
Quantifying uncertainty in deep-learning classification of radio galaxies

Fiona Porter^{*1} and Anna Scaife¹

¹Jodrell Bank Centre for Astrophysics – United Kingdom

Abstract

Fanaroff-Riley (FR) galaxies, a type of radio-loud AGN, are among the sources that are expected to see a drastic increase in known population with advent of SKA-scale surveys, providing a wealth of information about AGN and their local environments but necessitating the use of robust automated classification to be identified and labelled. While efforts have been made to create such classifiers for FR galaxies, there remains the issue that even the best-trained models are uncertain regarding some labels, just as human classifiers are. In this talk, we discuss methods of quantifying this uncertainty.

To investigate the uncertainty properties of the FR population, we used a dataset of FR galaxies labelled both for their FR class and the confidence of a human classifier in the assigned class, and trained a CNN using LeNet architecture modified to include a final dropout layer to approximate a Bayesian posterior on predictions. This was used to extract two types of uncertainty measures from the CNN: aleatoric (irreducible uncertainty resulting from traits of a source that make its class inherently unclear) and epistemic (reducible uncertainty resulting from the model having a limited quantity of data to learn class information from).

Entropy, a measure of the model's consistency in classification of an image for multiple classification attempts using different dropout configurations (Gal, 2016), was used as a measure of aleatoric uncertainty (per Mukhoti et al 2020); an entropy of zero corresponded to a classification with a probability of 100% across all iterations. We then used a Gaussian Mixture Model (GMM) to fit the logits predictions for each image to clusters using expectation maximisation. In this instance, it was assumed that each FR class could be represented as a 2D Gaussian in latent space, with the datapoints found by the CNN representing a sample from these clusters, and with datapoints known to have low entropy weighted as more significant when fitting. From this GMM, a score was obtained that represented the probability of a datapoint belonging within the pair of clusters; this score was used as a measure of epistemic uncertainty (per Mukhoti et al 2020).

We find that the GMM fitted shows a region of overlap between clusters for the two FR classes, and that presence in this overlap region correlates well with high entropy. Additionally, while both confidently- and uncertainly-labelled sources are present in this region, the density of uncertainly-labelled sources is greater, and confidently-labelled sources commonly show unusual morphology (e.g. wide-angle tail) or possible contamination by background sources. A high GMM score corresponds to a source occupying a dense region of latent space where the model is confident in the accuracy of its predicted probability scores. Regions with a low GMM score may be separated by considering their entropy values; high-entropy sources with low GMM score are typically ambiguous, while low-entropy sources with low GMM score typically show very clear FR morphology.

*Speaker

Keywords: radio astronomy, deep learning, CNN, uncertainty, epistemic uncertainty, aleatoric uncertainty