NCRA-TIFR Project Proposal

Archit Sakhadeo
Rathin Desai
Shadab Shaikh
Shubhankar Deshpande

Mentor
Dr. Yogesh Wadadekar
Dr. C. H. Ishwar Chandra

1 Morphological Classes of Radio Galaxies

Radio galaxies with active nuclei can be distinguished based on their radio luminosity or brightness of their radio emissions in relation to their hosting environment. Some of the basic morphological classifications include point sources, extended sources i.e. sources with extended contours, double radio sources, jets, and lobes.

2 Problems faced with current classification

Currently Radio astronomers manually classify galaxies based on visual inspection of the images which, we are afraid, is slow, introduces uncertainities and something we believe is not a worthy activity for a radio astronomer to be engaged in.

Also, most algorithms have been used to classify just two object classes at a time. What we aim to achieve is build a robust model capable of handling more than 2 classes.

3 Objective

- Exploring rare galaxies by classifying them into a special category other than the ones mentioned in section yy.
- Effectively reduce the efforts put in by radio astronomers without affecting accuracy.

4 Approaches used for this problem

As far as we know, no method has been employed for this task.

5 Our Proposed method

5.1 Extracting sources from the FITS image using PyBDSM

By Shubhankar

5.2 Cropping each individual source to work on it individually

On the basis of RA(Right Ascension) and Dec (Declination) values generated from section 5.1, we then convert it to its corresponding pixel values in the original fits image. Based on these pixel values we crop a 10*10 patches of image with the source coordinates at the centre.

5.3 Preprocessing of the data

- Image Processing techniques

5.4 Analytical Method

Two broad steps that we plan to use:

- Statistical modelling of data to manually extract features. We plan to employ Scale-invariant feature transform (SIFT) algorithm to detect the features.
- Classification of the radio galaxies based on these extracted features. Possible approaches : Naive Bayes, SVM and Random Forests.

5.5 Empirical Approach

We plan to deploy a Convolutional Neural Network model for classification which reduces the manual feature engineering part, and has achieved significant successes in object recognition and image classification tasks. (Give References to papers)

6 How we predict it will solve the problem

7 Approximate timeline of completing the whole project

- 26th April to 11th May - literature survey
- Mid-August to October - working on basic prototype model individually by trying out multiple approaches
- October to November - choosing the approaches which work, and implementing them on all data, validation of the results.
- November to December - Refining the system, cleaning and commenting the code

8 Data used

TIFR GMRT Sky Survey data shall be used along with the data processed from GMRT cycle 20. We plan to use the data from other cycles once it has been run through our SPAM pipeline.