Deep Learning for neutrino telescopes

Stefan Geißelsöder
21. Feb. 2017
Big Data Science in Astroparticle Physics
HAP Workshop, RWTH Aachen

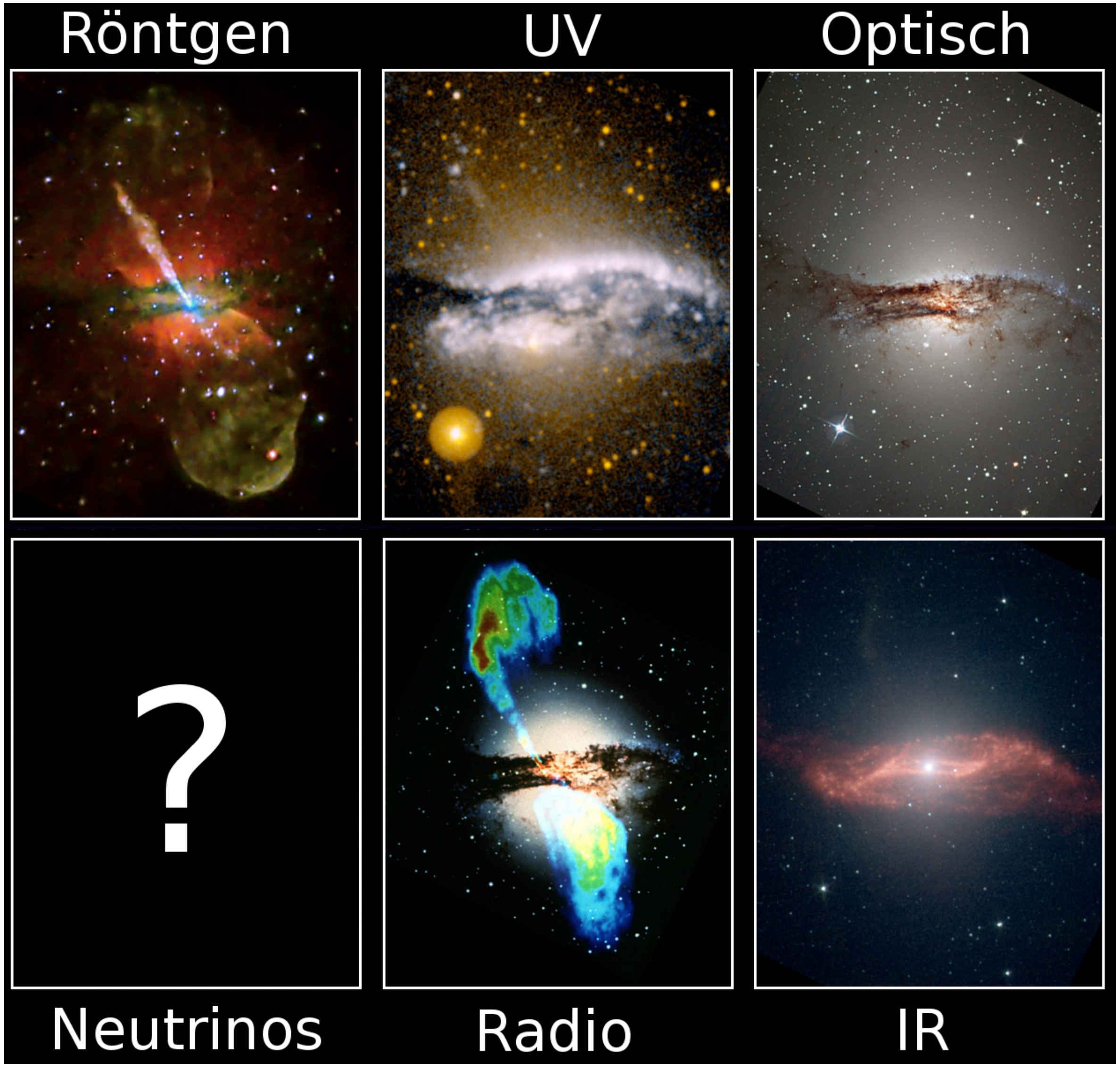
KM3NeT

ERLANGEN CENTRE FOR ASTROPARTICLE PHYSICS

FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG
NATURWISSENSCHAFTLICHE FAKULTÄT

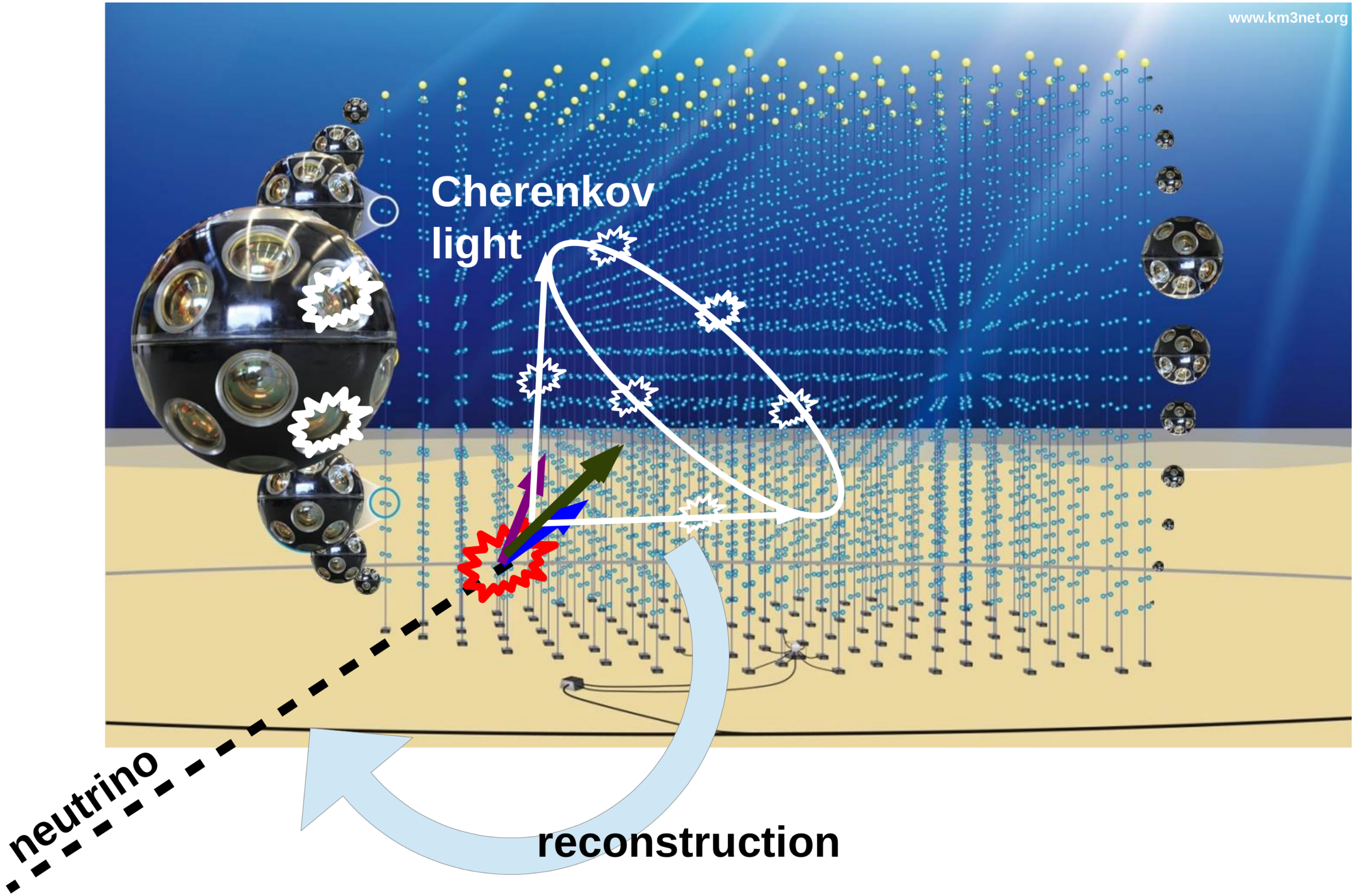
Motivation neutrino astronomy

- Learn about high-energy cosmic rays
- Find sources of high-energy neutrino flux
- Do multi-messenger astronomy

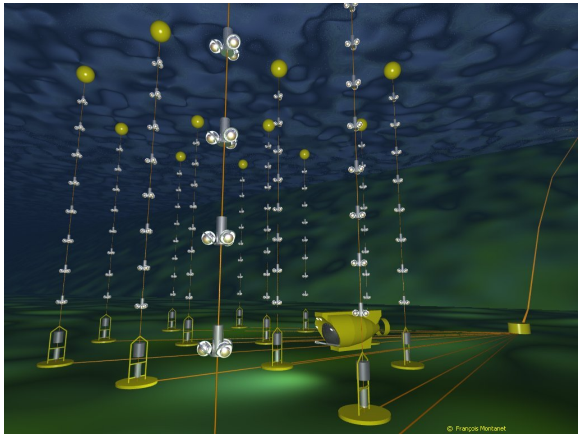


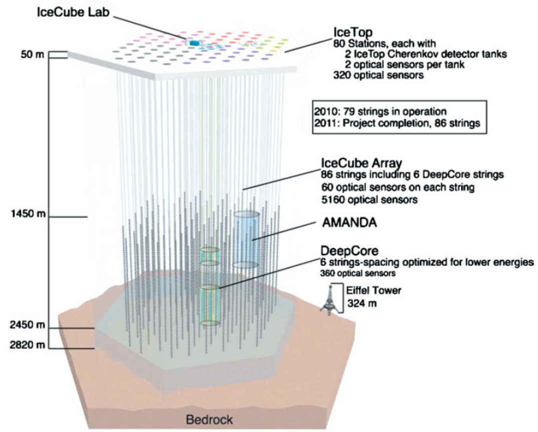
(Centaurus A)

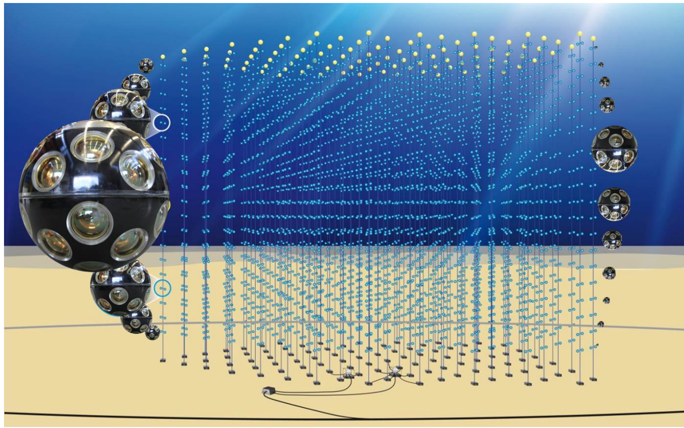
Detection principle



ECAP involved in





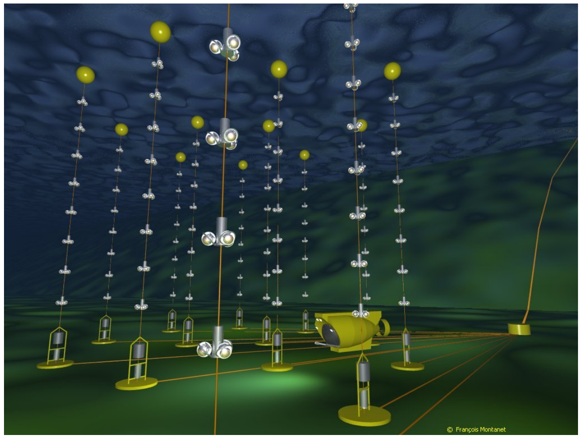


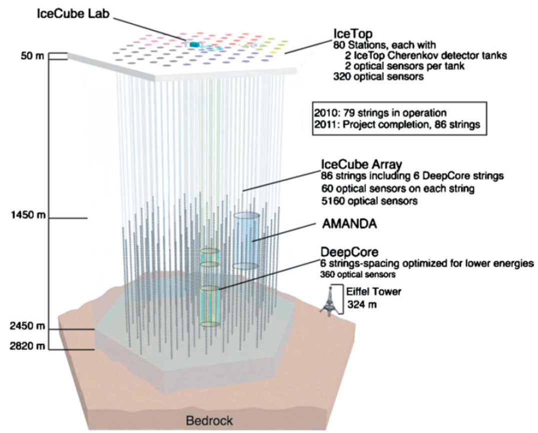
ANTARES
- Operational
- Mediterranean Sea
- 885 OMs
- 1 PMT/OM

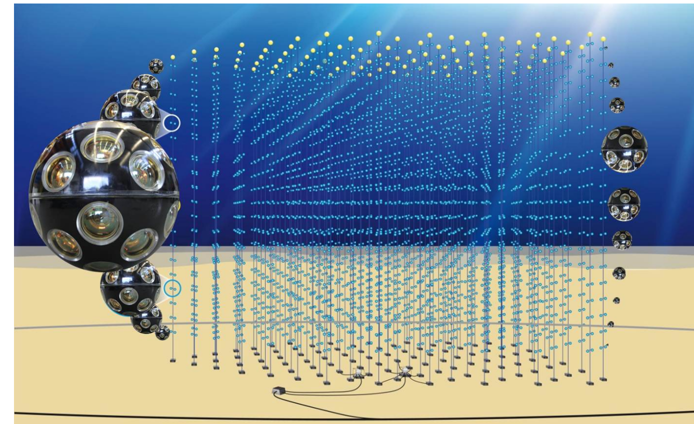
IceCube
- Operational
- Antarctic Ice
- 5160 DOMs
- 1 PMT/DOM

KM3NeT
- Being build
- Mediterranean Sea
- 2070 DOMs/block (up to six blocks)
- 31 PMTs/DOM

ECAP involved in







ANTARES

- Operational
- Mediterranean Sea
- 885 OMs
- 1 PMT/OM

IceCube

- Operational
- Antarctic Ice
- 5160 DOMs
- 1 PMT/DOM

KM3NeT

- Being build
- Mediterranean Sea
- 2070 DOMs/block (up to six blocks)
- 31 PMTs/DOM

Main focus of this talk

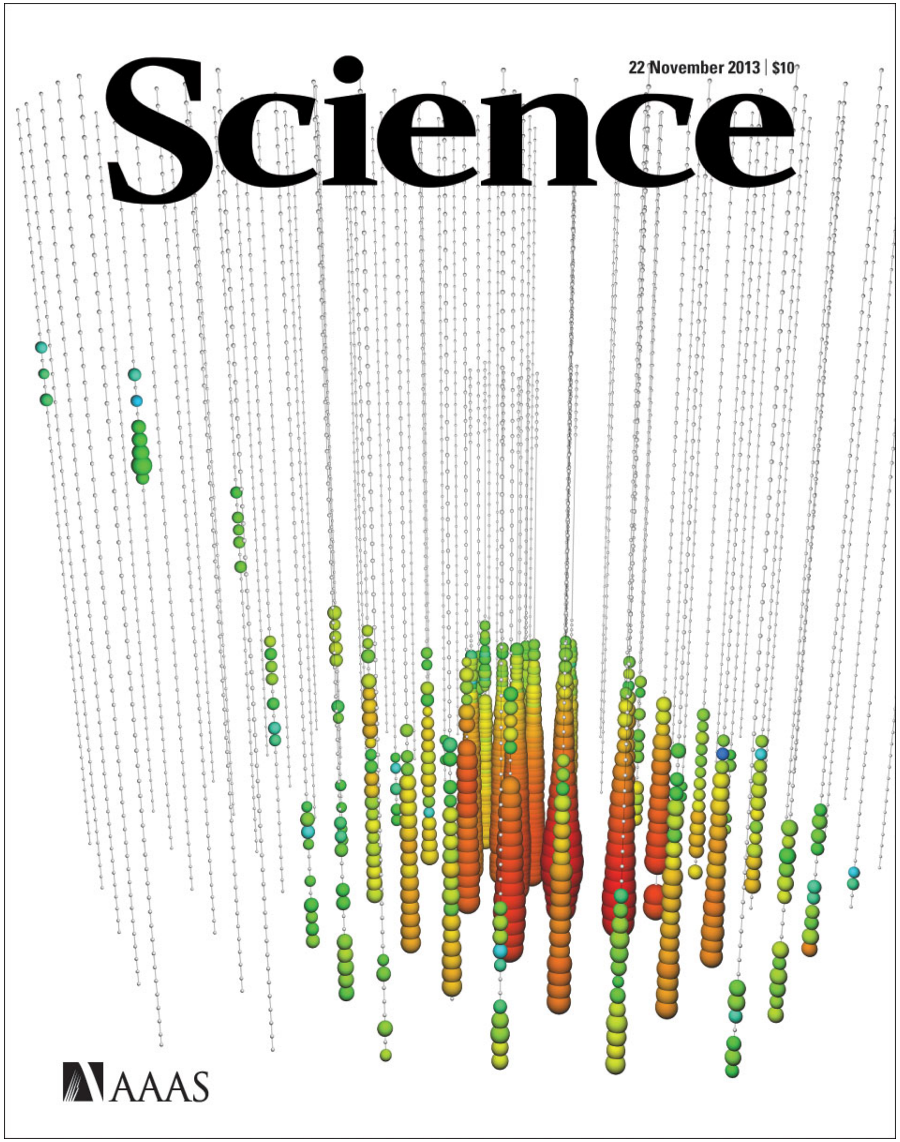
Measured information per event

Event consists out of list of "hits".

- Time
- OM (PMT*) number $\rightarrow$ x,y,z
- (observable correlating with number of photons)

Allows reconstruction of

- direction
- energy
- (interaction)

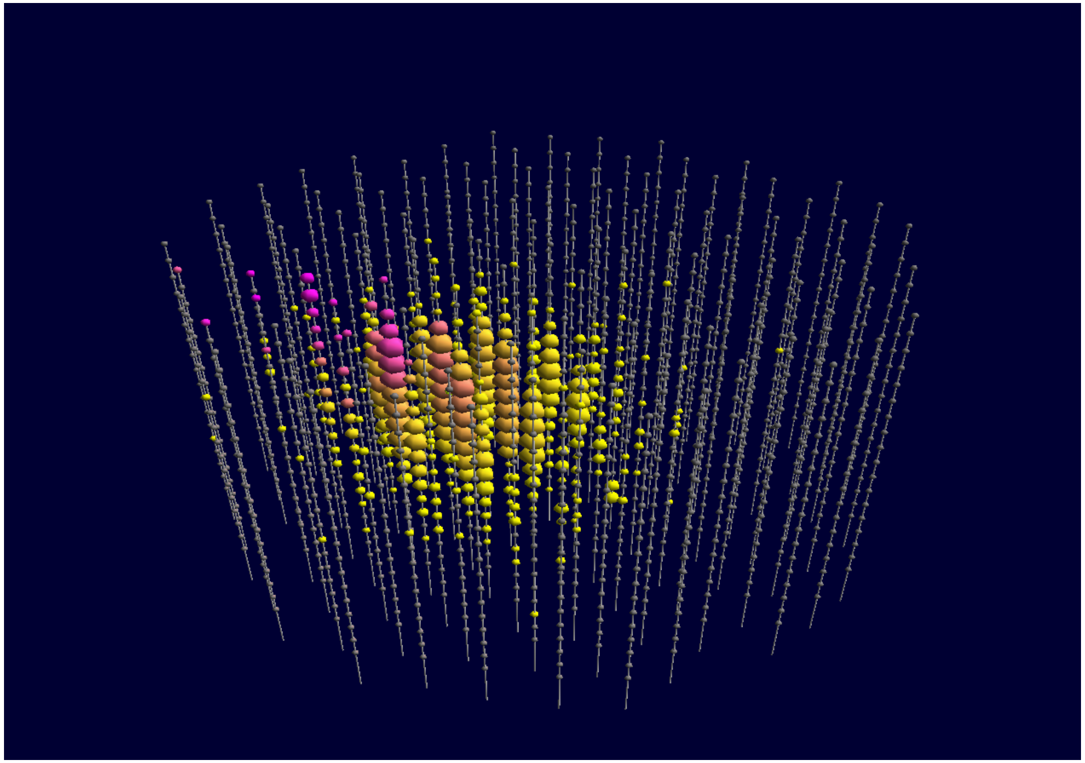


Hits to high-level information

- Reconstructions
  (e.g. Maximum likelihood)

- Classical pattern recognition
  - Background suppression
  - Energy reconstruction

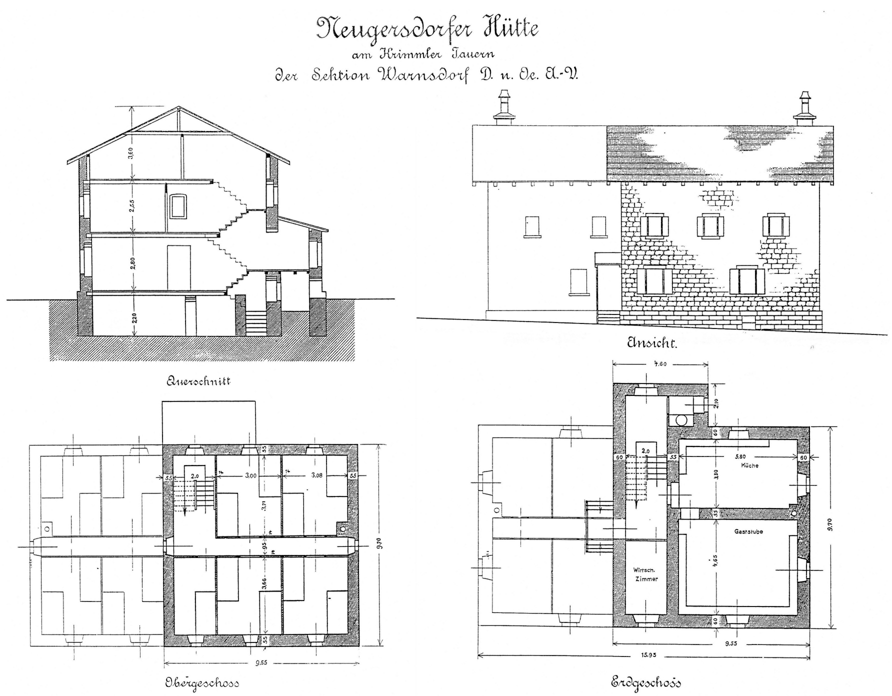
- Deep Learning:
  - Improve pattern recognition?
  - New applications feasible? (e.g. direction regression)
  - Faster developments?

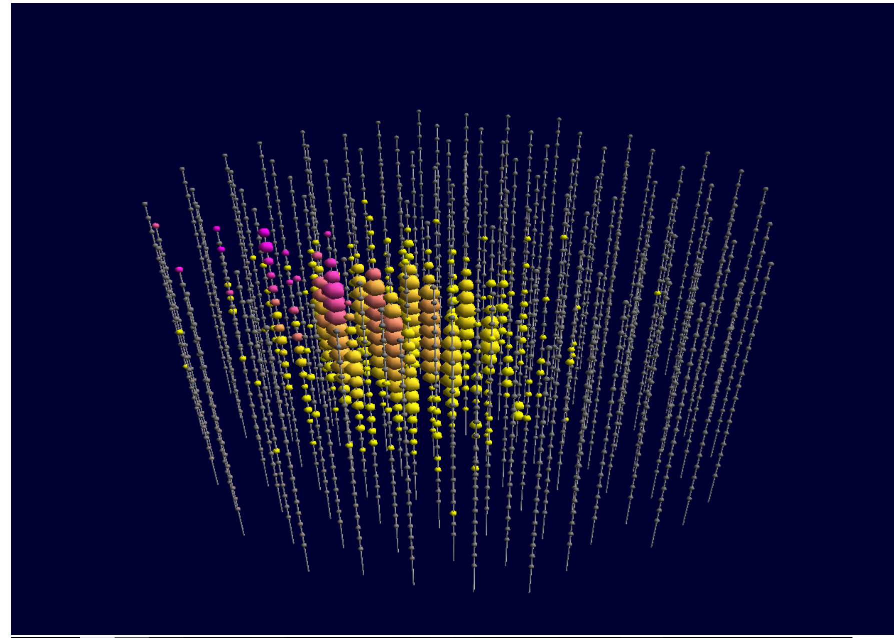


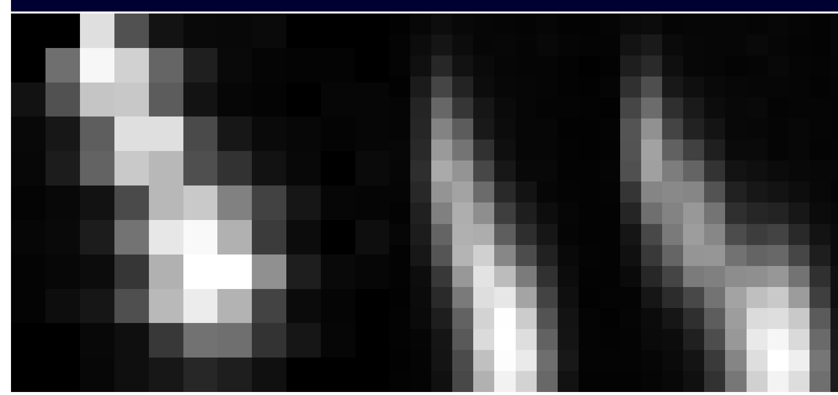
How to best feed x,y,z,t data to a framework most often used for images?

Hits to high-level information

Two dimensional projections?

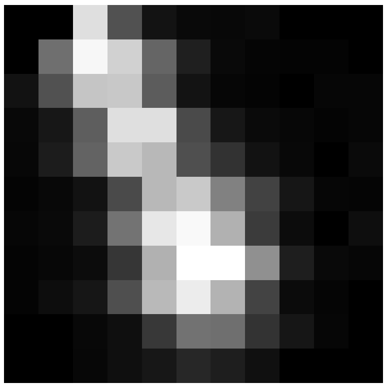




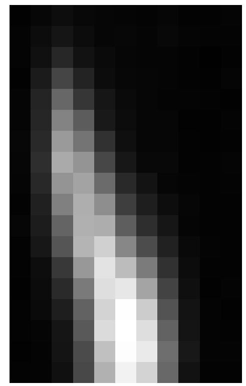


Ansatz I

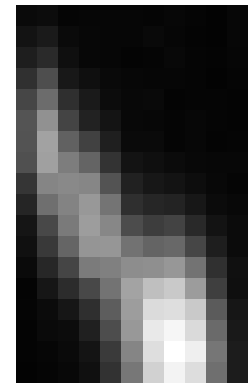
OM number + time → x,y,z + time → 2D projection (histogram)



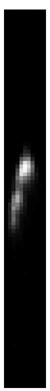
x-y, 11 x 11 bins



x-z, 11 x 18 bins

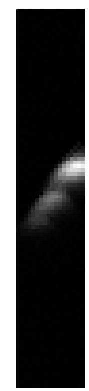


y-z, 11 x 18 bins



x-t, 11 x 100 bins

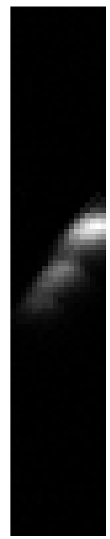
y-t, 11 x 100 bins

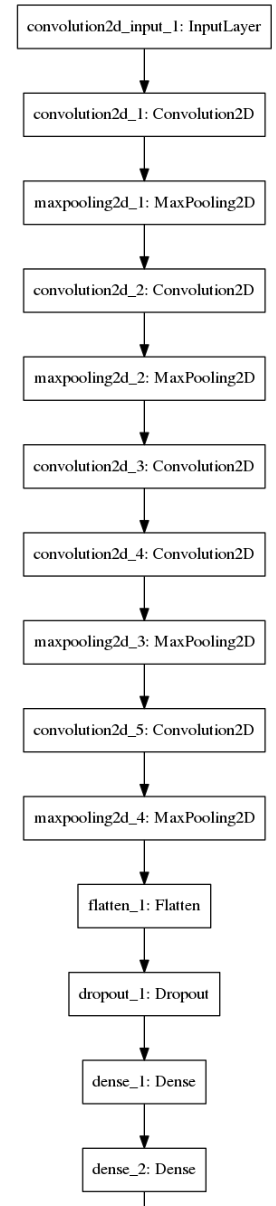


z-t, 18 x 100 bins

Testing Ansatz I

- Up/Down classification
- z-t as input
- CNN
- $\approx$ 93 % accuracy
- Work in progress, suggestions welcome

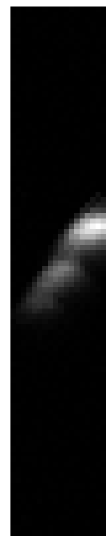


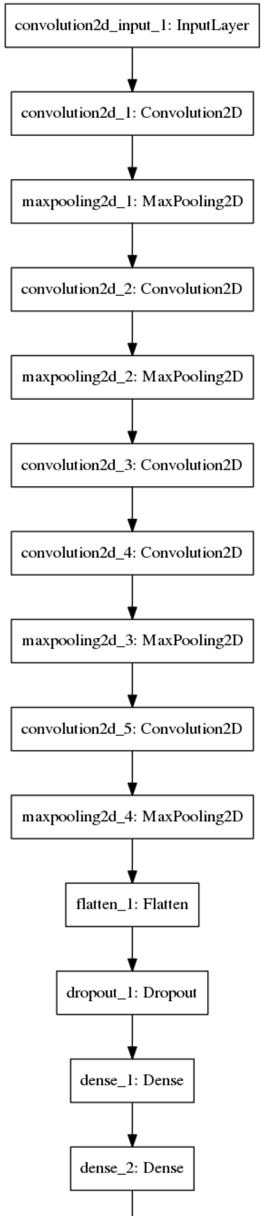


Testing Ansatz I

- Up/Down classification
- z-t as input
- CNN
- $\approx$ 93 % accuracy
- Work in progress, suggestions welcome

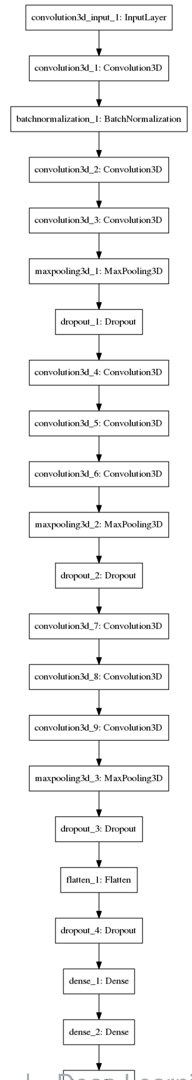
- As good as a Master's student (me) at this task with shallow learning after one year!





Ansatz II

OM number + time → x,y,z + time → Three dimensional projections



- x-y-z
  11 x 11 x 18 bins
- x-y-t
  11 x 11 x 100 bins
- x-z-t
  11 x 18 x 100 bins
- y-z-t
  11 x 18 x 100 bins
- r-z-t
  11 x 18 x 100 bins

- Up/Down classification
- r-z-t $\Rightarrow$ 96.5 %
- Better than z-t
- About as good as one year Master's and first year PhD student with shallow learning*

*proves how well CNNs work!

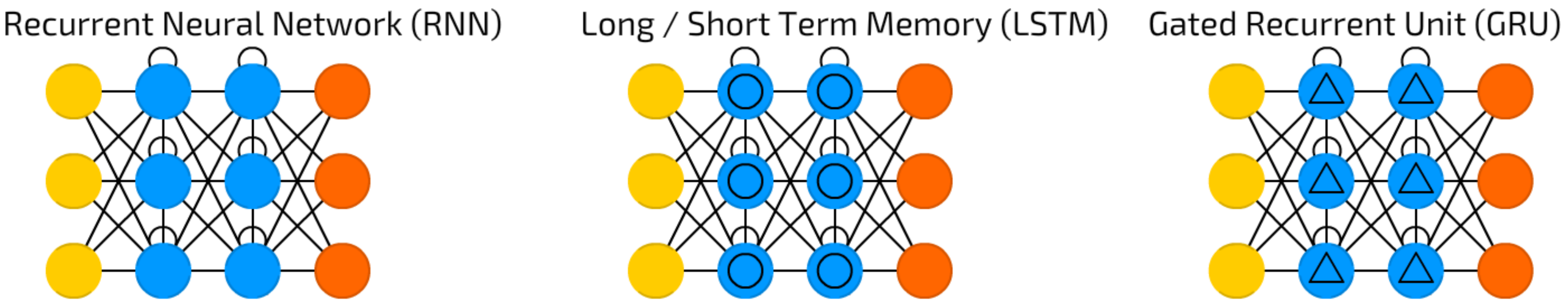
Obvious next step - Ansatz III

OM number + time → x,y,z + time → Full four dimensions

- x-y-z-t
  11 x 11 x 18 x 100 bins
- No 4D convolution implemented (yet)
- Maybe test CNTK

Ansatz IV

OM number + time → RNN / LSTM / GNN / ...

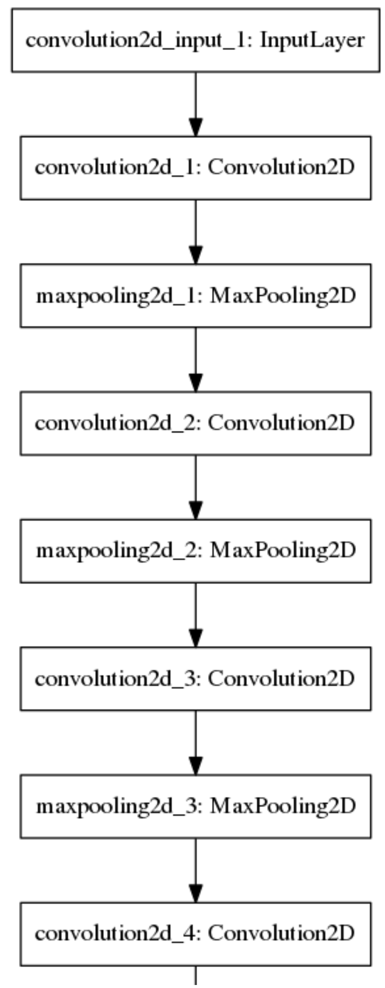


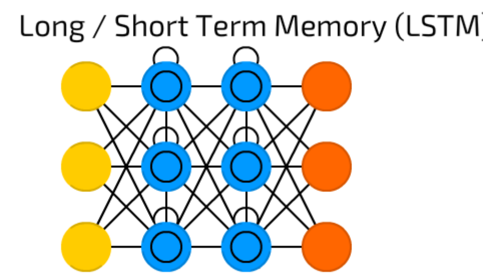
Not tested yet.

Ansatz V

OM number + time → x,y,z + time → Time series of 3D projections

- 100 x-y-z histograms, one for each time bin
- 100 times 11 x 11 x 18
- CNN on x-y-z
  ⇒ feature reduction for each time bin





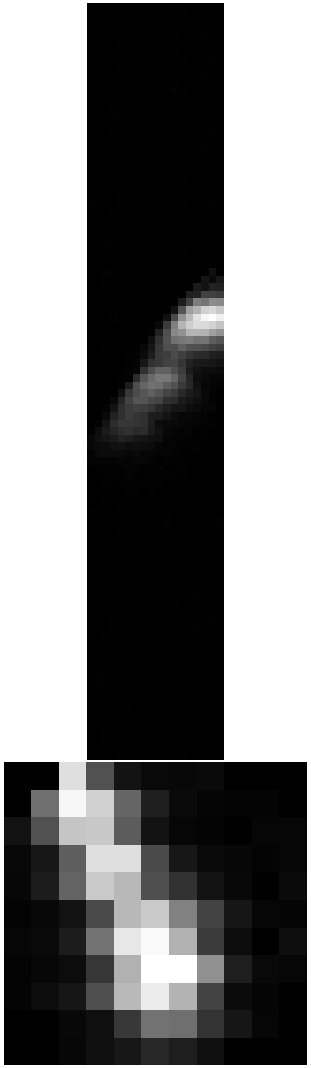
- LSTM for time series of CNN outcome
- Not tested yet

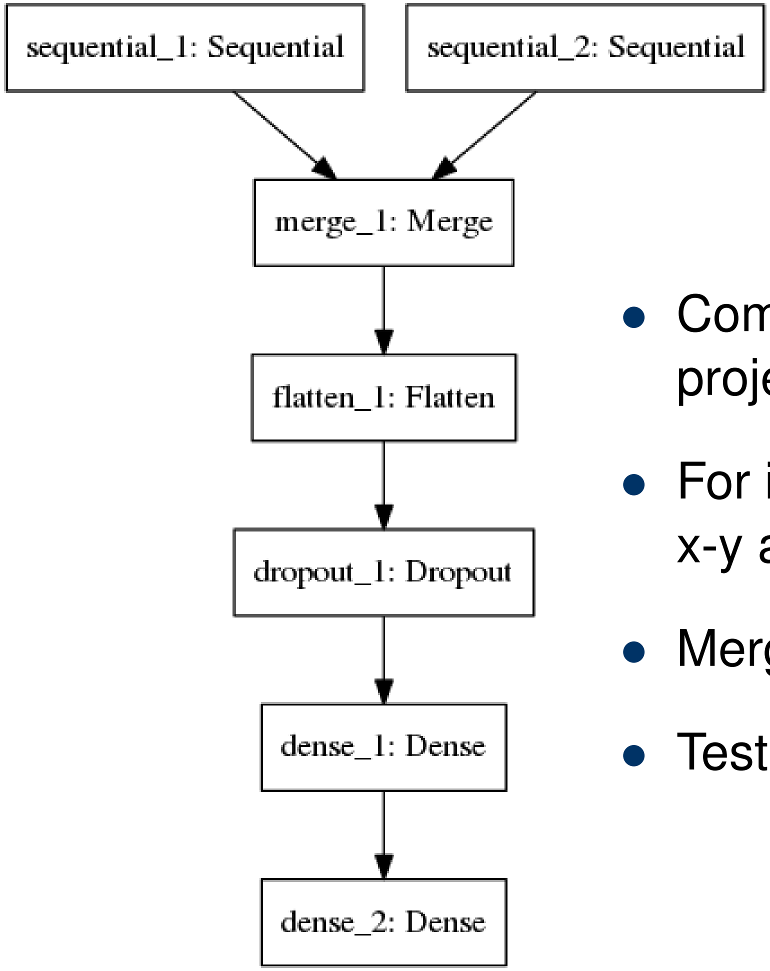
Summary

- Neutrino telescopes produce (at least) four dimensional data
- Many options to handle this data
- Deep convolutional nets applied successfully
- Recurrent nets will be investigated
- More applications to come
- You are highly welcome to suggest improvements

Thank you for your attention!

Ansatz VI

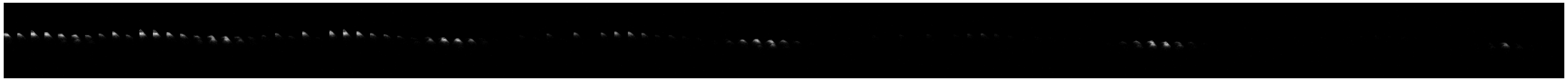




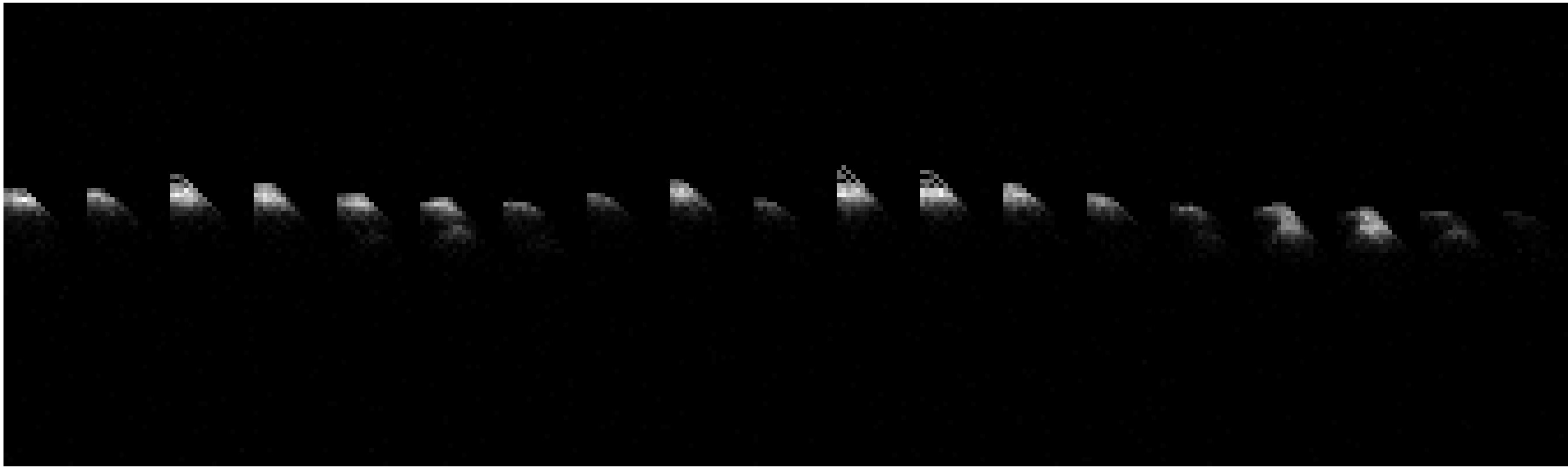
- Combine multiple 2D projections
- For instance:
  x-y and z-t as input
- Merged CNNs
- Testing currently

Ansatz VII

OM number + time → Two dimensional projection

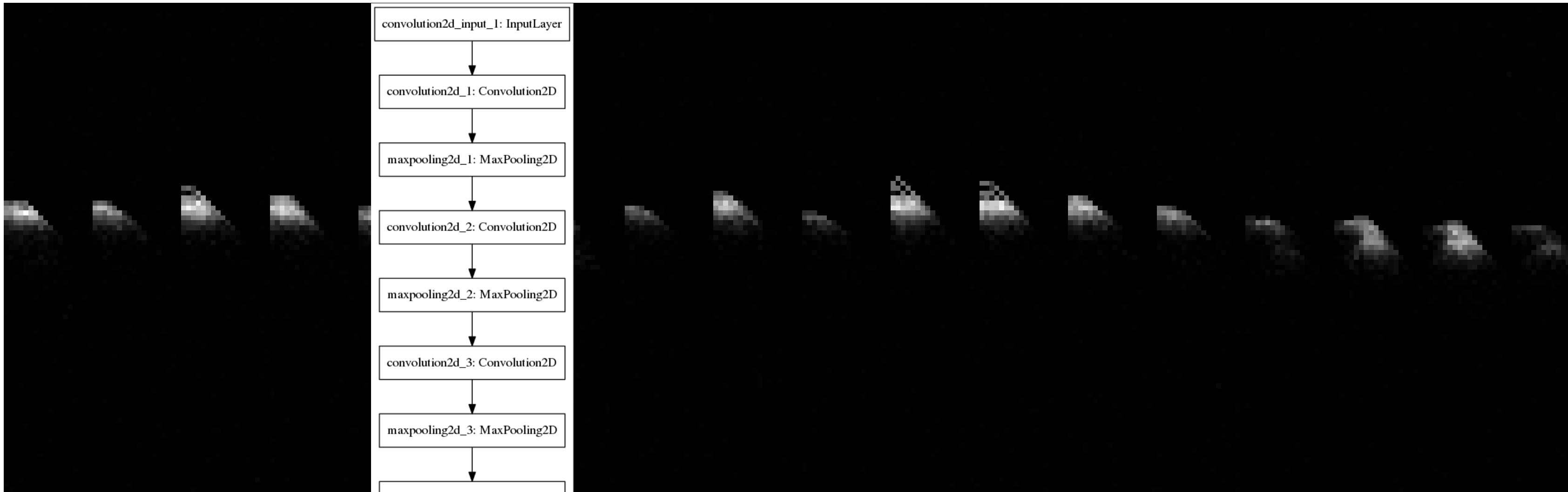


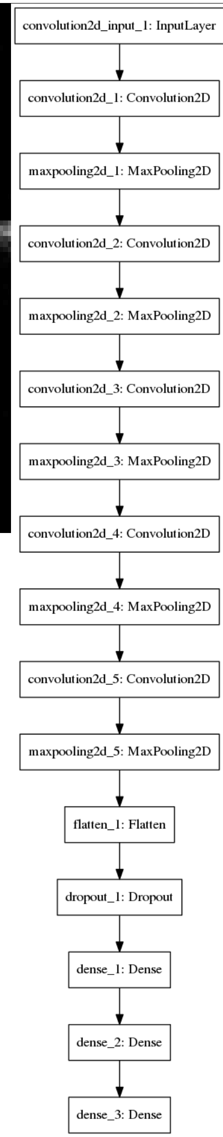
OM number - t, 2070 x 100 bins



Zoom

Ansatz VII





- CNN
- $<$ 86 % accuracy (Up/Down)
- Poor convergence
- Being investigated further

Other ideas so far

- Ditch convolution, use existing hand-crafted features
  - Tested with CNTK for particle / interaction identification
  - Not (yet) better than best shallow learning
- Use other projections (e.g. reconstructed direction)
  - Could introduce unnecessary uncertainty
- Regress energy and direction
- Estimate error of regression (Gaussian processes?)
- Investigate unsupervised (pre)training