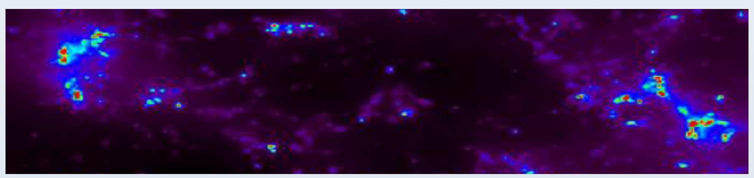
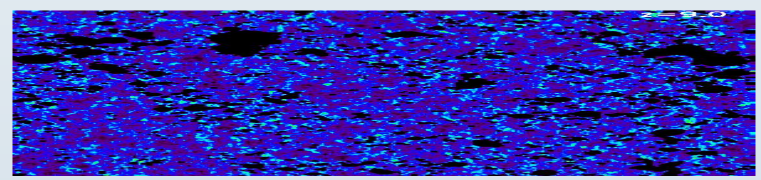
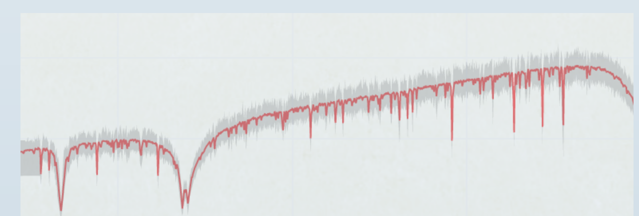
Networks Learning the Universe: From 3D (hydrogen tomography) to 1D (classification of spectra)







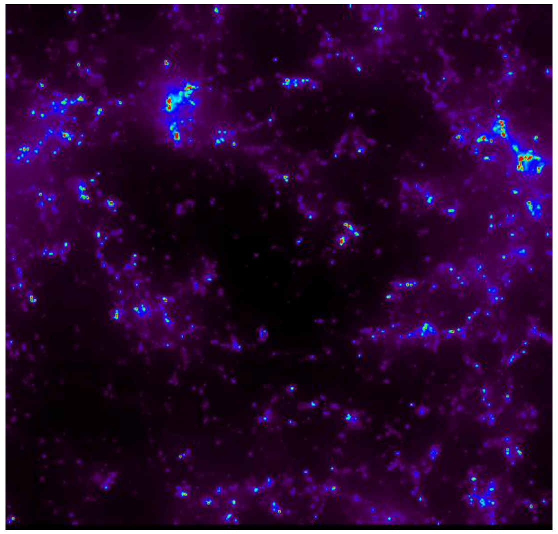
Caroline Heneka
Hamburg Observatory & Excellence Cluster Quantum Universe, Universität Hamburg

ESA/ESO SciOps Workshop 2022, May 17th 2022

Learning the Universe: Advances

Driven by large-scale + high-resolution surveys

Example: Intensity Mapping



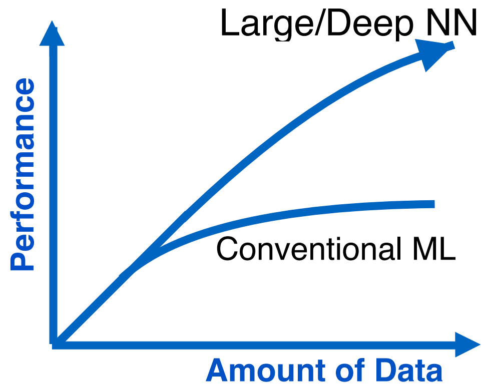
Lyα emission < 1 billion years after Big Bang

Extract more & less biased (?) information
Data mining

Efficient data reduction
Automation

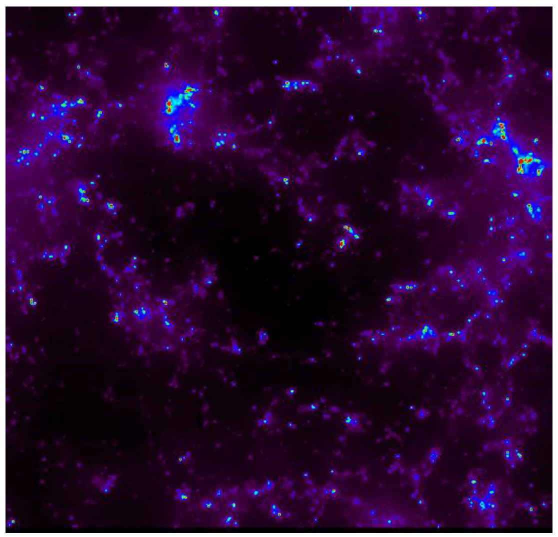
Complementarity to
previous probes + methods

Example: Deep Learning
Driven by ability to improve with large datasets



Learning the Universe: Intensity Mapping (IM)

Astrophysics & Cosmology ∞ **AI and Computer Vision**



Lyα emission < 1 billion years after Big Bang

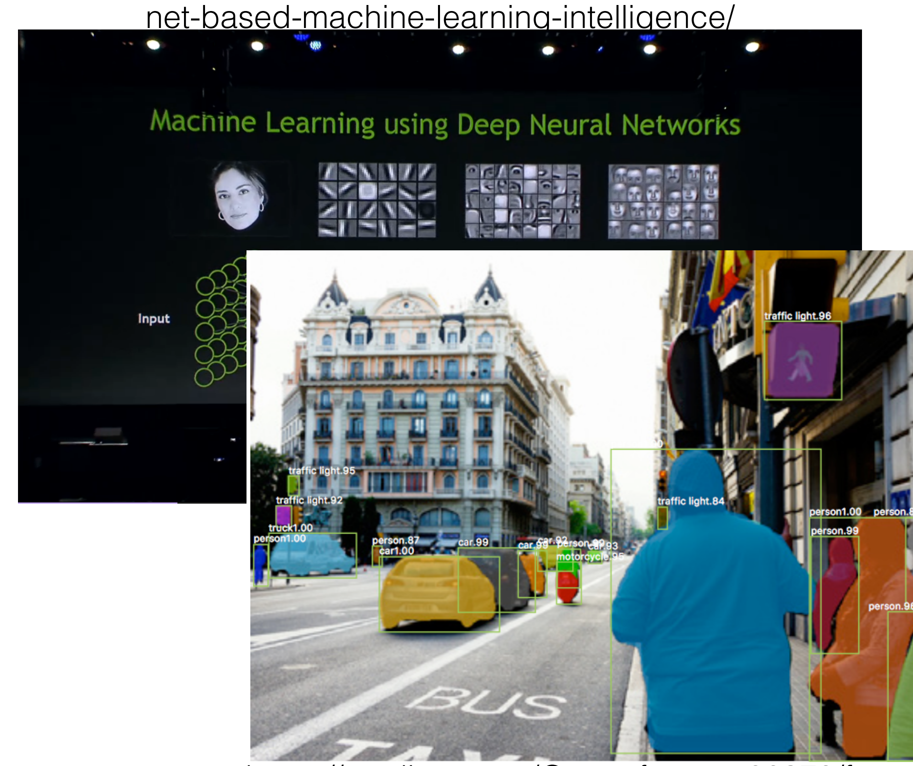
Modelling
LSS
Cosmological signatures
Sources

Data
Imaging
Spectral
3D cubes
Time series

Simulations
Computational Astrophysics
Reionisation

Neural Networks
Convolutional
Generative
Classifier
Regressive

https://wccftech.com/nvidia-demo-skynet-gtc-2014-neural-net-based-machine-learning-intelligence/



https://medium.com/@umerfarooq_26378/from-r-cnn-to-mask-r-cnn-d6367b196cfd

Intensity Mapping ∞ **Deep Learning**

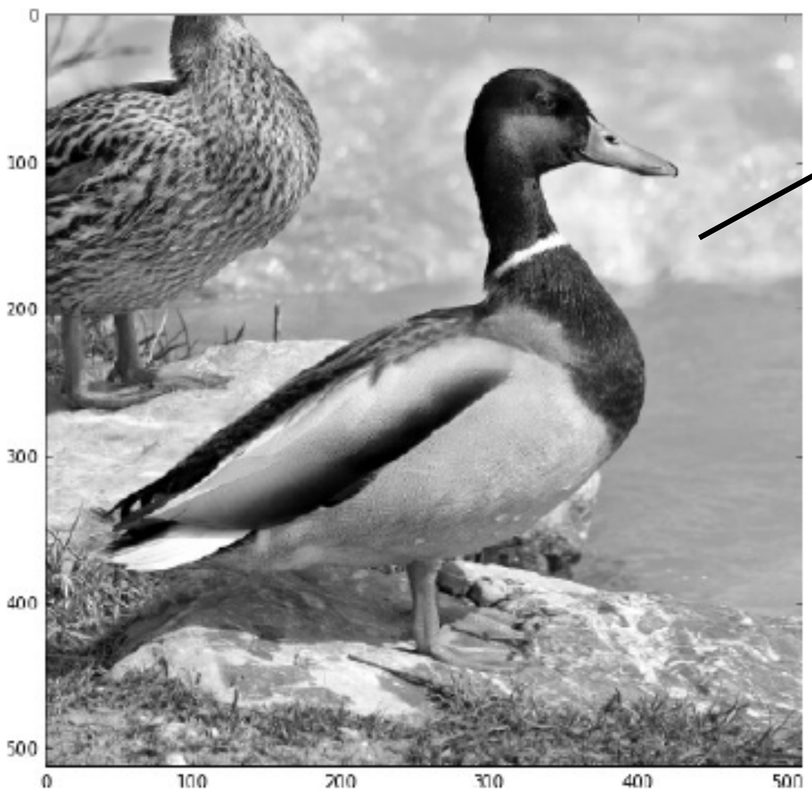
full-sky multi-line mappings pixel-by-pixel

1D, 2D, 3D treatment pixel-by-pixel

novel scientific life cycles

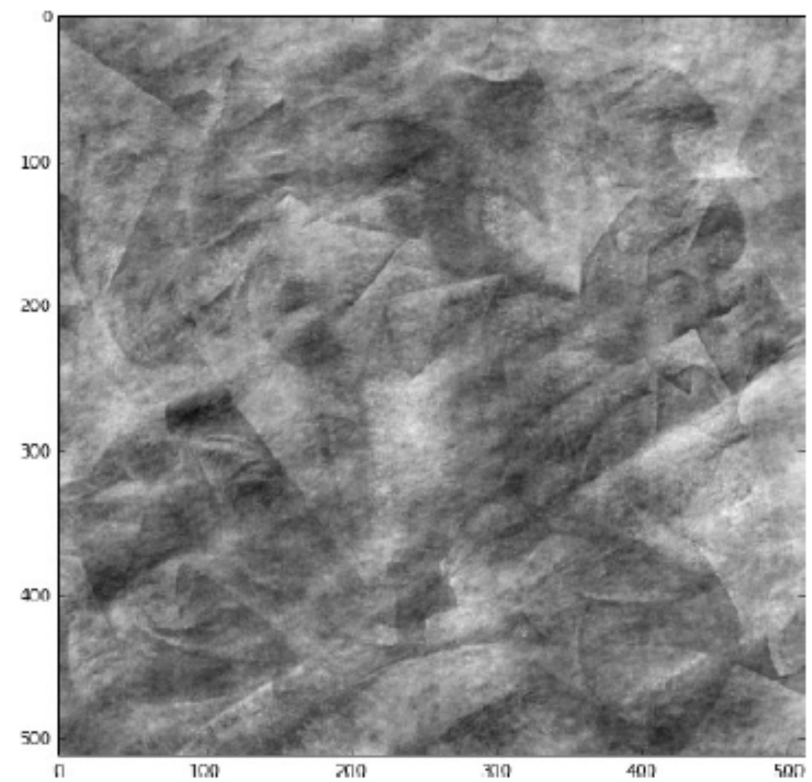
Moving on from summary statistics. Why deep learning?

The duck example of (Non-)Gaussianity



parameters
(size, color, background)

randomise phases



Credit: G. Bernardi

The duck: highly non-Gaussian

The Gaussian duck

Same 2D
power spectrum

- Picks up non-Gaussian information
- Representation learning

Applications:

1. Detect the duck (or galaxy, or signature)
2. **Inference** (what duck? what properties? what shapes?)

Learning the Universe: Setting sail to the first galaxies

3D imaging to track the history of the Universe

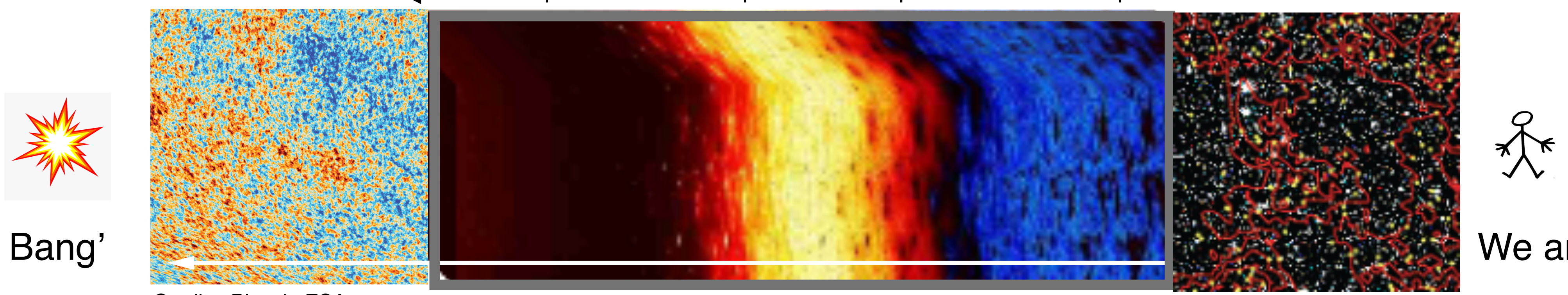
<u>cosmic dawn **CD**</u>
‘first stars and galaxies’

<u>reionisation **EoR**</u>
‘the last big phase transition’

(look-back time)
redshift z 20 15 10 7

<u>galaxy surveys</u>
e.g. LAEs at z~6

‘Big Bang’



Credits: Planck, ESA

Ouchi, Ono, Shibuya ‘20

We are here

<u>SKA - 21cm signal (neutral hydrogen)</u>

21-cm* 3D imaging**

* neutral hydrogen line

** blue = emission
red / yellow = absorption

<u>+ multi-line intensity mapping</u>
(21cm + Hα, Lyα, OIII, Hβ, CII, ..)

SKA EoR working group, synergies sub-group

Inference from 3D tomographic cubes

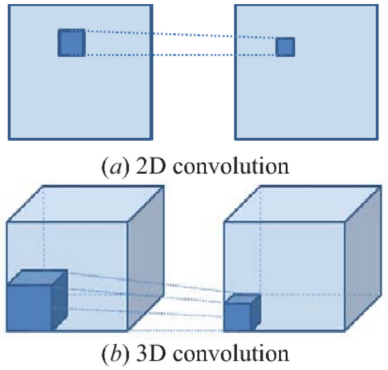
Direct likelihood-free inference from 3D tomographic cubes (21cm IM)

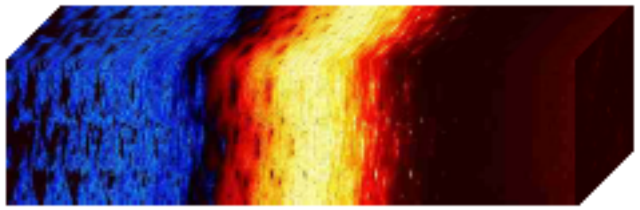
Various Options:

- Slicing and treatment with 2D CNN as image → ‘standard’ **2D CNN**, residual (skip connections) **ResNet**
- Time series (frequency) of co-eval images → **LSTM** network
- Full 3D convolution → 3D CNN:

[see e.g. Prelogovic+ arXiv: 2107.00018, Gillet+2019]

Moving from **2D to full 3D convolution**





Inference from 3D tomographic cubes

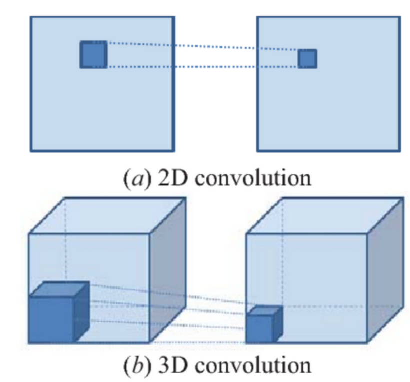
Direct likelihood-free inference from 3D tomographic cubes (21cm IM)

Various Options:

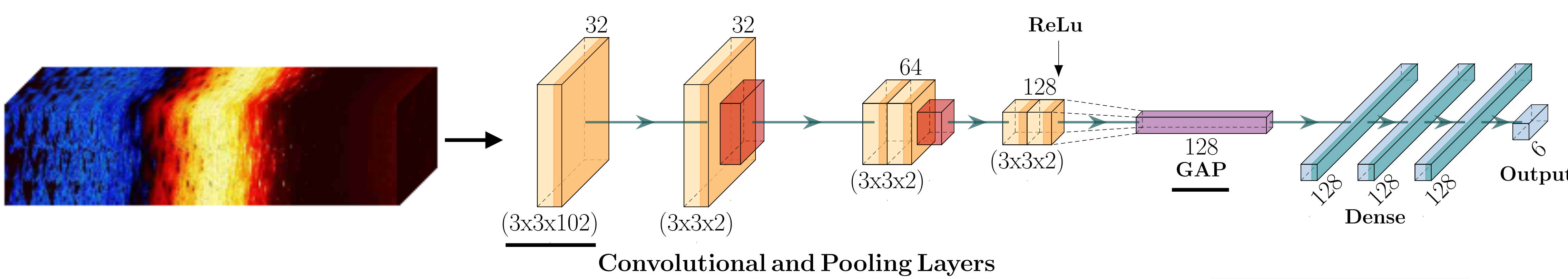
- Slicing and treatment with 2D CNN as image → 'standard' **2D CNN**, residual (skip connections) **ResNet**
- Time series (frequency) of co-eval images → **LSTM** network
- Full 3D convolution → 3D CNN:

[see e.g. Prelogovic+ arXiv: 2107.00018, Gillet+2019]

Moving from **2D to full 3D convolution**



Best-performing: simple Conv3D architecture



Database of ~5000 lightcones
140x140x2350 pix, 1.4 Mpc resolution
Training (single K80 GPU): ~20min/epoch, ~30 epochs

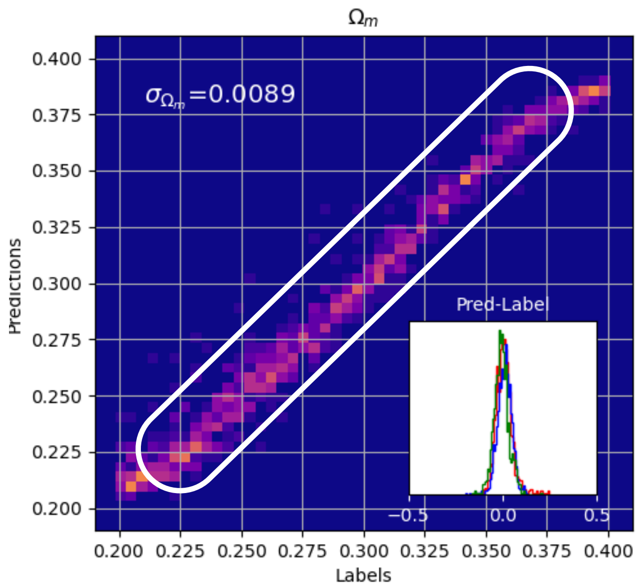
3D-21cmPIE-Net

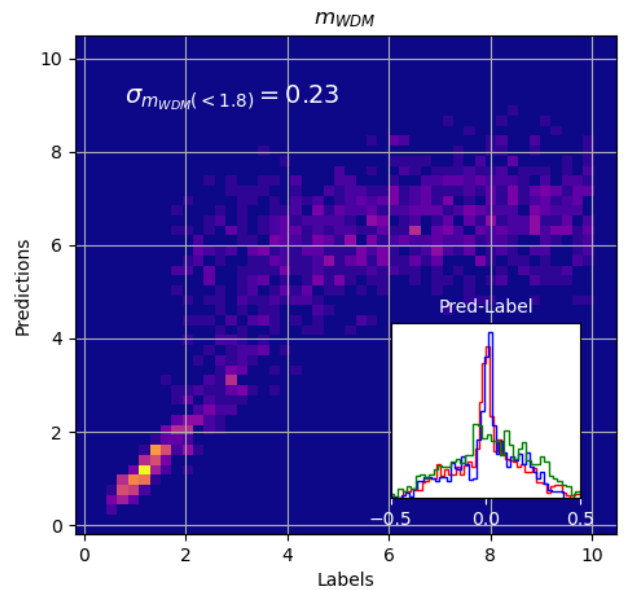
Neutsch, Heneka, Brüggen
MNRAS (2022)
arXiv:2201.07587

Inference from 3D tomographic cubes

Direct likelihood-free inference from 3D tomographic mock cubes (21cm IM)

$$(\Omega_{\rm m},\, \zeta,\, T_{\rm vir},\, L_{\rm X},\, E_0,\, m_{\rm WDM})$$

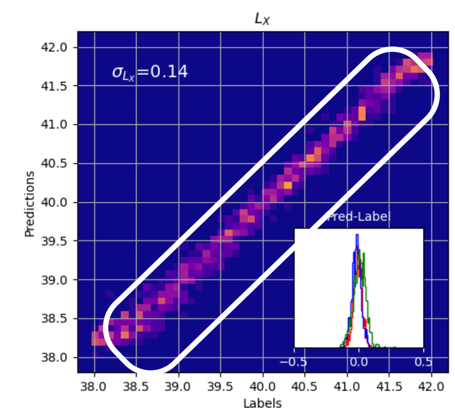


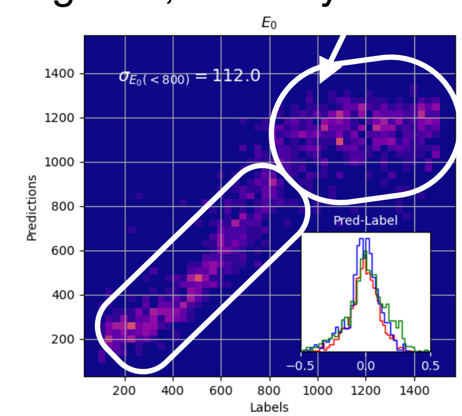


Cosmology

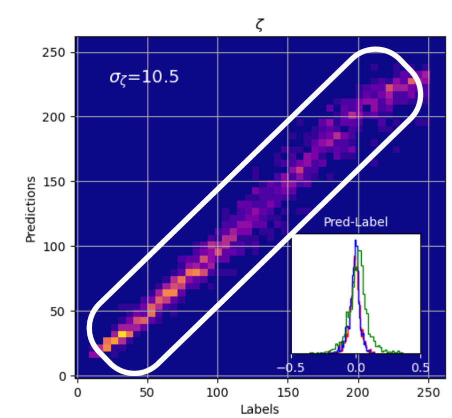
Cosmic Dawn

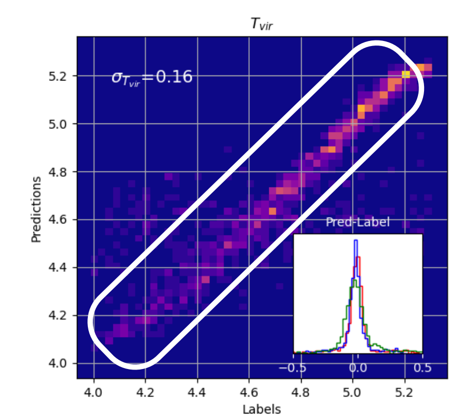
high E0, difficulty to escape





Reionisation





Directly constrain cosmology, CD & EoR astrophysics

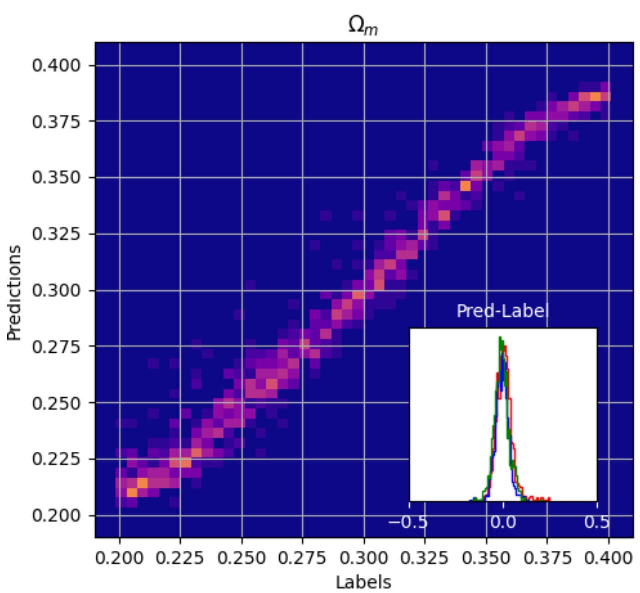
Mostly: small scatter, unbiased

arXiv:2201.07587

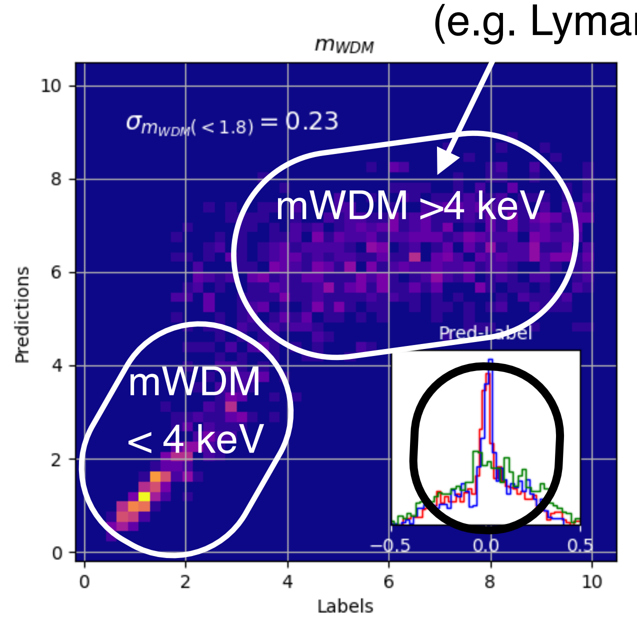
Inference from 3D tomographic cubes

Direct likelihood-free inference from 3D tomographic mock cubes (21cm IM)

$$(\Omega_{\rm m},\, \zeta,\, T_{\rm vir},\, L_{\rm X},\, E_0,\, m_{\rm WDM})$$

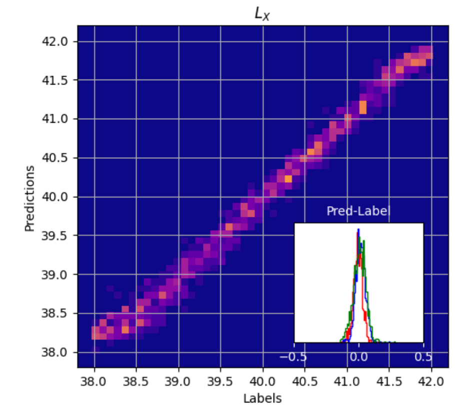


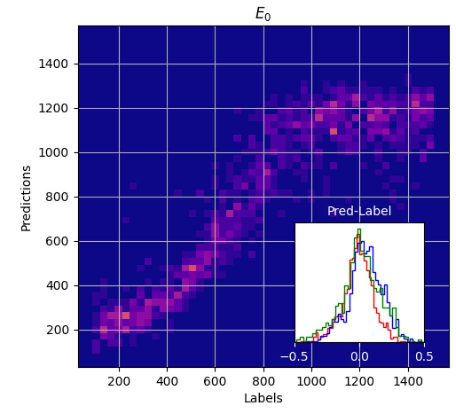
~results from other studies
(e.g. Lyman-forest)



Cosmology

Cosmic Dawn

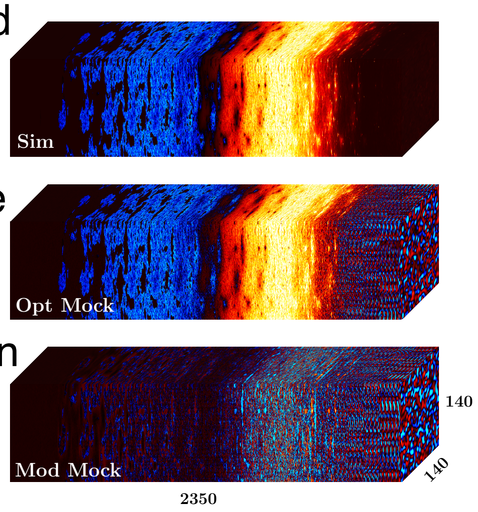




red

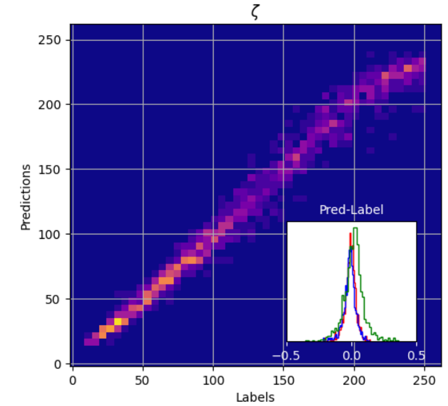
blue

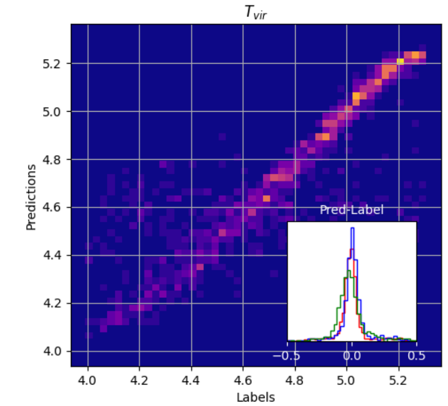
green



All, except mWDM, robust to increasing levels of foreground

Reionisation

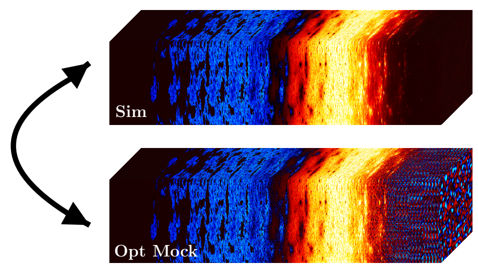




arXiv:2201.07587

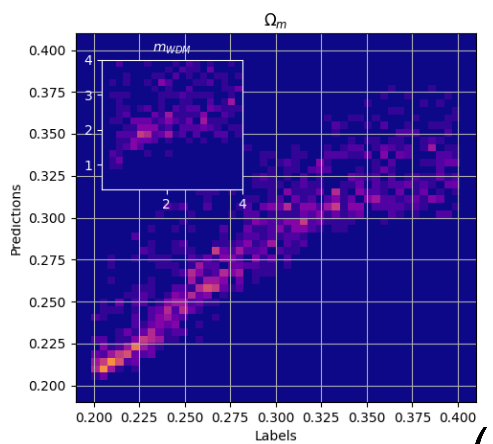
Testing robustness & interpretability

1) Transfer learning Sims & Mocks

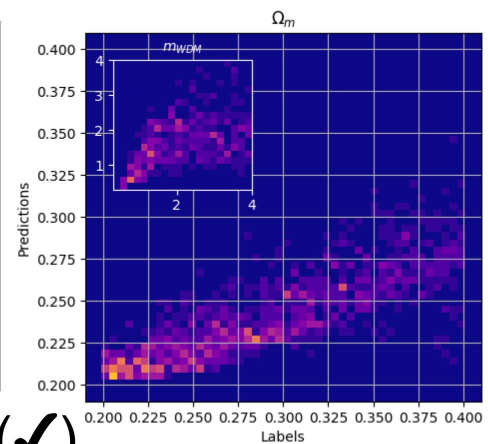


Sim -> Mock

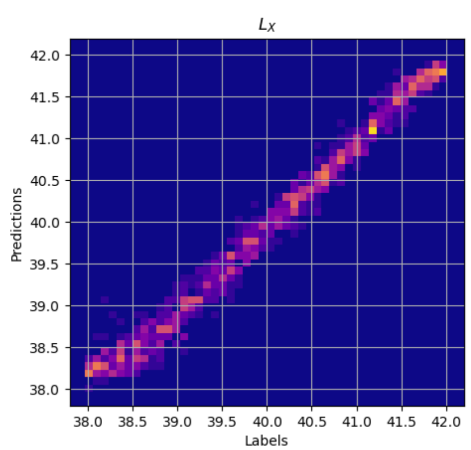
Mock -> Sim



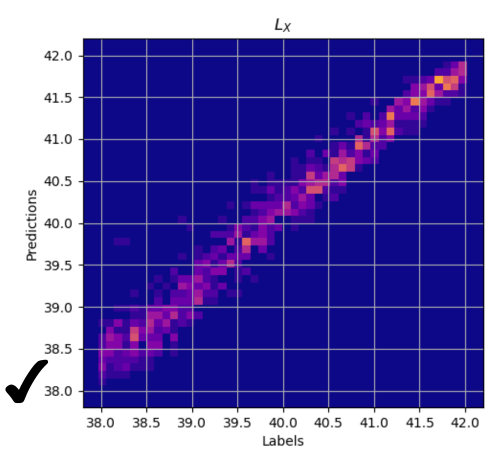
(✓)



✗



✓



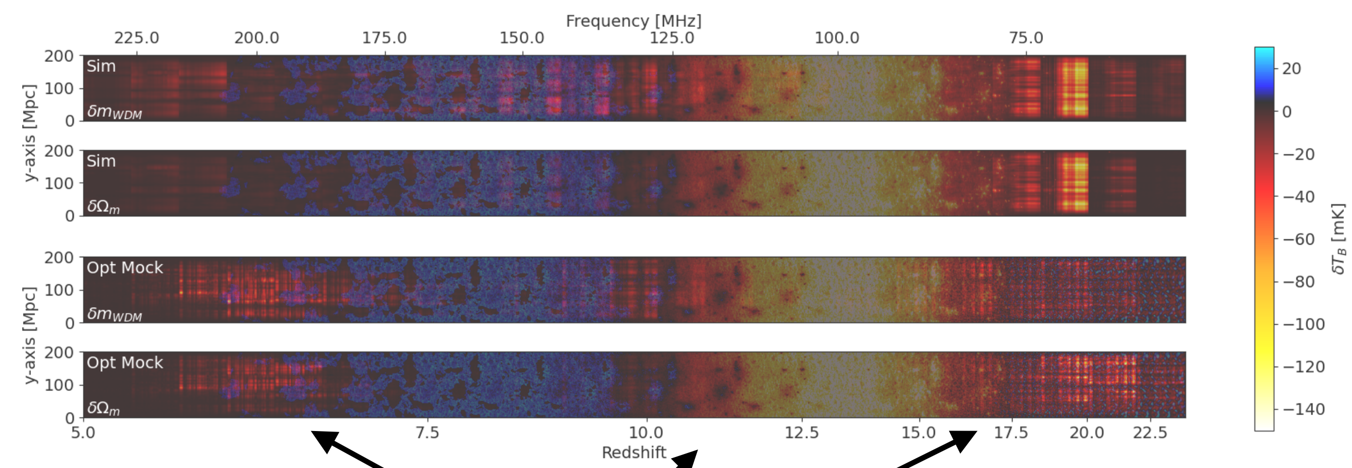
✓

✓ Robust to foregrounds & systematics

2) Gradient-based saliency maps

Sim

Mock



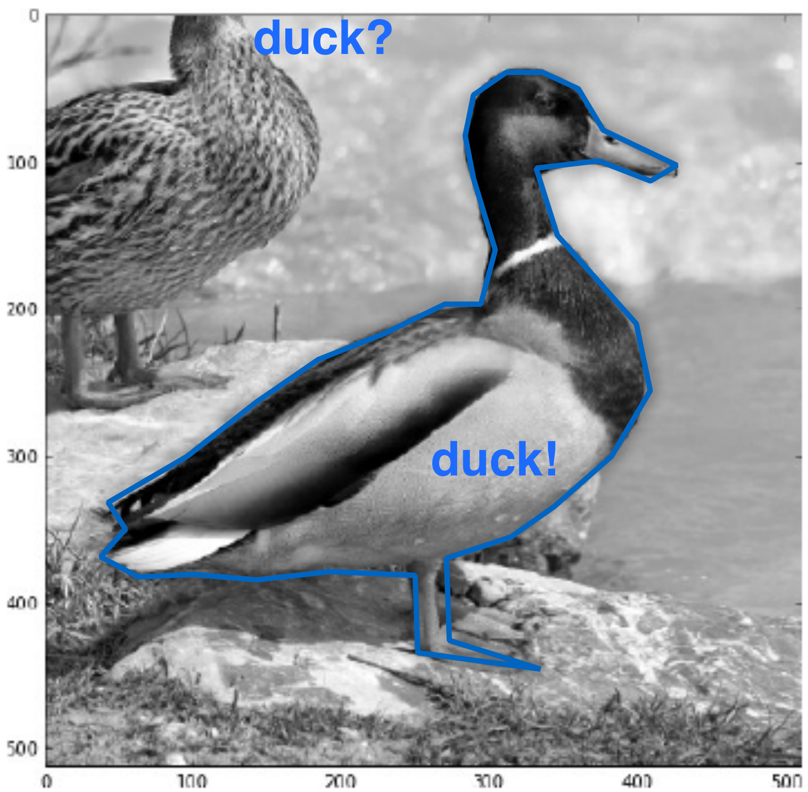
'transitions'

also seen for Fisher forecasts:
Heneka & Amendola 2018
Liu, Heneka, Amendola 2020

arXiv:2201.07587

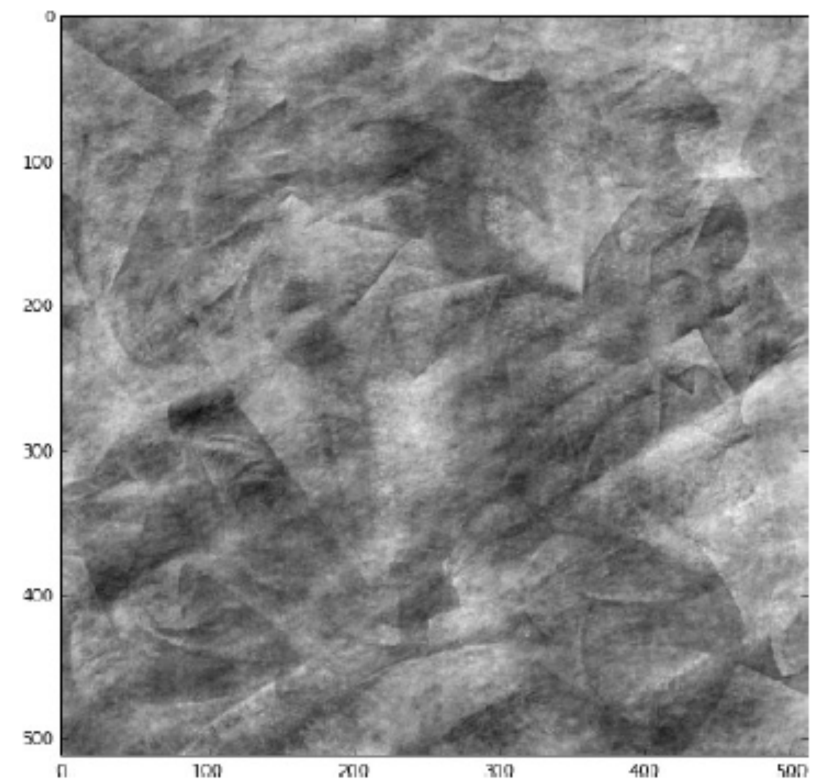
Moving on from summary statistics. Why deep learning?

The duck example of (Non-)Gaussianity



The duck: highly non-Gaussian

randomise phases



Credit: G. Bernardi

The Gaussian duck

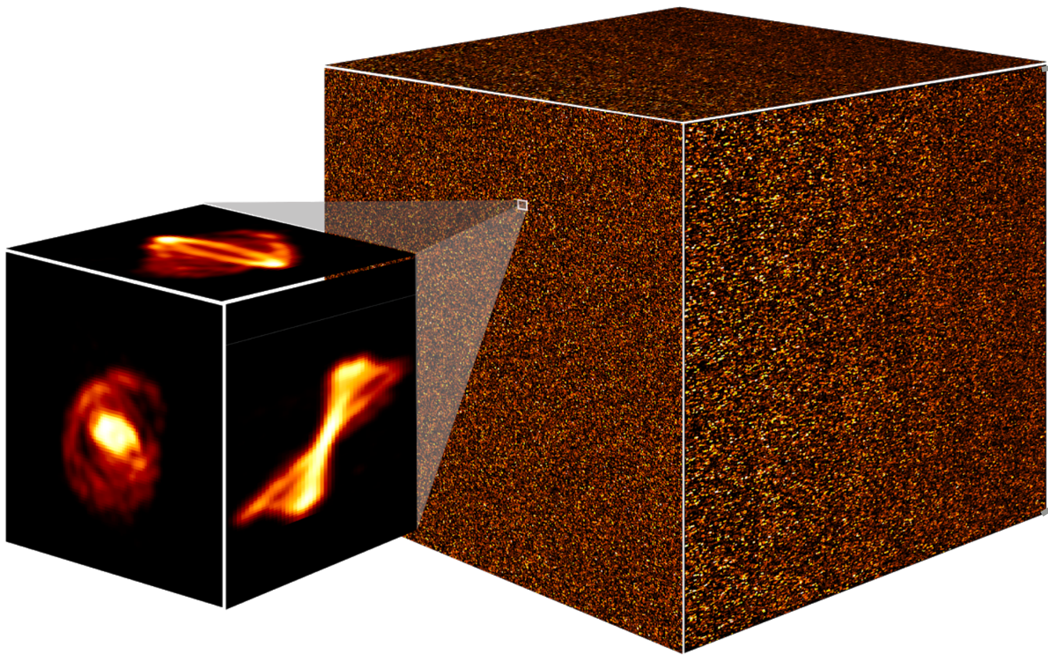
Same 2D
power spectrum

- Picks up non-Gaussian information
- Representation learning

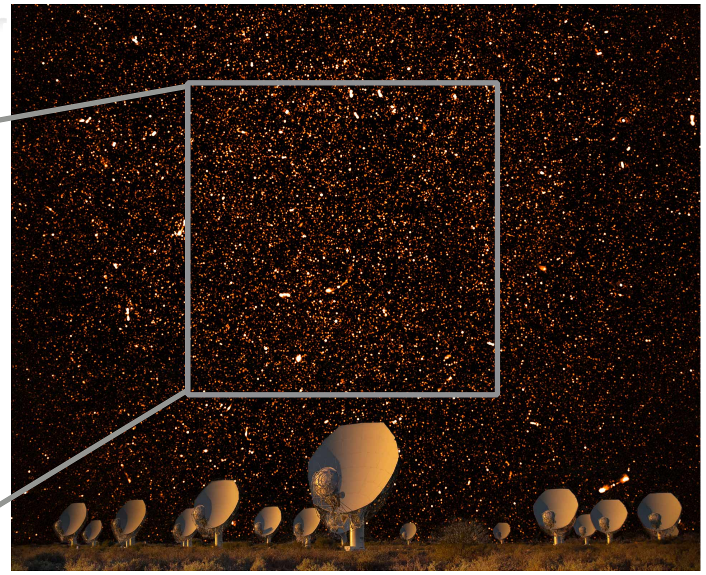
Applications:

1. **Detect** the duck (or galaxy, or signature)

Detection in 3D: SKA Science Data Challenge



Credit: https://sdc2.astronomers.skatelescope.org/sdc2-challenge/data



Composite MeerKAT dishes and observations.
Credit: South African Radio Astronomy Observatory (SARAO)

SKA -
The Square Kilometre Array

An international effort to build the world's largest radio telescope

Expected data rate in full operation: 1 TB/s

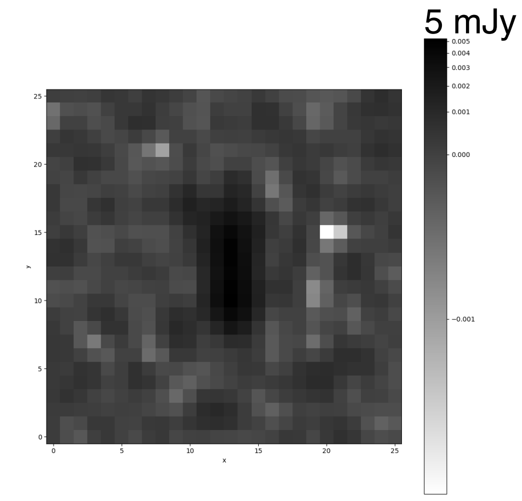
Key science goals include:
Galaxy Evolution, Reionisation, Cosmology, Astroparticles

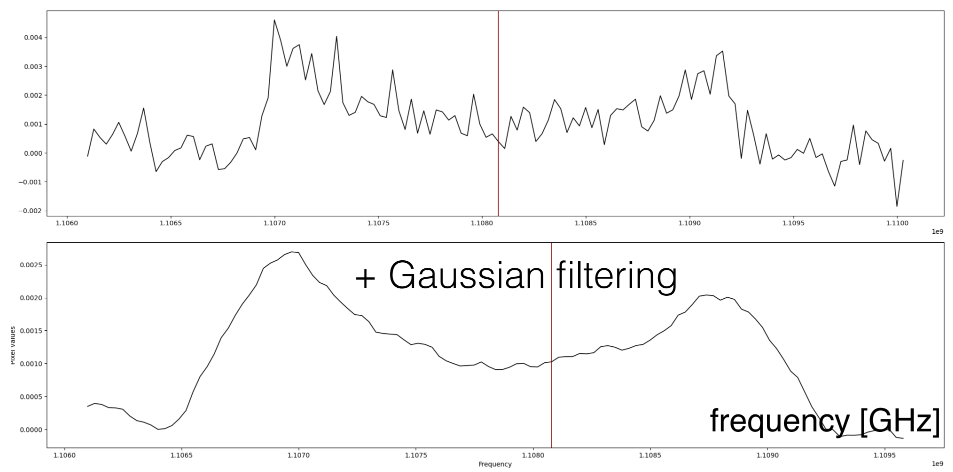
Goal is both source finding and characterisation & test for new SKAO Regional Data Centers

The challenging HI sources:

- **low S/N**
- **small spatial size**
- **systematics**

The brightest HI source





Detection in 3D: SKA Science Data Challenge

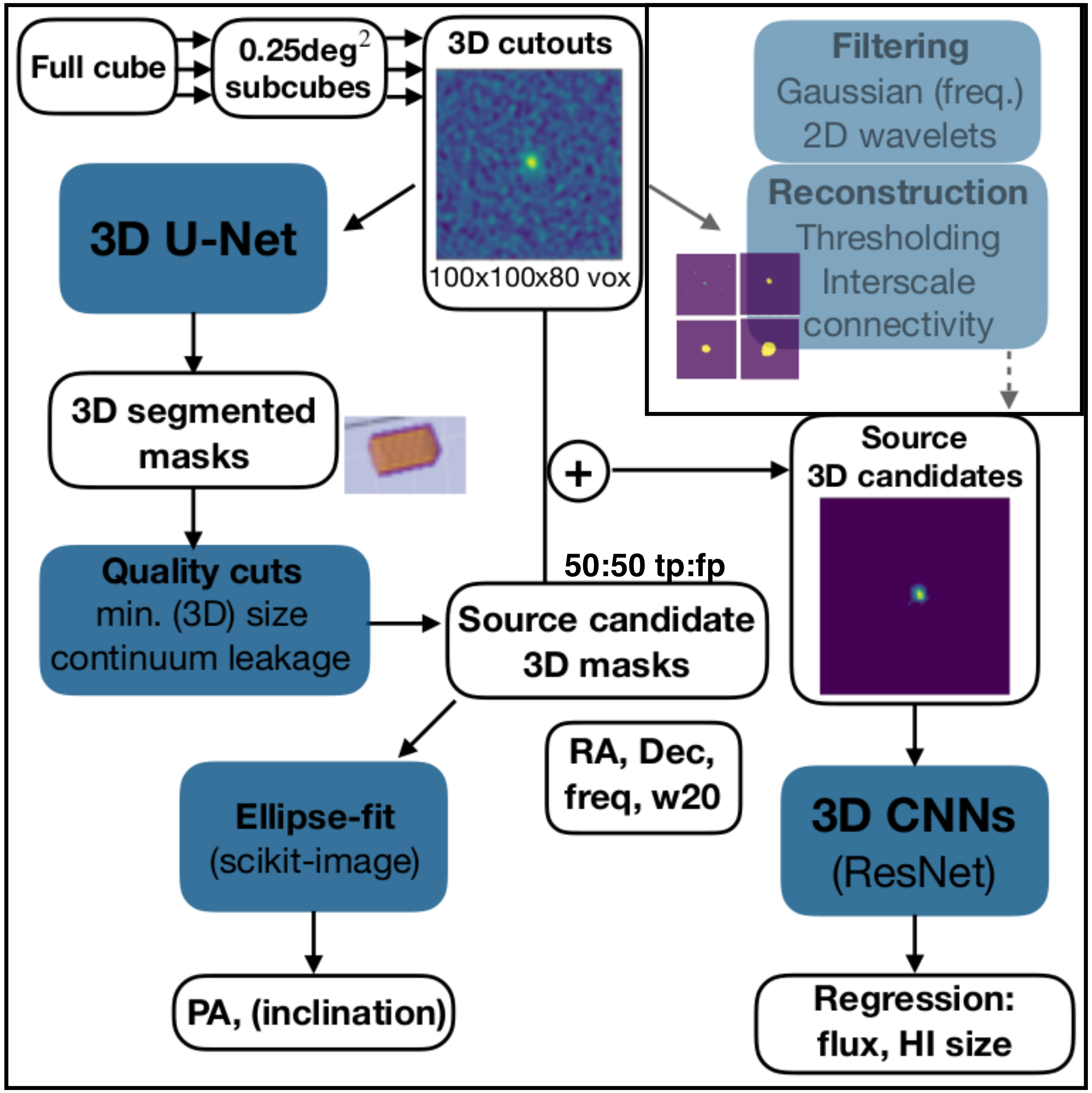
Machine learning and deep learning come together?

Team: Michelle delli Veneri, Andrew Soroka, Bernardo Fraga, Fedor Gobanov, Clecio de Bom, Alex Meshcheryakov

DL source detection & characterisation:

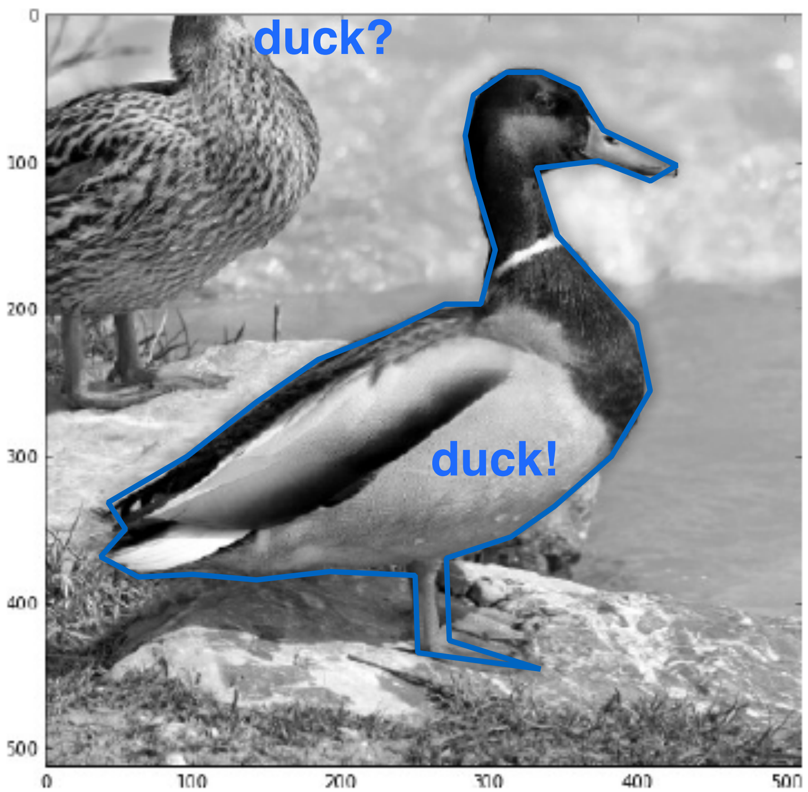
Pitfalls & Take-aways:

- Pre-processing, noise model(s)
- High sparsity
- **Choice of training set self-supervised**
- **Multi-step and/or ensemble decision**



Why deep learning?

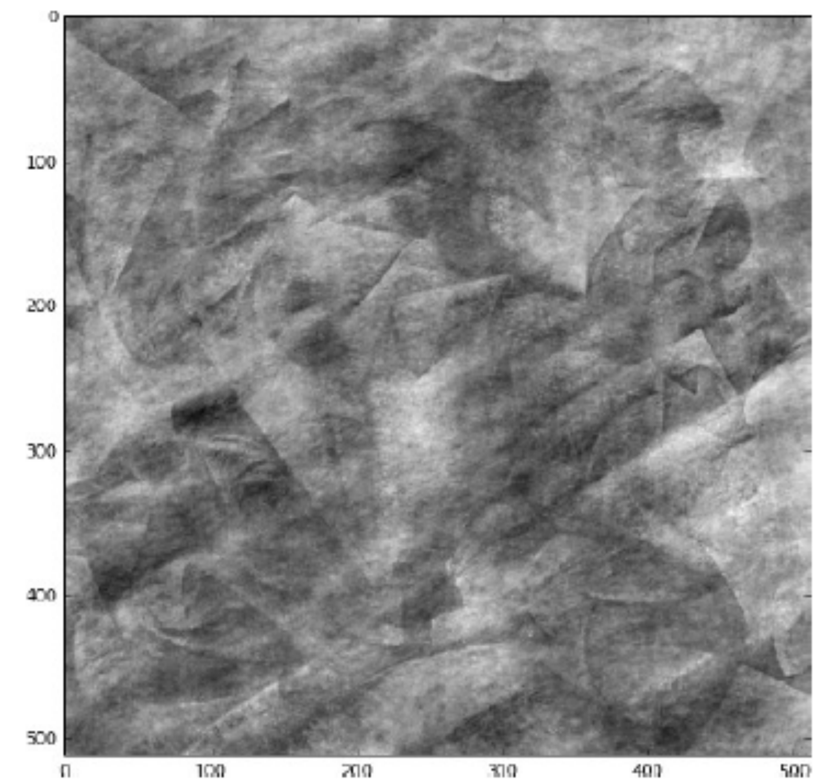
The duck example of (Non-)Gaussianity



The duck: highly non-Gaussian

randomise phases

Same 2D
power spectrum



Credit: G. Bernardi

The Gaussian duck

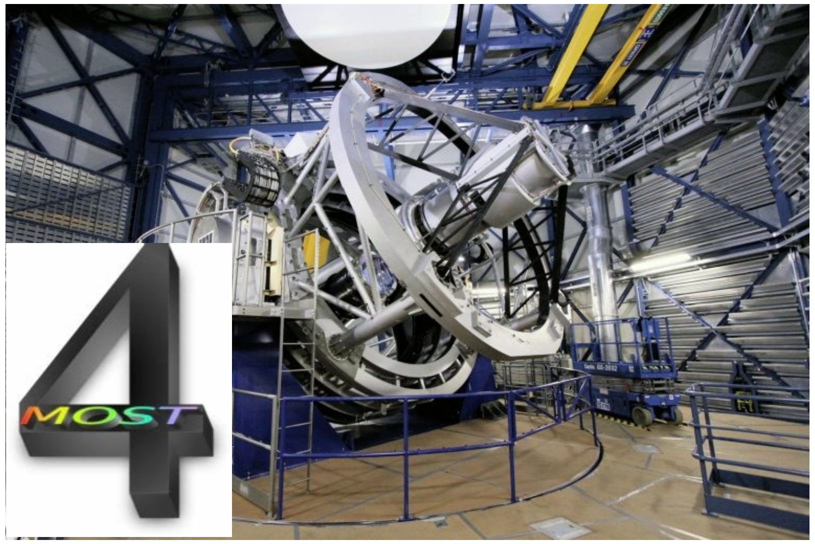
- Picks up non-Gaussian information
- Representation learning

Applications:

1. ~~Detect~~ the duck (or galaxy, or signature)

duck! classify.

Classification in 1D: Spectroscopy



Credit: ESO

https://www.4most.eu

Building a classifier for 4MOST - Classification IWG9

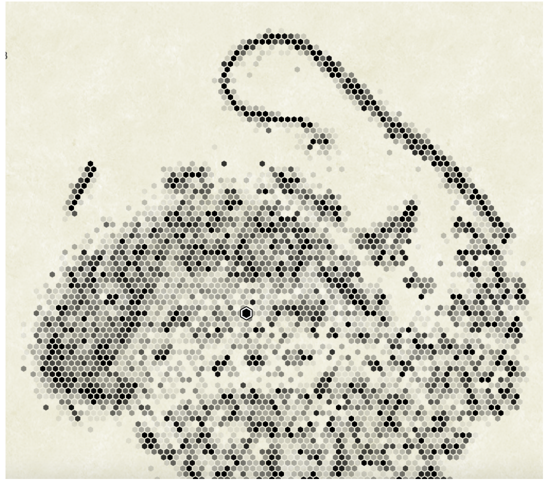
- 5-year survey
- wide-field, fibre-fed, optical spectroscopy
- on ESO's 4-m-class telescope VISTA
- 2.5-degree diameter field-of-view, 2436 fibres
- HRS R ≈ 18000 – 21000, LRS R ≈ 4000 – 7500
- 20mio. (LRS), 3mio. (HRS) sources

Goal: Data-driven classification layer between L1 and L2 pipelines

- **Basic target classification.**
  Classes: star, galaxy, AGN, quasar, .. → Probabilistic multiclassifier
  also: lowres vs. highres
  (low S/N vs. high S/N)

- **Galactic & extragalactic source classification.**
  Sub-classes matching L2 sub-pipeline
  galactic: FGKM, OBA, WD sub-pipelines
  supplement to metadata based decision
  → Probabilistic multiclassifier II
  (sub-classes)

- **Feedback on a) targets, b) 'unknown' class**
  Currently set-up:
  4MOST explorer t-SNE
  (Gregor Traven, Gal Matijevic)
  arXiv: 1612.02242
  → **a)** match with expectation
  b) clustering, dimensionality reduction

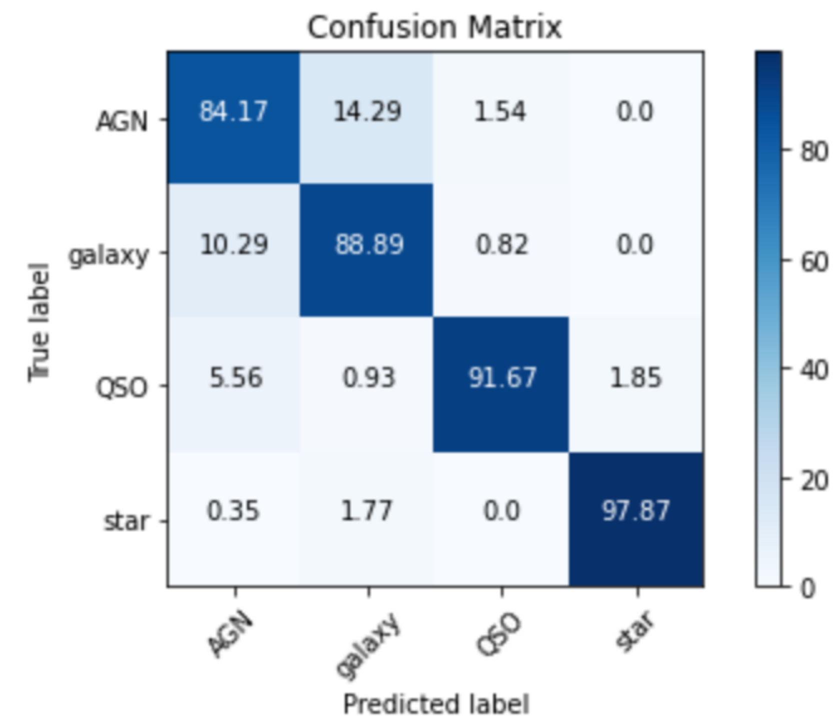


Classification in 1D: Spectroscopy

Benchmark tests with SDSS spectra

Convolutional network
Random Forest
Support Vector Machine
Logistic Regression
Gaussian Naive Bayes

2 convolutional + 2 dense
Currently: softmax-output
Ongoing: 'true' probablistic

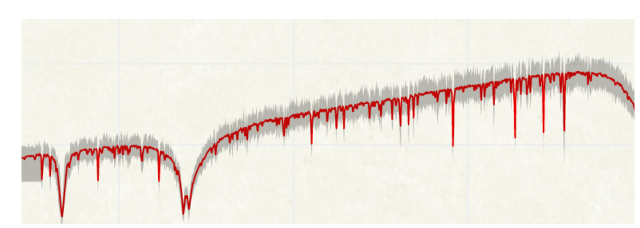


Advantage:
- complements template-based approaches
- once trained, extremely fast evaluation
- transfer of learning - training on archival data

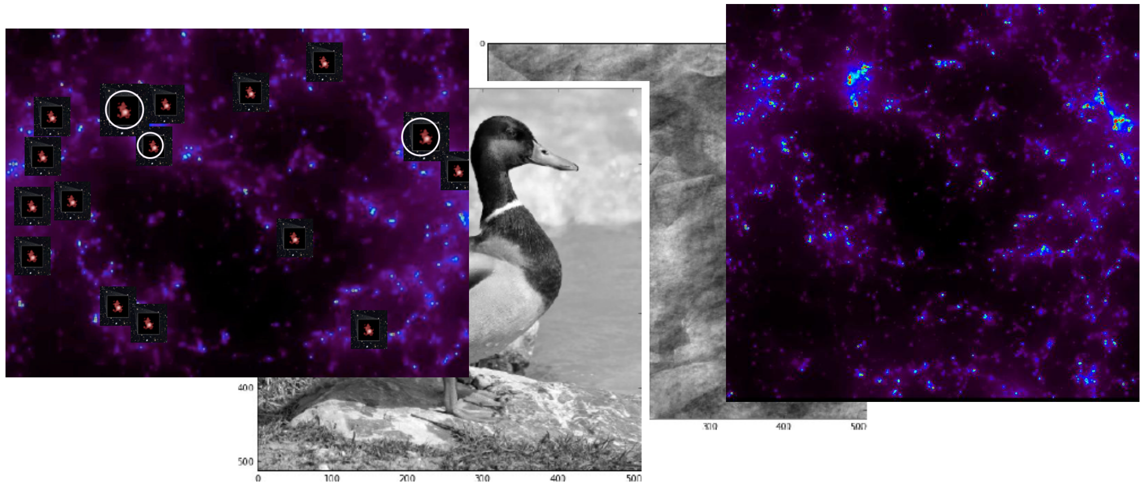
A supervised classifier (to start with) relies on 'good' training data..
Classification working group was re-activated in 2021,
under discussion: focus sub-group for training data

Training sets
(not connected
to daily data-flow,
but subject to updates)

Options:
SDSS, SDSS superset, GALAH, APOGEE, WEAVE (start '22)
templates / synthetic spectra, OpR simulated spectra



Learning the Universe from 3D to 1D



Main take-aways:

- Exploit techniques and methods such as intensity mapping and deep learning
- Avenue to jointly constrain astrophysics and cosmology (at Cosmic Dawn and Reionization)
- ..and more: exploit synergies and complementarity with e.g. galaxy surveys

Goals:
Map-based approach to astrophysics
Representation learning
Beyond summary 'let the net choose'
Automated, fast, effective
New signatures?

Ongoing:

- Public on Github: 3D-21cmPIE-Net
- Test of posterior estimate, probabilistic inference
- Test on data from e.g. SKA precursors

Interesting open questions:
When does ML improve over 'traditional' methods?
What needs to be done in terms of architecture, reproducibility, interpretability?
Training data vs. self-supervised?
Knowledge of uncertainties (epistemic & aleatoric)

Thank you!

caroline.heneka@uni-hamburg.de