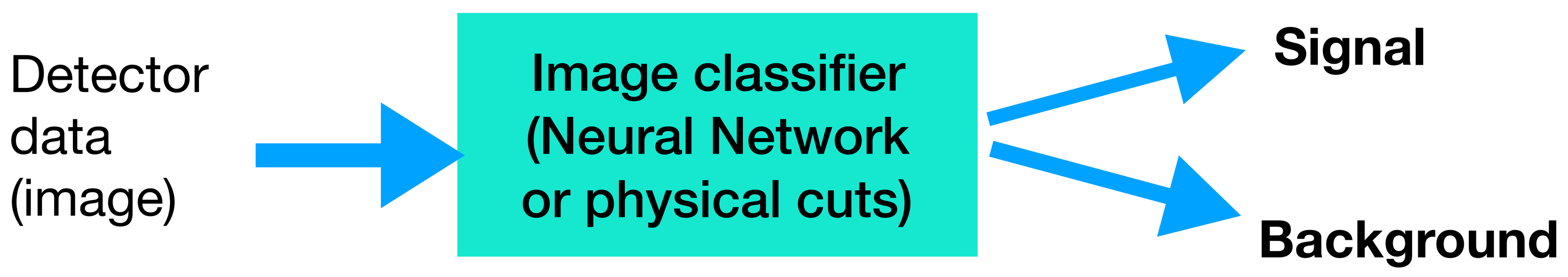
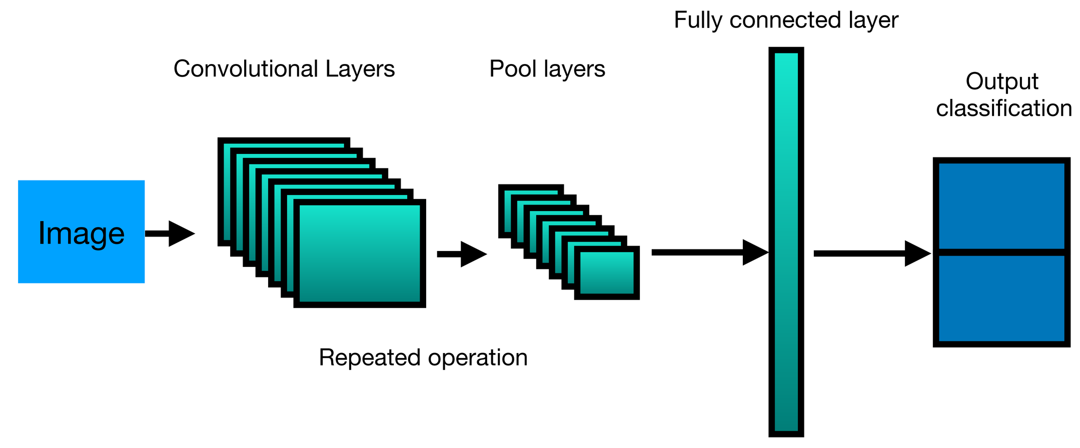
# Use of CNNs for signal-background classification in High Energy Physics

Venkitesh Ayyar[1], W. Bhimji[1], N. Choma[1], Lisa Gerhardt[1], Sally Robertson[1,2], Zahra Ronaghi[3]
[1]Lawrence Berkeley National Lab, [2]University of California, Berkeley, [3]NVIDIA

## Signal-background classification in HEP

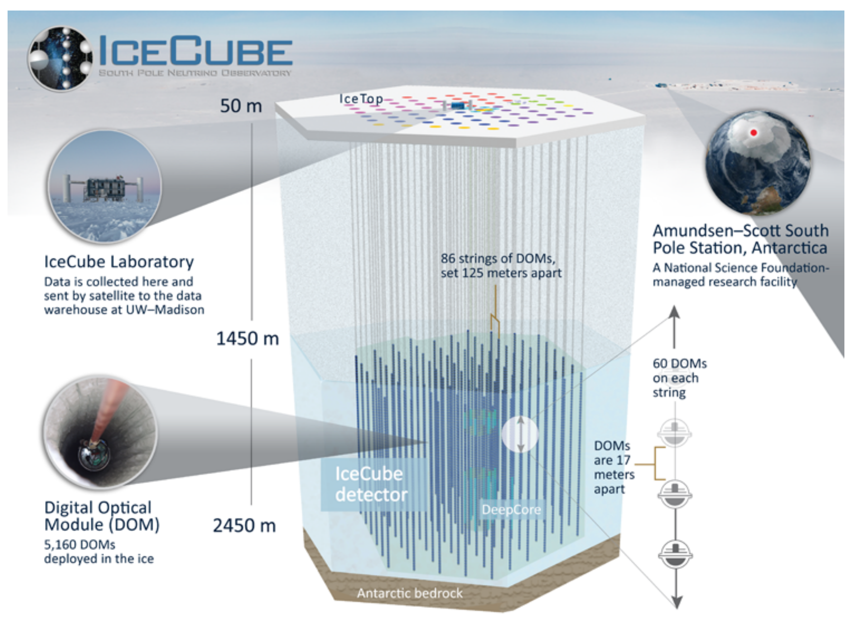



- Currently done using cuts motivated by physics
- Convolutional Neural Networks (CNNs) have proved successful for image classification
- CNNs could provide a quicker and more efficient way for signal-background classification

We describe our study using data from two experiments: Ice Cube and ATLAS
Expanding on previous work, we demonstrate the success of compact CNNs in identifying signal from background.
This work is not part of either collaboration

## Ice Cube experiment




Neutrino observatory based in the South Pole

Searching for high energy neutrinos

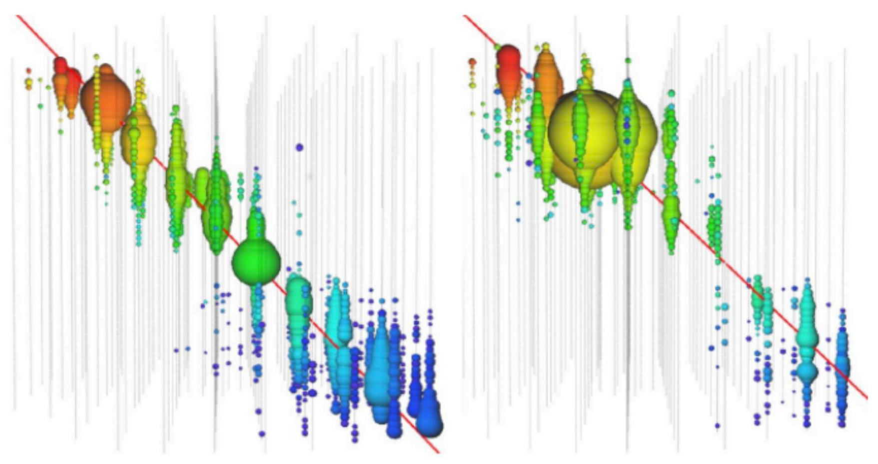
Detects charged particles by observing Cherenkov radiation

### Dataset[2]

Simulated data with astrophysical muon neutrino signal and cosmic ray shower background

NN input included charge and timing from each optical sensor

### Light deposition patterns for Background and Signal



**Background**: Muon bundles (multiple muons lose energy evenly)

**Signal**: Single high-energy muon (stochastic light emission along track)

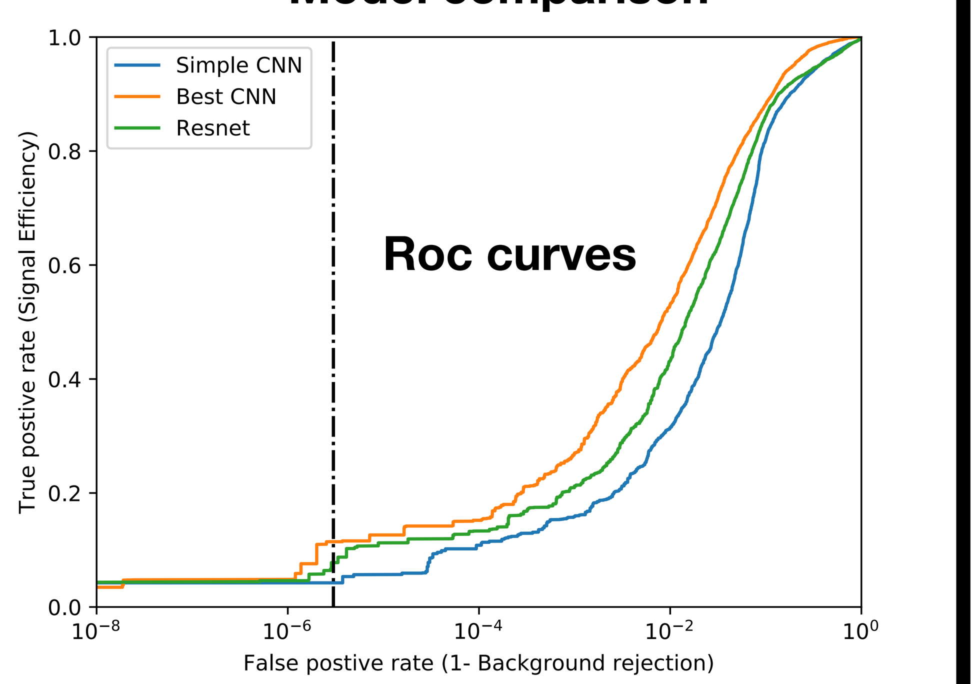
### Prior work[1]

Use of 3D GNNs and CNNs (resnet) showed improved performance compared to Physics cuts

### This work

- Broader architecture exploration of layered CNNs
- Filtered out High Energy Starting Events from signal

### Model comparison

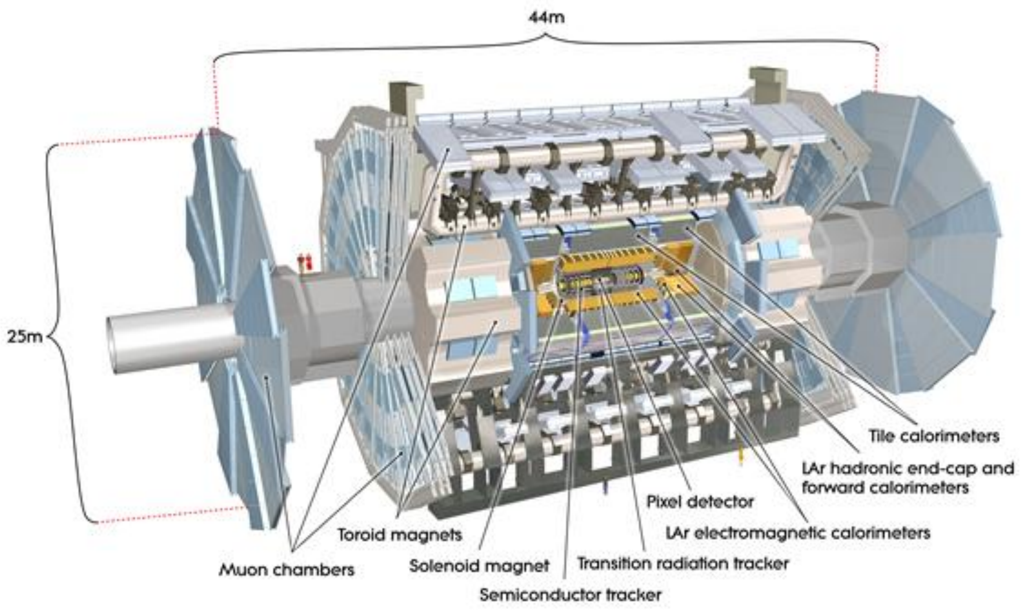



### Result

- Number of model parameters reduced by 1 order of magnitude
- Best CNN model was layered CNN with 5 convolutional layers and 2.1 million parameters

| Model type | Number of parameters | Training time per epoch on GPU (in seconds) | AUC score | True positive rate at fpr=3e-6 |
|---|---|---|---|---|
| Simple CNN | 14,789 | 23 | 0.922 | 0.04 |
| Best CNN | 2,116,881 | 120 | 0.960 | 0.116 |
| Resnet | 26,486,126 | 115 | 0.935 | 0.078 |

## ATLAS experiment




One of the major experiments at the LHC in CERN

Testing Standard Model of Physics and looking for Physics beyond the Standard Model

### Dataset[4]

Search for new massive SUSY particles in multi-jet final states, focussing on gluino-cascade decays
Simulated data with RPV-SUSY signal with QCD background

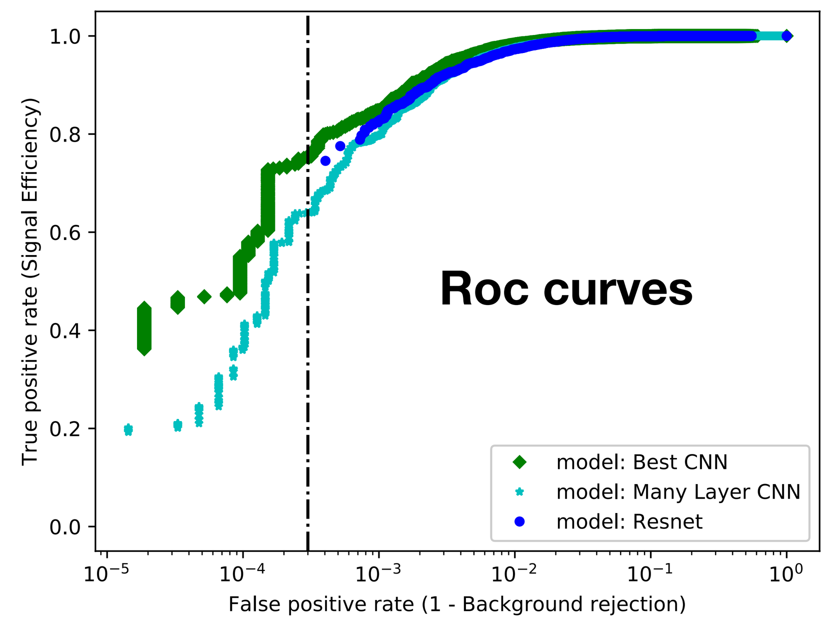
### Prior work[3]

Use of layered compact 2D CNNs to demonstrate improved performance over physics cuts

### This work

More thorough exploration of architecture using compact layered CNNs and Resnets

### Model comparison




| Model type | Number of parameters | Training time per epoch on GPU (in seconds) | True positive rate at fpr=3e-4 |
|---|---|---|---|
| Many Layer CNN | 34,515,201 | 294 | 0.641 |
| Resnet | 23,597,826 | 515 | - |
| Best CNN | 43,009 | 40 | 0.746 |

### Result

- Number of model parameters reduced by 2 orders of magnitude
- Best model was a CNN with around 43k parameters

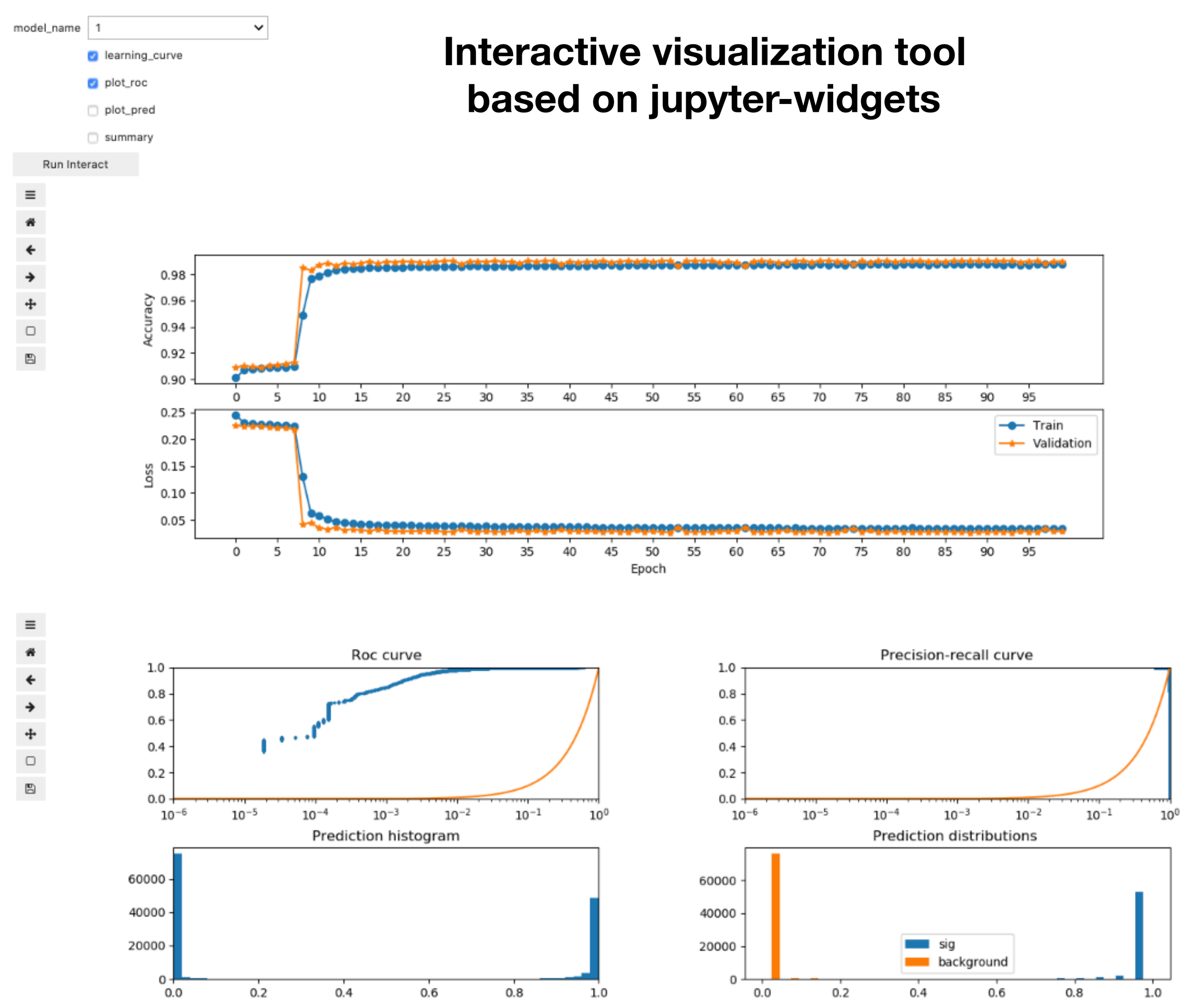
Developed a package for
- Training CNNs
- Visualizing CNN performance

https://github.com/vpayyar/layered_CNNs_for_ATLAS_data

### Best model

```
Model: "model"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
input_1 (InputLayer)         [(None, 64, 64, 1)]       0
_________________________________________________________________
conv2d (Conv2D)              (None, 64, 64, 10)        100
_________________________________________________________________
max_pooling2d (MaxPooling2D) (None, 32, 32, 10)        0
_________________________________________________________________
dropout (Dropout)            (None, 32, 32, 10)        0
_________________________________________________________________
conv2d_1 (Conv2D)            (None, 32, 32, 10)        910
_________________________________________________________________
max_pooling2d_1 (MaxPooling2 (None, 16, 16, 10)        0
_________________________________________________________________
dropout_1 (Dropout)          (None, 16, 16, 10)        0
_________________________________________________________________
conv2d_2 (Conv2D)            (None, 16, 16, 10)        910
_________________________________________________________________
max_pooling2d_2 (MaxPooling2 (None, 8, 8, 10)          0
_________________________________________________________________
dropout_2 (Dropout)          (None, 8, 8, 10)          0
_________________________________________________________________
flatten (Flatten)            (None, 640)               0
_________________________________________________________________
dropout_3 (Dropout)          (None, 640)               0
_________________________________________________________________
dense (Dense)                (None, 64)                41024
_________________________________________________________________
dense_1 (Dense)              (None, 1)                 65
=================================================================
Total params: 43,009
Trainable params: 43,009
Non-trainable params: 0
_________________________________________________________________
```

### Interactive visualization tool based on jupyter-widgets




## Summary

- Studied the use of compact layered CNNs to classify signal and background
- Model architecture search finds models that can achieve similar performance with substantially more compact network
- Developed packages for general purpose CNN training and visualization
- Potential for CNNs to replace physics cuts in data filtering

## References:

1. *Graph Neural Networks for IceCube Signal Classification*, Choma et al, arxiv: 1809.06166
2. D. Heck, G. Schatz, J. Knapp, T. Thouw, and J. Capdevielle, “COR- SIKA: A Monte Carlo code to simulate extensive air showers,” Tech. Rep. FZKA-6019, 1998
3. Deep Neural Networks for Physics Analysis on low-level whole-detector data at the LHC, Bhimji et al, arxiv: 1711.03573
4. *ATLAS Collaboration 2016 Search for massive supersymmetric particles in multi-jet final states produced in pp collisions at s = 13 TeV using the ATLAS detector at the LHC Tech. Rep. ATLAS-CONF-2016-057 CERN Geneva URL http://cds.cern.ch/record/2206149*