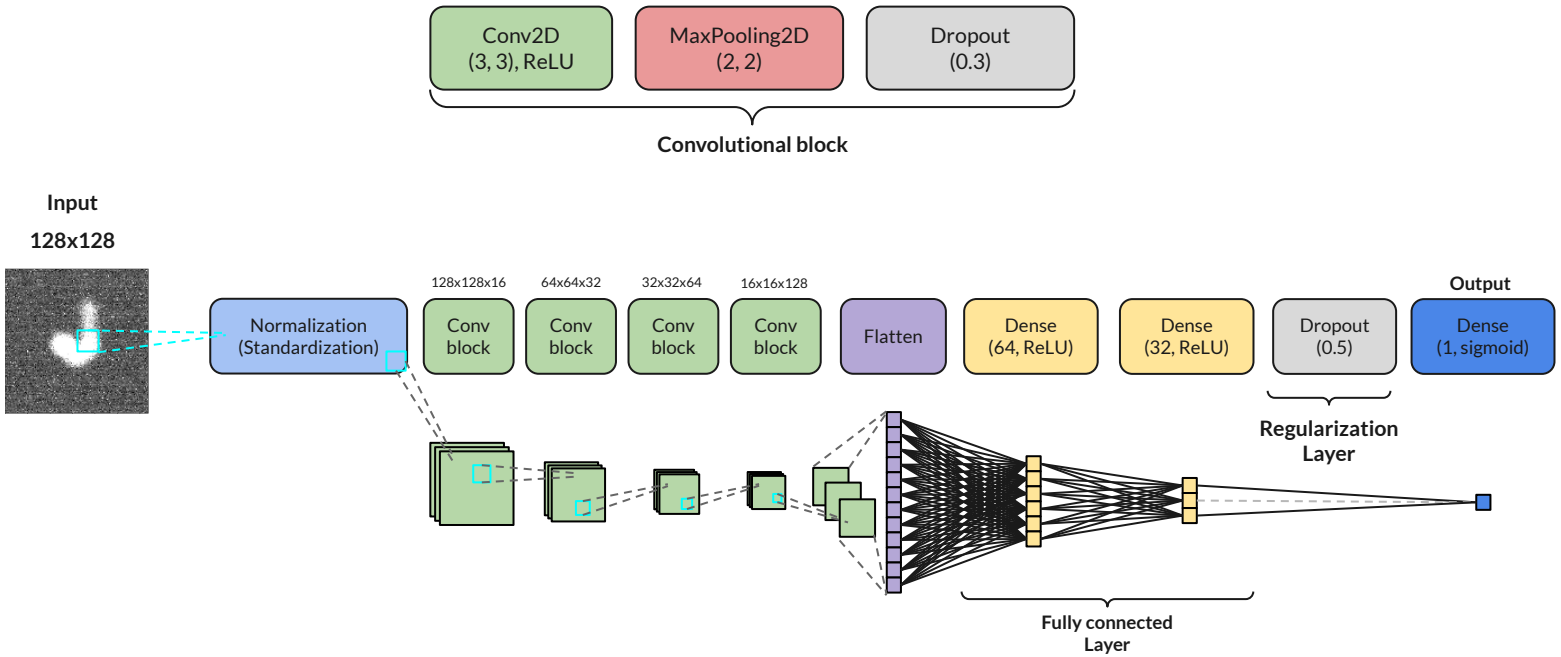






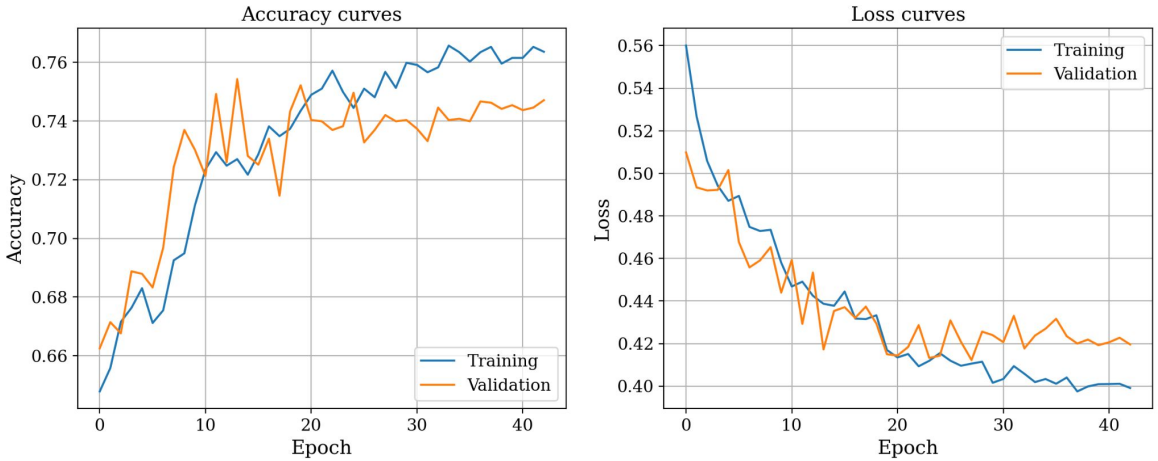
The model will still be able to predict NR in high energies

Architecture for the model



Training the model

Training curves



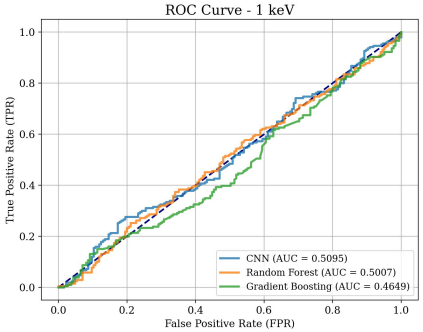

```
--- Test dataset evaluation ---
75/75 [==============================] - 4s 46ms/step -
loss: 0.4070 - accuracy: 0.7646 - auc: 0.8829
```

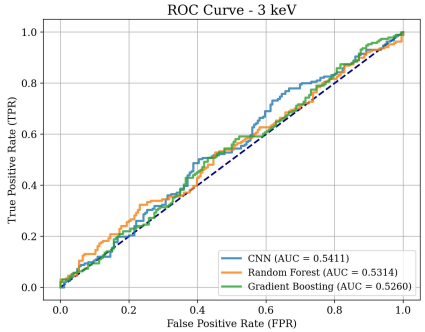
```python
early_stopping = callbacks.EarlyStopping(
    monitor='val_loss',
    patience=15,
    restore_best_weights=True,
    verbose=1
)
```

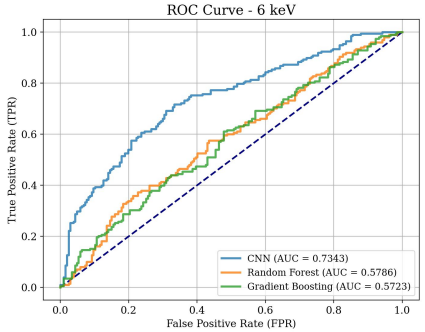
Learning rate: 0.001

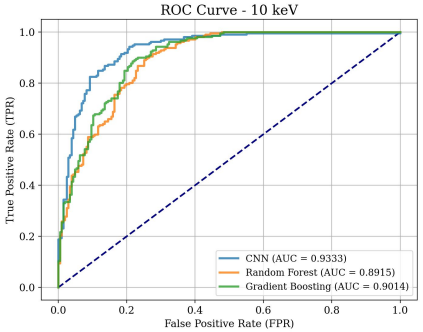
Results

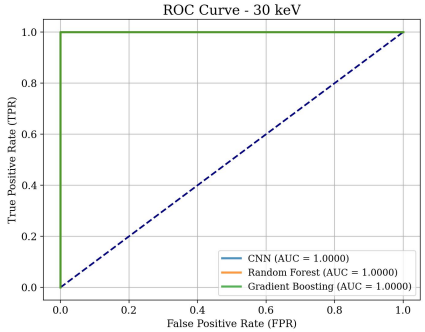
ROC Curves for the test dataset

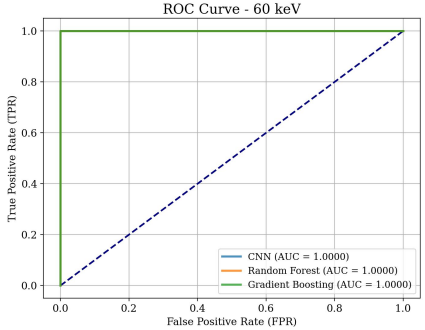












Conclusions

Current approach

- Strategy:
  - sc_xmean and sc_ymean -> centered cluster in the 128x128 cropped image -> CNN Input
- Results:
  - Performance better than Random Forest and Gradient Boosting classifiers (6 - 60 keV)
  - The CNN demonstrated potential for discrimination
- Challenges and limitations:
  - Currently is a binary classification problem: ER vs NR
  - Noise clusters are identified by the reco. algorithm -> Train the model with pedestal subimages
    - Expand to a multiclass problem
    - Binary problem: NR vs background (ER + noise)

} Test and evaluate

Conclusions

Next steps

- Start using the new simulated data
  - Test configurations:
    - Low gain
    - High gain
    - Quest/Fusion

  } Evaluate performance

- Improve the current pipeline
  - Try applying pedestal subtraction to the subimages
- Explore advanced architectures
  - Mask-RCNN

Thank you for your attention!

Questions, suggestions?