iDark
The intelligent dark matter survey

Radboud University

netherlands eScience center

ANALYSING ɣ-RAYS OF THE GALACTIC CENTER WITH DEEP LEARNING

LUC HENDRIKS
RADBOUD UNIVERSITY NIJMEGEN

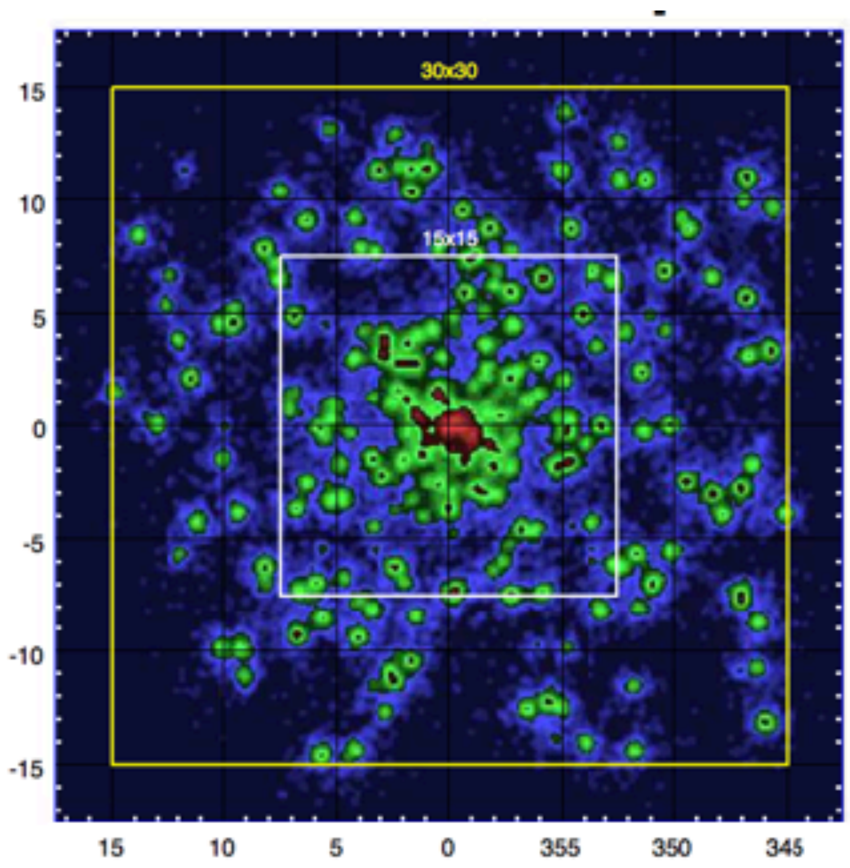
OUTLINE

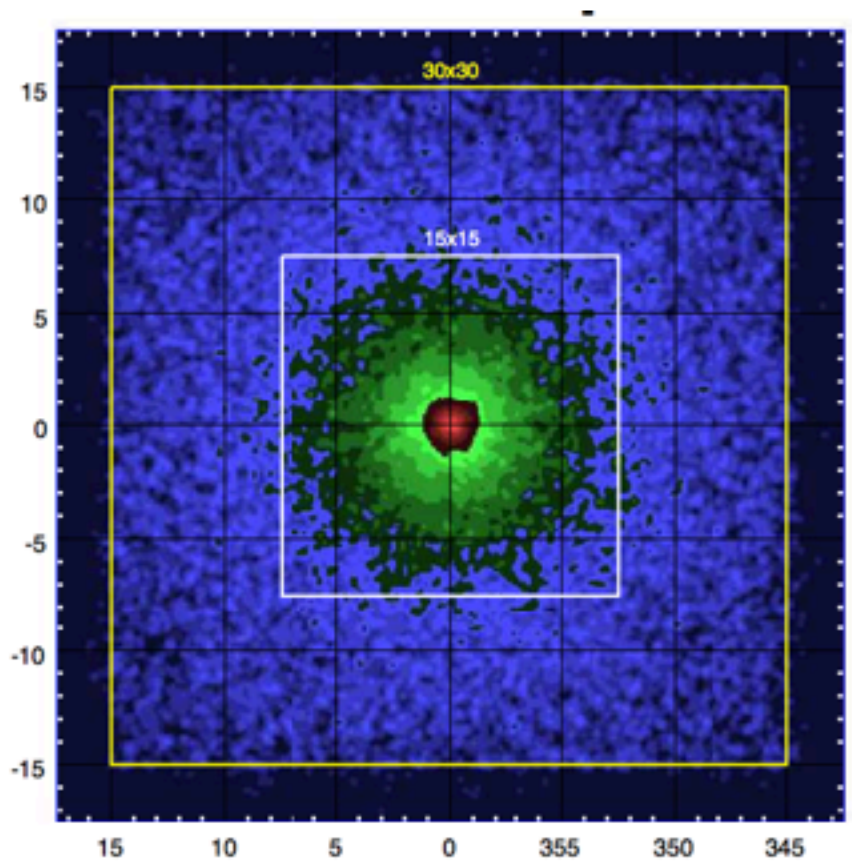
- Introduction
- Deep learning & ConvNets
- Analysing ɣ-rays using deep learning
- Conclusions + outlook

- Together with:
  - Sascha Caron
  - Germán Gómez-Vargas
  - Roberto Ruiz de Austri
- Arxiv: 1708.06706

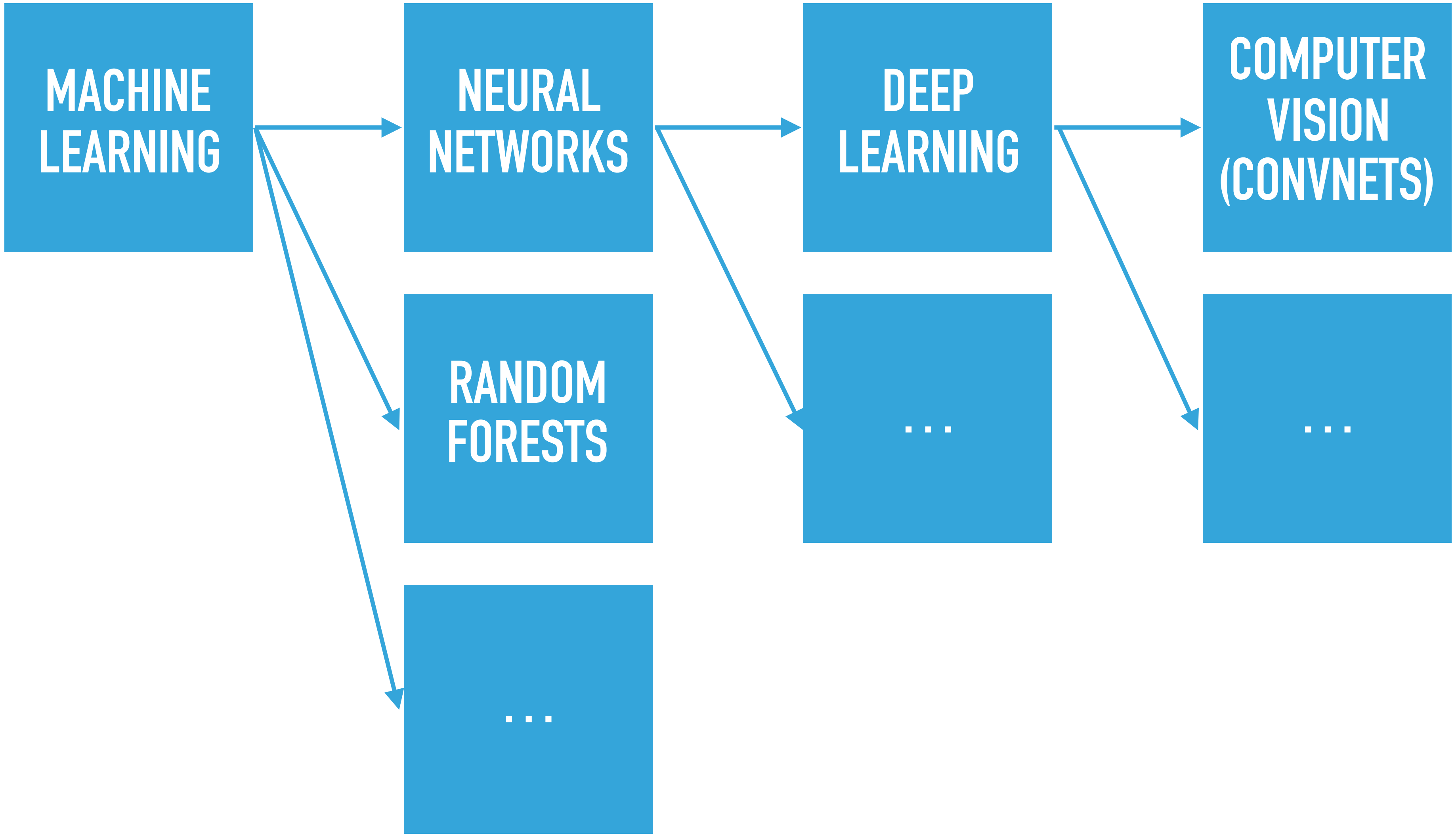
INTRODUCTION

- Is the GC excess from a diffuse source or collection of unresolved point sources?
- Can we answer this question with deep learning?
- Assume diffuse and point source component distributed by gNFW profile





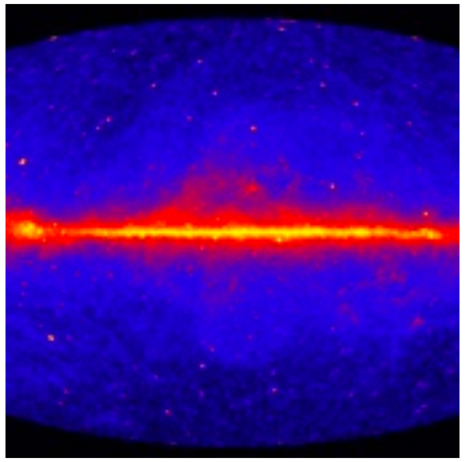
WHAT IS DEEP LEARNING?

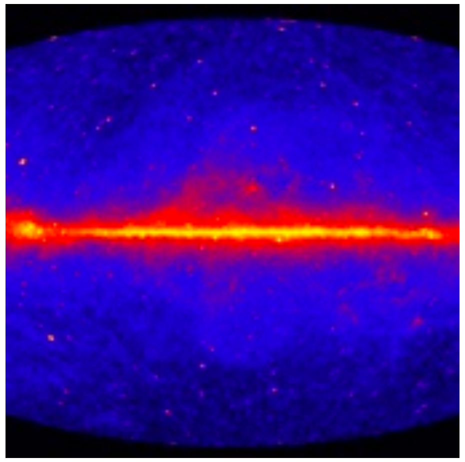


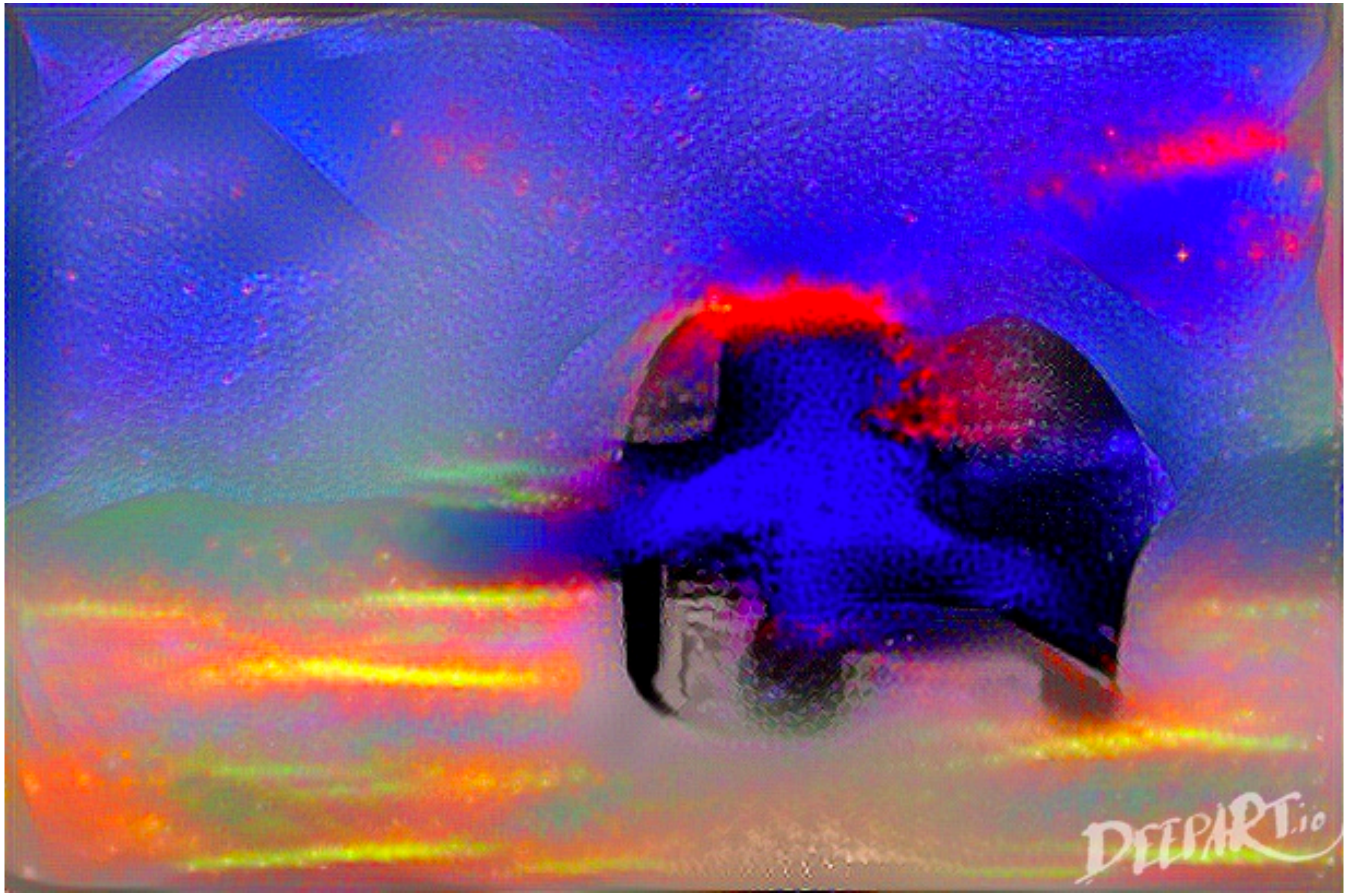
Instead of the programmer defining what the computer should do, supply an objective and learn from data

WHAT CAN YOU DO WITH DEEP LEARNING?

- Classification & regression
- Face detection, galaxy classification, …
- Other cools things like:
  - Create the best Go player
  - Generate music
  - Make photo books
  - Style transfer
  - Generate music
  - …

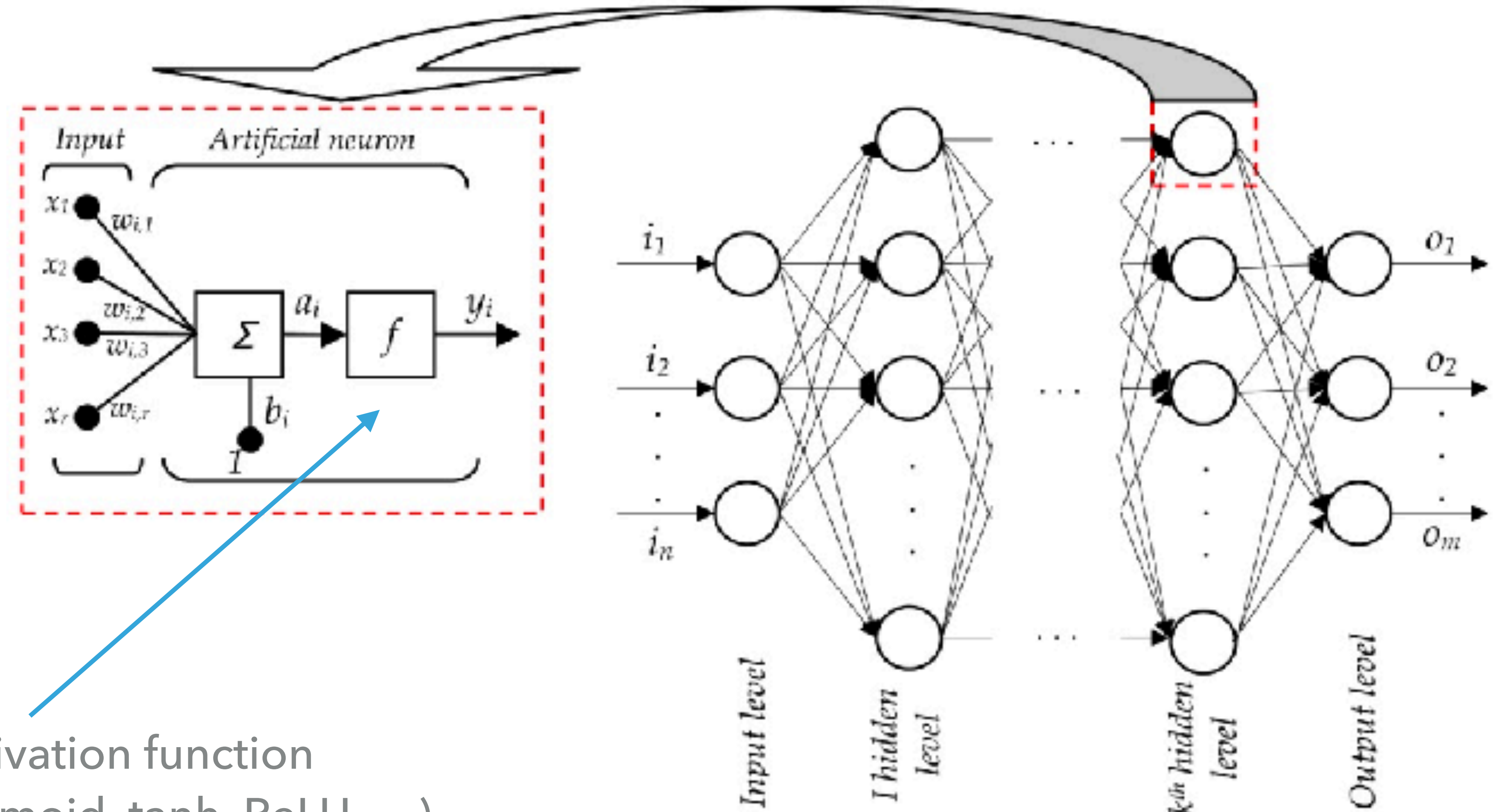






deepart.io

NEURAL NETWORKS & DEEP LEARNING

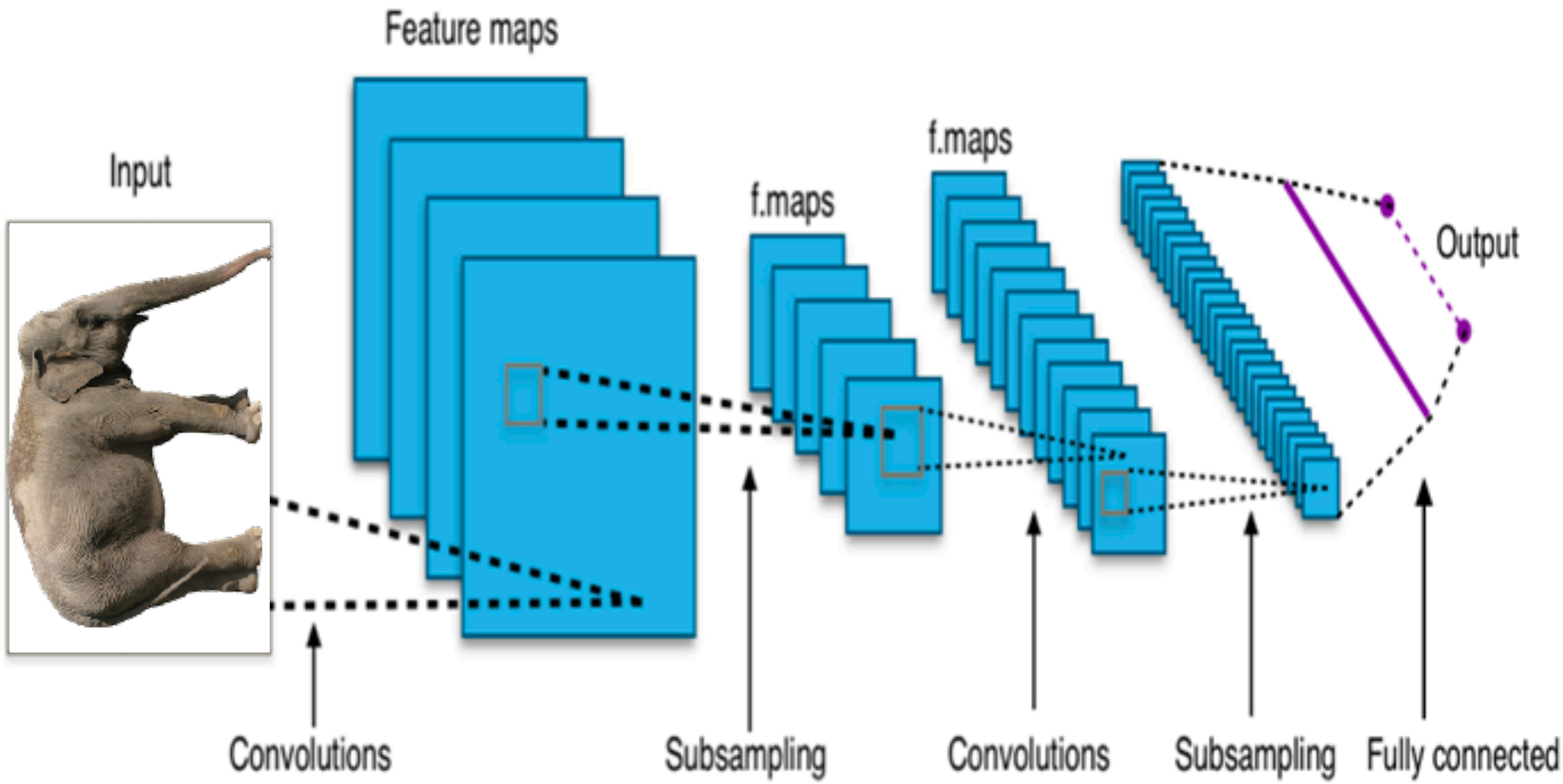


Activation function
(sigmoid, tanh, ReLU, …)

During training:

Backpropagate error & update weights using gradient descent

CONVOLUTIONAL NEURAL NETWORKS



Input (3 layers!)

CNN Layer 1 | CNN Layer 2 | CNN Layer 3

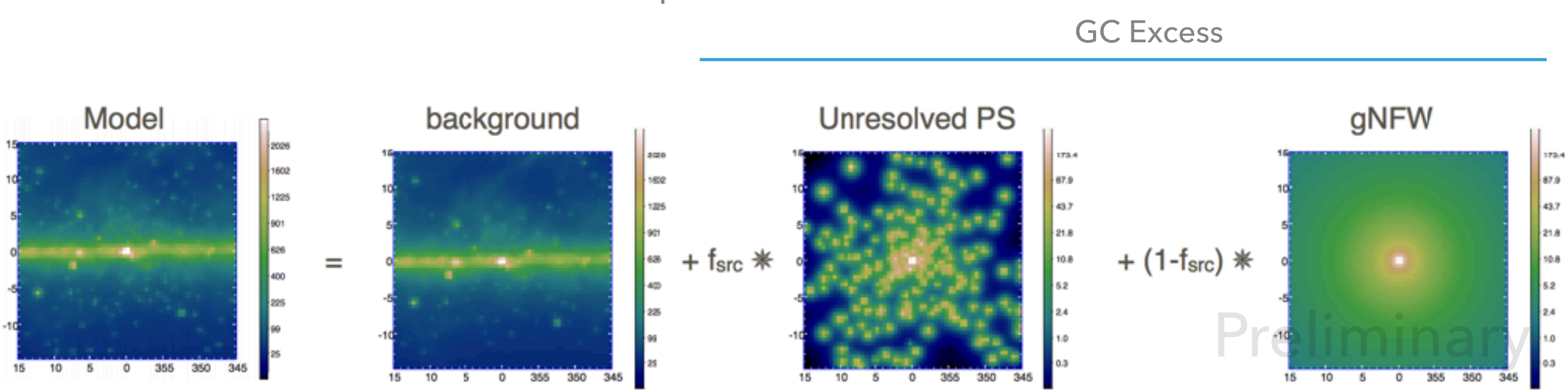
Output:

Elephant: 0.97
Grass: 0.01
Dog: 0.0001
...

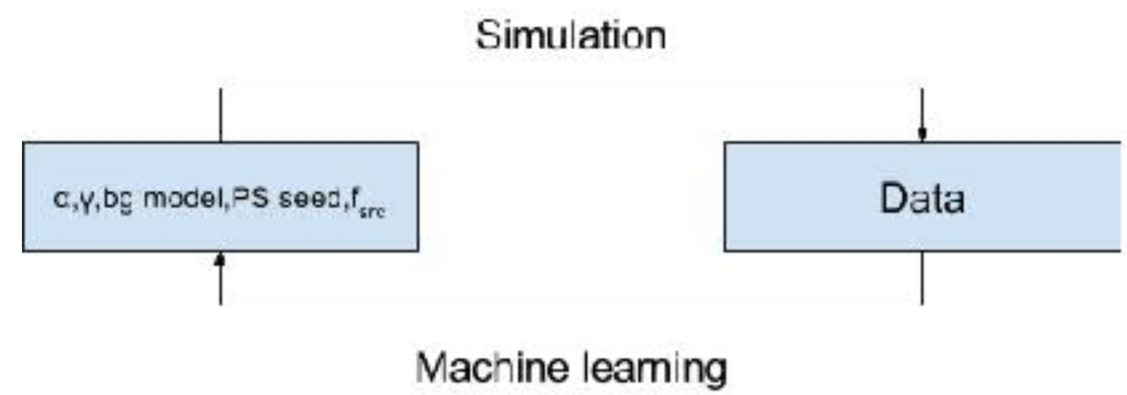
- Pros:
    - Only define data and output, no human assumptions in algorithm
    - NNs are universal approximators
- Cons:
    - They are black boxes
    - Training a ConvNet is still an "art"

HOW TO APPLY THIS TO ɣ-RAY DATA?

- Goal: determine the component of point sources vs diffuse source of the GC excess – f_{src}
- Simulate GC using Fermi tools (5 parameters)
  - Output is photon count map of photons between 1-6GeV
    (no spectrum information, will be improved in new version)
- Sample from simulations in 5D parameter space
- Train network to predict f_{src} accurately in all scenarios of the other components
- Apply on real data – sample big enough so that reality is somewhere in 5D space
- Network trained on simulated data to predict f_{src} – **simulation inversion**

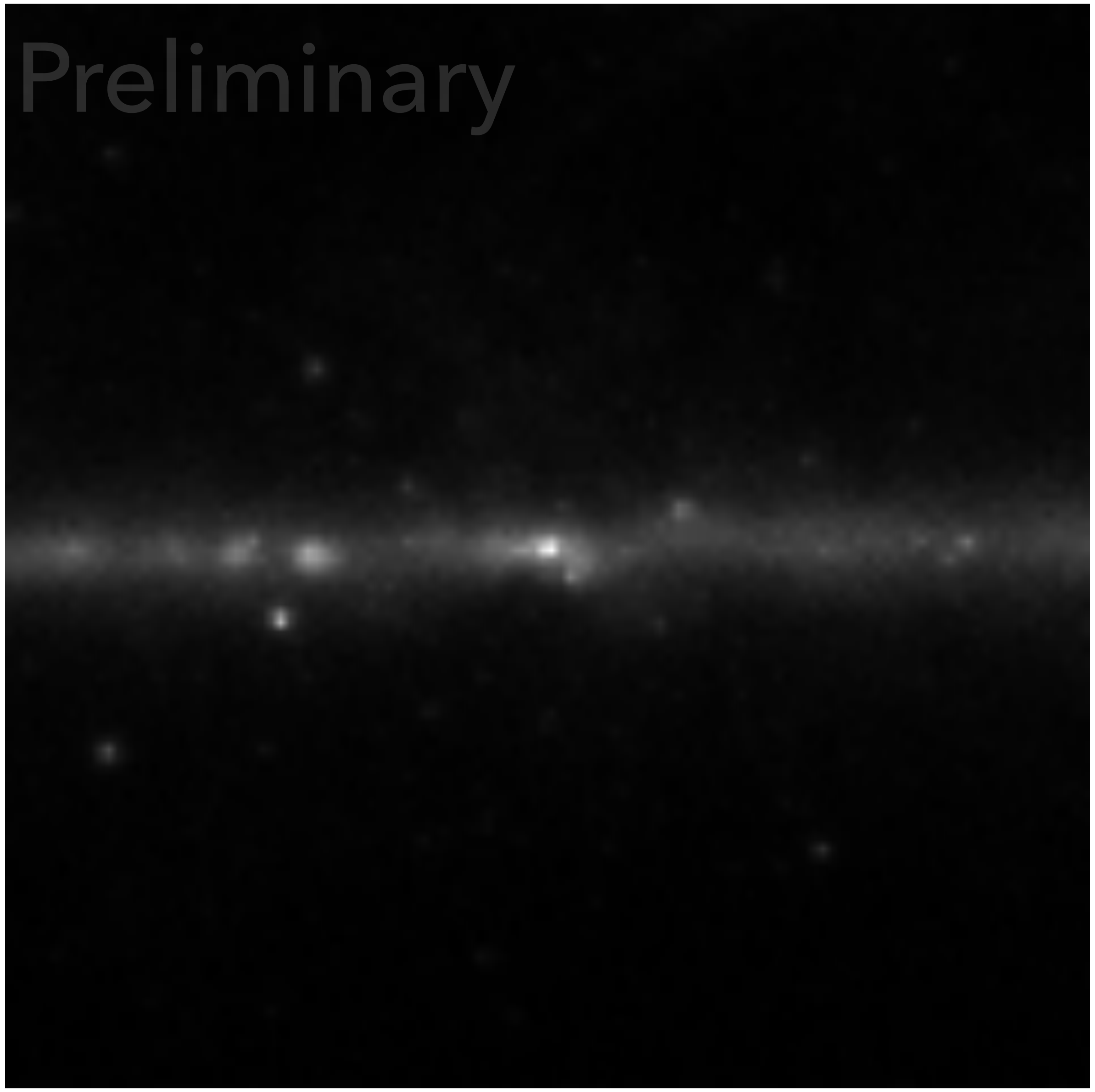


HOW TO APPLY THIS TO γ-RAY DATA?

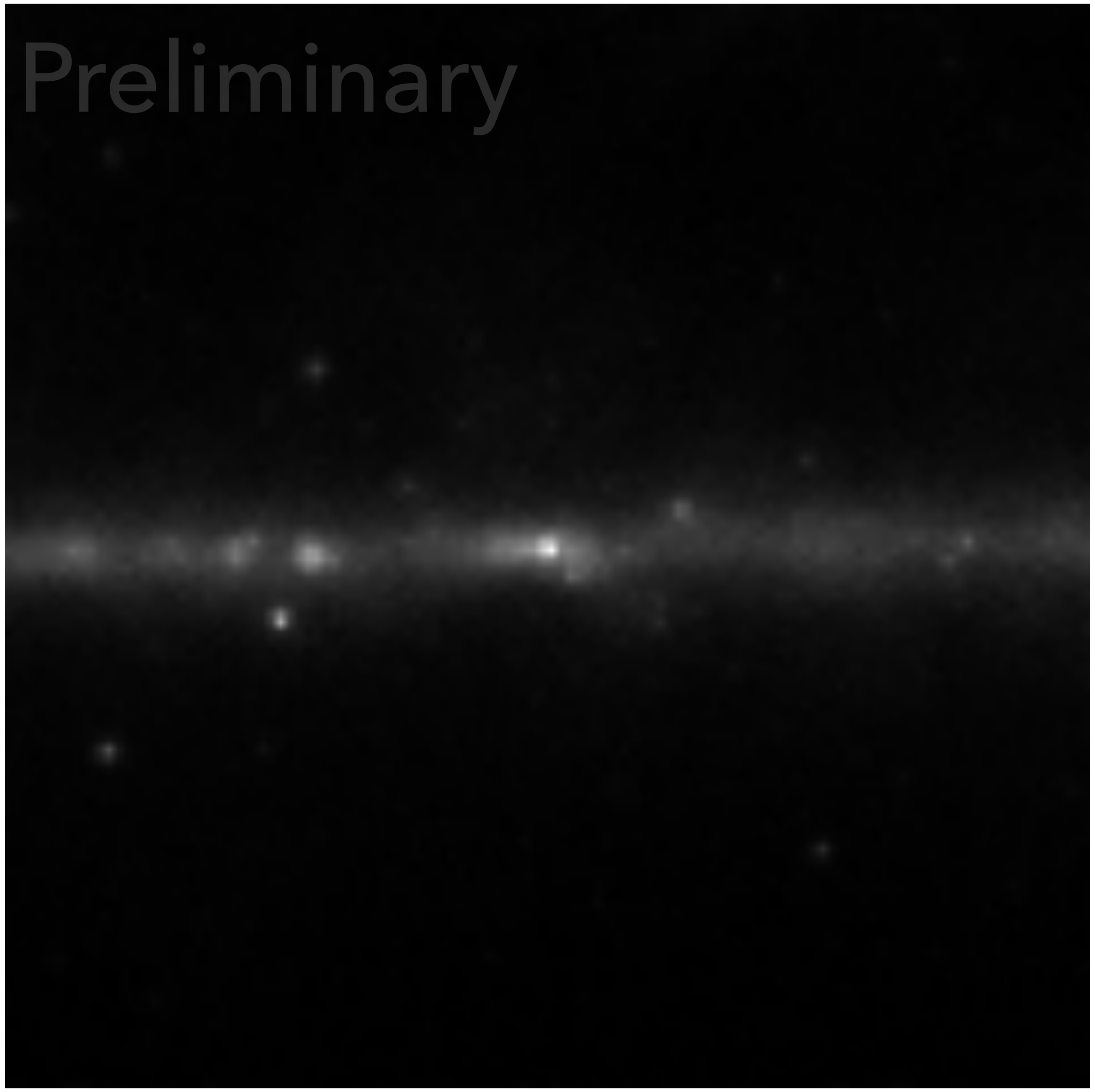


- 5 parameters are:
  - The fraction of diffuse vs point sources – f_{src}
  - Point source flux distribution – α $$\frac{dN_{src}}{ds} = As^{\alpha}$$ Determined by total excess
  - The value of γ in the gNFW profile - 1.1
  - The background model used (3 for train, 2 for testing)
  - The randomised locations of the point sources
- Using simulation go from parameters to image
- Using ConvNets from image to parameters

EXAMPLE OF TWO SIMULATIONS



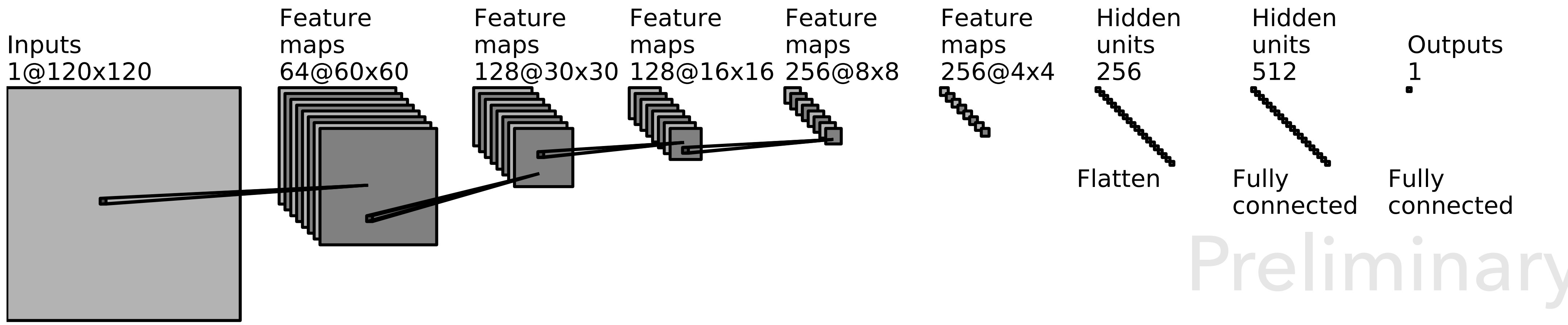
$f_{src} = 0.9883$



$f_{src} = 0.0275$

Try yourself! http://fermiai.s3-website-eu-west-1.amazonaws.com/

CONVOLUTIONAL NEURAL NETWORK



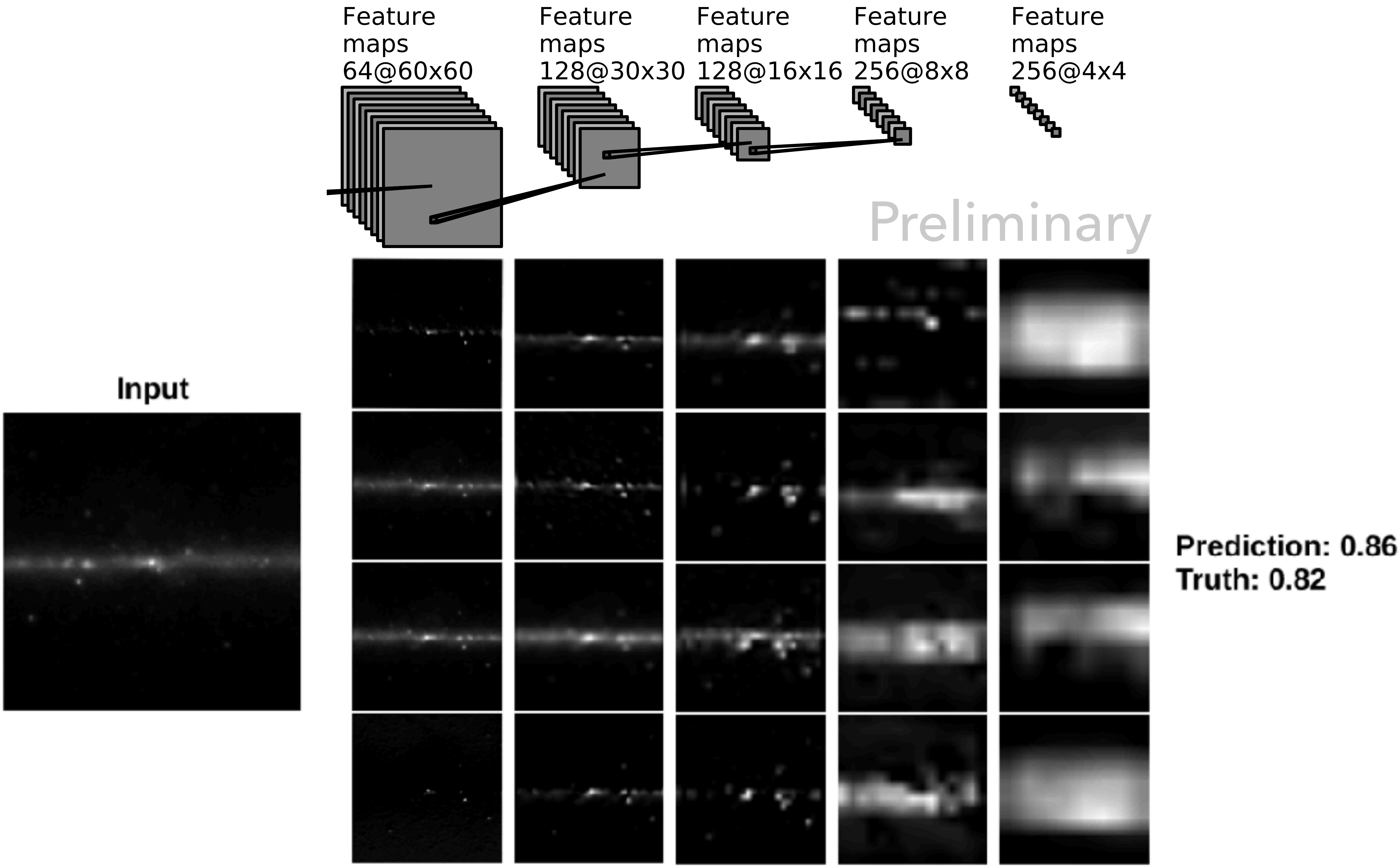
Max-pooling after every convolution
Local response normalization after every other convolution

$f_{src} = 0.89$

- Every layer has L2 regularisation (penalise high weights to prevent overfitting)
- 1.2 million images of 120x120 values
- ~10 million internal parameters
- 1 day to train each network (TensorFlow, 2x GTX1080, >5000 cores, ~16 TFLOPs)

2x as fast as the fastest supercomputer in 2000!

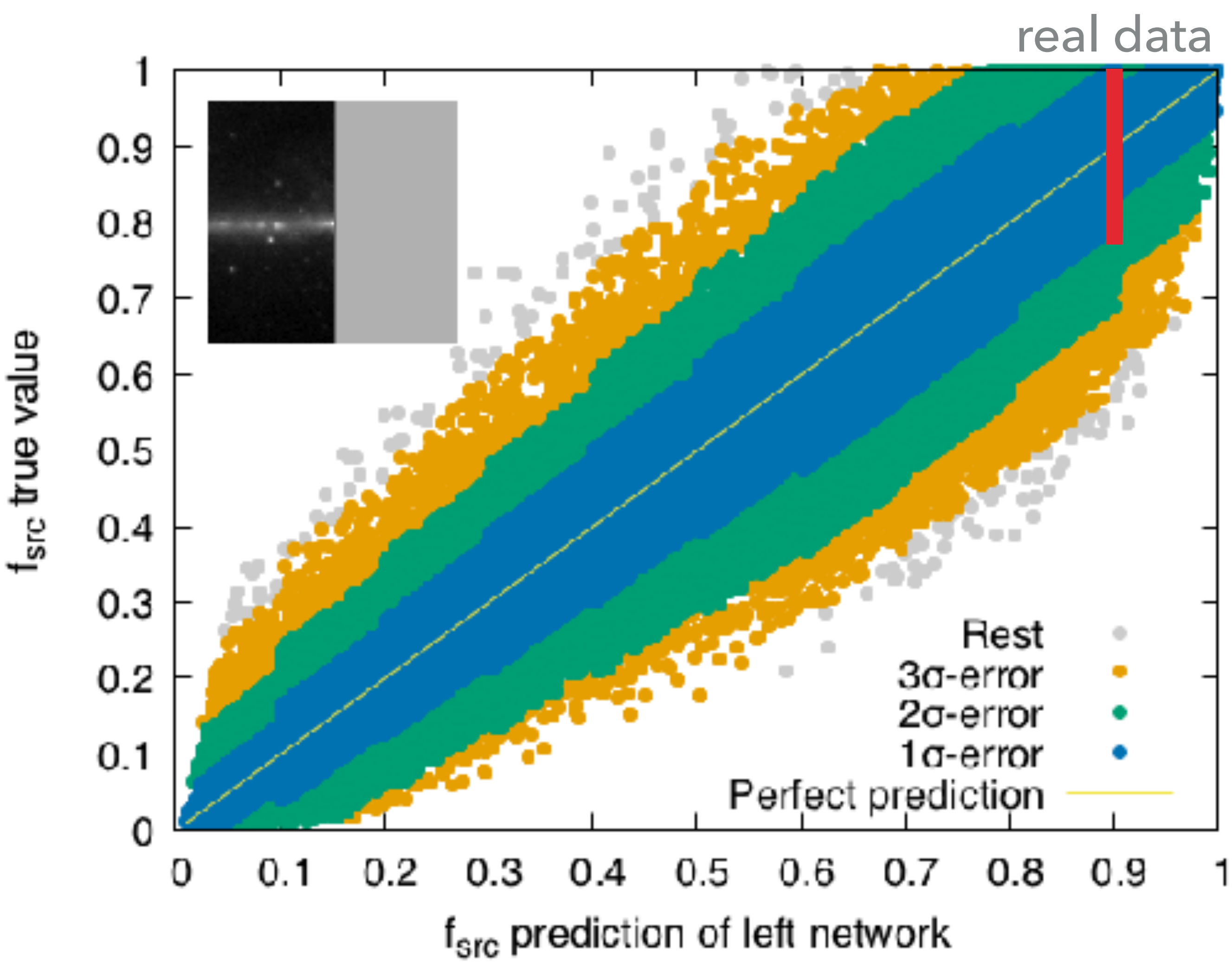
A LOOK INSIDE THE NETWORK

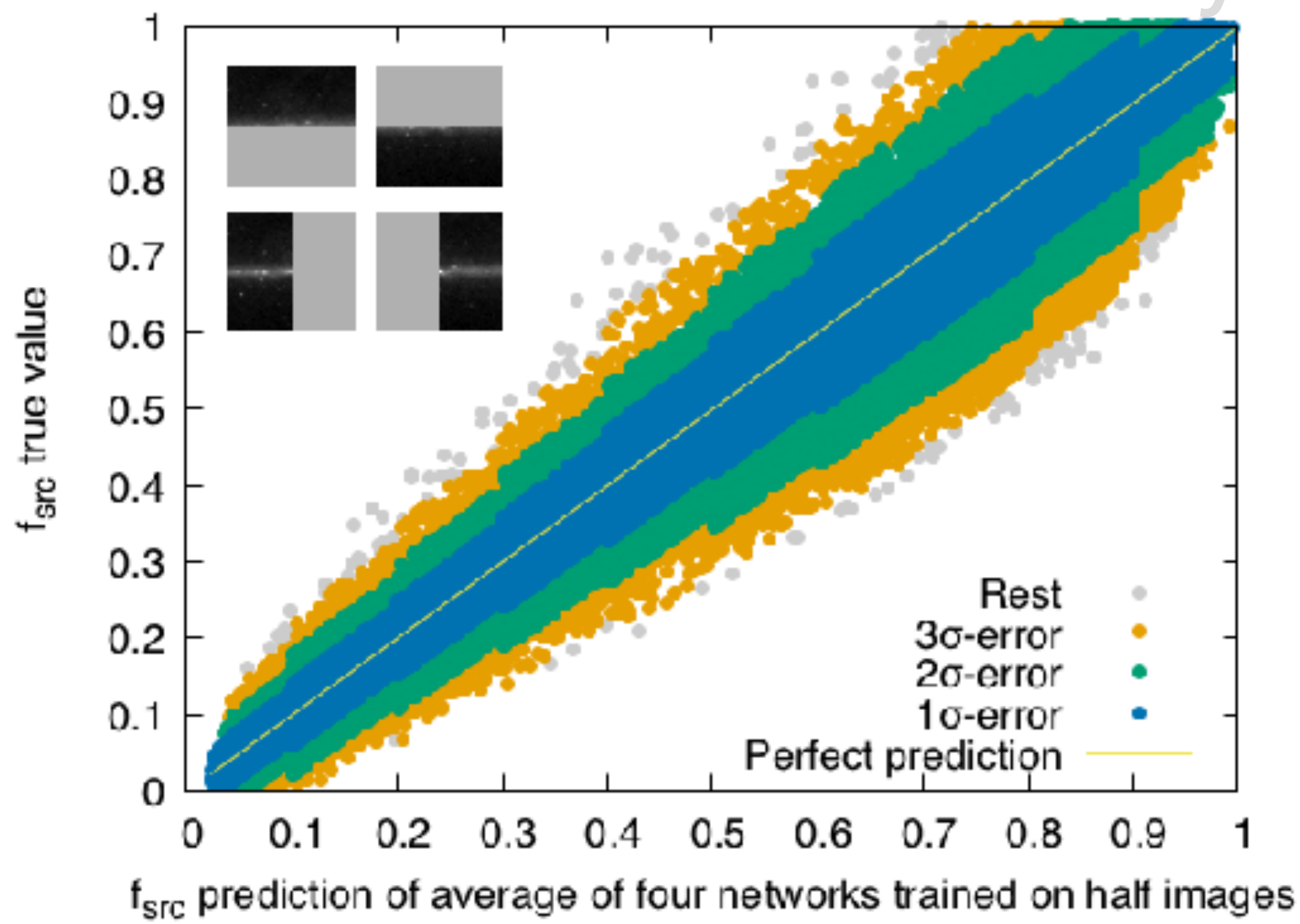


RESULTS

- Train using 3 background models, test on 2 others
- Test data: 2x30000 test points

Preliminary





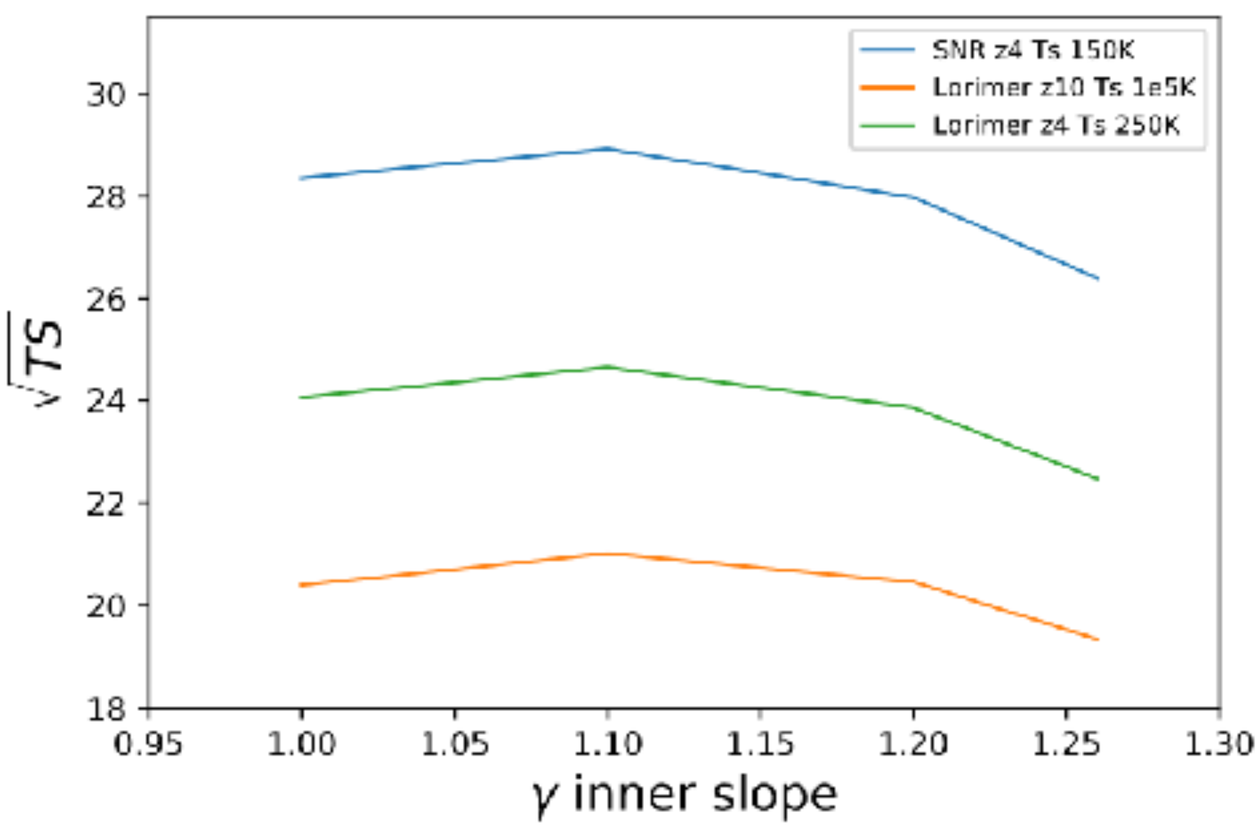
Because we are doing a followup study, real result is only evaluated on the left network to not bias ourselves

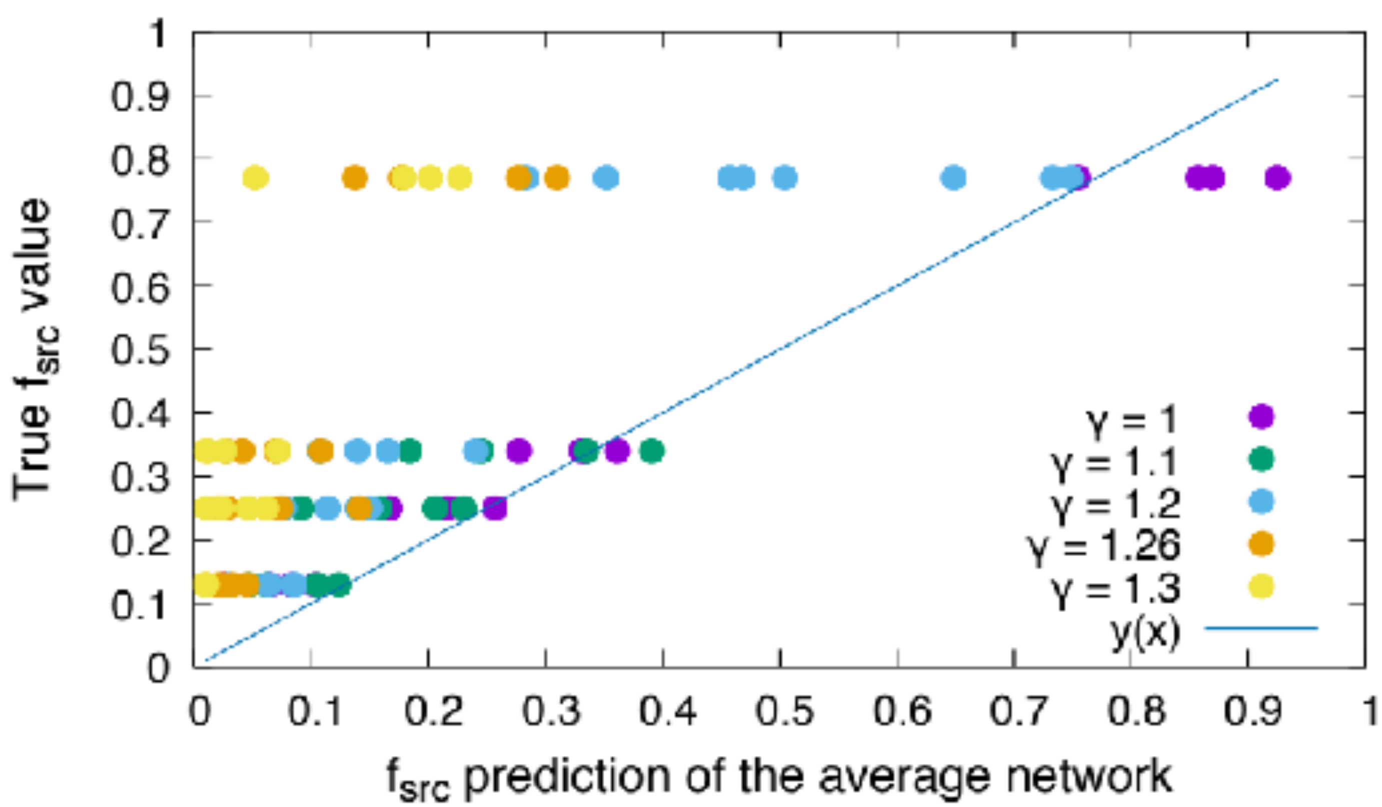
CONCLUSIONS + OUTLOOK

- Network successfully generalised over all training parameters
    - Utilise research by Google, Facebook, OpenAI, etc…
    - Applicable to many other problems as well – regression & classification
- Strongly disfavours 100% diffuse (in agreement with other studies)
    - Give a prediction of actual mixture
- Followup study needed, improvements necessary in data generation
    - Use updated catalog (now 3FGL was used with pass 8 data. Does not include all sources)
    - Extend range of α from [-1.05, 1.05] to [-1.2, 1.2]
    - Allow for different γ values
    - Use multiple energy bins in the input data (now only 1 bin with 1-6 GeV)

BACKUP SLIDES

GAMMA DEPENDENCE





BG MODELS USED

Usage	CR distribution	Halo height z (kpc)	T_S (K)	Log$\mathcal{L}$	$\langle\sigma v\rangle \times 10^{-27}$ cm^3/s
Training A	SNR	10	150	-442855	45.59
Training B	Lorimer	10	1×10^5	-442304	33.61
Training C	Lorimer	4	150	-442357	39.32
Testing A	Lorimer	10	150	-442539	39.63
Testing B	SNR	4	1×10^5	-442664	42.67

SUSY DM – SIGMA SD