AI for Astronomy in the SKA Era
Anna Scaife
Jodrell Bank Centre for Astrophysics

Swiss SKA Days 2021
What are the factors that motivate the use of AI for SKA?

- Large astrophysical populations;
- High data rates;
- Complex relational dependencies;
- Computational barriers to participation;
- Scientific (analysis);
- Scientific (operational);
- Operational (e.g. predictive maintenance, scheduling).
• Much of modern astrophysics is driven by population analyses;
• Populations need to be extracted from observational data;
• New discoveries need to be separated from known populations.
What are the issues that face the systematic use of AI for SKA?

- A paucity of labelled data for training - paradoxically, although we have too much data, we don't have enough;

- A clear understanding of the effect of biases introduced due to observational and intrinsic astrophysical selection effects in the training data;

- The quantitative statistical representation of outcomes from decisive AI applications that can be used in scientific analysis.
- Large archival databases but only small labelled data sets;
- Significant and variable class imbalances;
- Need carefully calibrated uncertainties on model outputs;
- Need for biases in model outputs to be quantitatively estimated.
Standard CNN

Group-equivariant CNN
Bayesian Neural Networks

- Allow us to recover posterior probability distributions;
- Provide a measure of how uncertain the model is in its prediction;
- Allow us to rank the importance of different weights within the model.

\[ ELBO = \mathbb{E}[\log p(D|w)] - KL(p(w|D)||p(w)) \]
Generative AI for SKA Data Challenges

Encoder

Decoder
Real or Fake?
Real or Fake?
Conclusions

• The digital revolution has vastly increased the sensitivity and utility of radio telescopes;

• New radio telescopes are big data machines;

• AI is essential for extracting scientific impact in a timely fashion –

• However, quantitative estimates of uncertainty and bias on AI outputs are essential for scientific use;

• We must not “automate out” the potential for DISCOVERY.