UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

DOCTORAL THESIS

An unsupervised search of non-thermal diffuse emission in extended sources

---

Author:
Mr. Siphiwe Anthony THWALA

Supervisor:
Dr. Geoff BECK

A thesis submitted in fulfillment of the requirements for the degree of Doctor of Philosophy

in the

Faculty of Science
School of Physics

January 13, 2022

Declaration of Authorship

I, Mr. Siphiwe Anthony THWALA, declare that this thesis titled, "An unsupervised search of non-thermal diffuse emission in extended sources" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed: 

Date : 13 January 2022

"I learned that every mortal will taste death, But only some will taste life."

- Rumi

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

Abstract

Faculty of Science
School of Physics

Doctor of Philosophy

An unsupervised search of non-thermal diffuse emission in extended sources

by Mr. Siphiwe Anthony THWALA

Low-surface brightness diffuse radio emission is a useful probe for studying the connection between non-thermal processes in radio sources and their environment, as well as gaining insights on galaxy formation. The presence of this emission in different astrophysical and cosmological signals is not well understood in the literature due to challenges in detection caused by the low-surface brightness of diffuse objects and their origins, particularly in relation to galaxy feedback and formation. We explored the utility of publicly available data to detect, characterise, and study diffuse emission in extended radio sources at the scales of radio galaxies and galaxy clusters. We performed multiple analyses on extended radio continuum sources found in large publicly available radio surveys. The analyses range from in-depth multi-frequency examinations of individual sources to novel approaches for grouping and classifying radio sources with unsupervised machine learning. The analysis of individual sources brought together a range of datasets and techniques to provide insight about their nature. To group and classify continuum radio sources at scale, a novel unsupervised machine learning approach was designed to combine self-organising maps with convolutional neural networks for automatically detecting and clustering similar sources in radio surveys. To the best of our knowledge, this is the first implementation of an architecture that allows for the training of a machine learning model using multi-frequency and multi-scale radio continuum data cubes, as input, to automatically detect and cluster similar sources in radio surveys. This comes at a time when radio astronomy is undergoing a transformation and data mining methods are critical for optimum scientific utilisation of data from telescopes like MeerKAT, JVLA, MWA, and LOFAR (as well as the upcoming SKA). The different works showcase the most interesting radio sources with diffuse emission observed in these investigations.

Acknowledgements

"If you want to go fast, go alone. If you want to go far, go together."

- African proverb

To my mother, words will never be enough to express my gratitude for the support you have gifted me through this journey and the sacrifices you have endured without hesitation. I deeply love and appreciate you! To my late gogo, it is as you said it will be so, ngiyabonga. To my late supervisor and mentor, thank you for gifting me the best environment for intellectual and personal growth. Your approach and discussions on working with the unknown continue to inspire my reasoning. To Dr. Geoff Beck, thank you for showing up for me in so many ways! You truly allowed me to be myself, were calm enough to keep me on track, and certainly supported me beyond expectations.

To my research colleagues, peers, and collaborators: Nthabiseng Khanye, Raesetje Sefala, Nebiha Shafi, Charles Takalana, Ashlynn Le Ray, Prof. Nukri Komin, Prof. Andrew Chen, Ayman Kudoda, Claudio Paulo, Justine Beck, Tanita Ramburuth-Hurt, Ahmed Ayad, Tebogo Mashile, Charissa Button, and Michael Sarkis, thank you for the countless conversations and shared insights; you all definitely brought the sanity checks and some more.

To my parents, family, and friends, THANK YOU for playing your role in making this vision happen. Everything has definitely added up to this conclusion of a chapter in my life. I appreciate you all!!!

This work was supported by the South African Radio Radio Astronomy Observatory (SARAO) through the South African Research Chairs Initiative (SARChI) under the grant No. 77948.

"One finger cannot lift a pebble."

- Malawian Proverb

Contents

List of Figures

List of Tables

List of Abbreviations

AGN	Active Galactic Nuclei
CMB	Cosmic Microwave Background
CS	Compton Scattering
DSS	ESO-Digitized Sky Surveys-2-Red
ESO	European Southern Observatory
FIRST	Faint Images of the Radio Sky at Twenty-cm
FSRQ	Flat-Spectrum Radio Quasar
GLEAM	GaLactic and Extragalactic All-sky MWA Survey
GMRT	Giant Metrewave Radio Telescope
IR	Infrared
ICS	Inverse Compton Scattering
Jy	Jansky
JVLA	Jansky Very Large Array
LOFAR	Low Frequency Array
LoTSS	LOFAR Two-metre Sky Survey
MWA	Murchison Widefield Array
NRAO	National Radio Astronomy Observatory
NVSS	NRAO VLA Sky Survey
RG	Radio galaxy
SDSS	Sloan Digital Sky Survey
SED	Spectral energy distribution
SKA	Square Kilometre Array
SSDC	Space Science Data Center
TGSS	TIFR GMRT Sky Survey
TIFR	Tata Institute of Fundamental Research
VLA	Very Large Array
VLASS	VLA Sky Survey
VLBA	Very Long Baseline Array
VLSS	VLA Low-Frequency Sky Survey
VLSSr	VLSS Redux
WENSS	Westerbork Northern Sky Survey
WISE	Wide-field Infrared Survey Explorer

Dedicated to
My beyond amazing mother, Busisiwe Thwala
and late greats, Bonginkosi Thwala,
Nomgqibelo Thwala and Sergio Colafrancesco

"The soul would have no rainbow if the eyes had no tears."

- Native American proverb

Chapter 1

Introduction

Radio astronomy, the study of astrophysical sources that emit radiation at radio wavelengths, is about to be one of the largest scale scientific data producers with the advent of the Square Kilometre Array (SKA; Dewdney et al. 2009; Weltman et al. 2020) project. The SKA has been projected to generate raw data on the order of an exabyte per day. Coupled with this expected data surge, is an increase in data volumes due to recent upgrades of existing telescopes, and new large modern telescopes coming online in other wavelengths (like MeerKAT (Jonas & MeerKAT Team, 2016), Low Frequency Array (LOFAR; van Haarlem et al. 2013), Australian SKA Pathfinder (ASKAP; Johnston et al. 2007; Hotan et al. 2020), Murchison Widefield Array (MWA; Beardsley et al. 2019) and Hydrogen Epoch of Reionization (HERA; DeBoer et al. 2017)). This then necessitates the use of analytics, especially data-intensive computing techniques, to better understand the statistical behaviour of astrophysical sources and facilitate the exploration of such large datasets.

For the reason that the morphology of radio galaxies is not well defined, the problem of detecting a radio galaxy is non-trivial and is known as source identification. This is the process of working out which components are associated together, and therefore identifying the radio galaxy. It is one of the open problems in radio astronomy, and a challenging one to address, since the evolution of a (radio) galaxy is largely shaped by the environment it coevolves in. This coevolution gives clues to the dynamics at play within the source and its interactions with the surrounding environment. The diversity in the observed features make it challenging to draw up simple rules for source identification (see Fig. 1.1 for some of the observed features when considering, the simplest case of, single sources). Although these may be easy for domain experts to pick up, the problem has not been solved at scale.

Machine learning (ML), the study of algorithms and statistical models that extract patterns and learn representations from raw data, is unlocking insights in domains that do not depend on or have hard-coded[1] knowledge. Using the fact that radio continuum images are tracers of dynamical states of

[1]Programmes/codes which have rules and/or parameters (that are explicitly given) in such a way that an alteration of a rule and/or parameter requires modifying the programme.

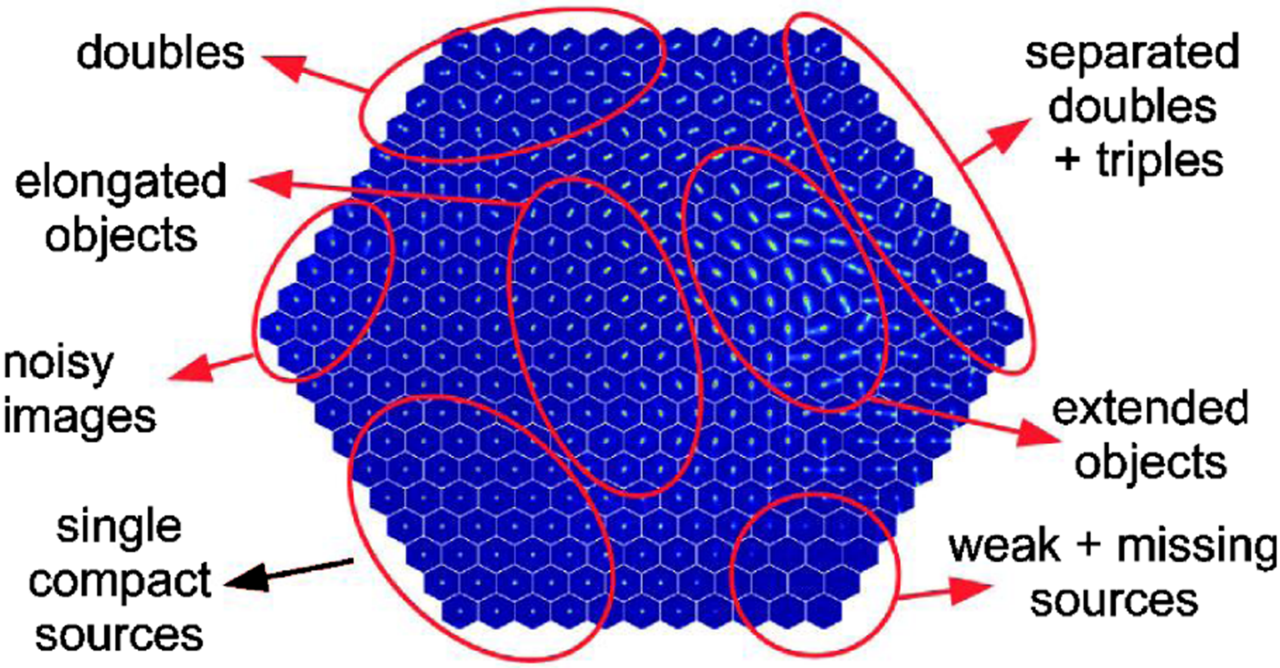


FIGURE 1.1: A view of the classified features from a self-organizing feature map (SOM) with morphological classes being marked and labelled in red (Polsterer, Gieseke & Igel, 2015).

radio sources at multi-frequencies; ML techniques provide a unique opportunity to find statistical tracers of these environments, and aid astronomers with source extraction (and possibly, characterization) on a greater scale than what can be achieved through visual inspection and heuristics-based methods. The observed extended features in radio sources are largely driven by non-thermal emission processes (described in detail in Chapter 1). We can extract these unique extended features and use them as the representations we want the ML algorithms to learn. Our particular focus will be on low-mid frequency radio continuum observations of radio galaxies, where these features are more conspicuous.

Low-frequency radio continuum observations of (radio) galaxies are of particular importance since the associated synchrotron emission probes the aged and low-energy electrons, which are less affected by energy losses, and therefore, can propagate further away from their site of origin (assuming that the diffusion coefficient does not vary with particle energy; Mulcahy et al. 2014). Depending on the magnetic field strength and spatial extension, one would expect to see a large radio synchrotron emission in extended sources (Mulcahy et al. 2014; see also the case of the $\sim$ 700 kpc remnant radio galaxy J1828+49 studied with the Low-Frequency Array (LOFAR) van Haarlem et al. 2013; Brienza et al. 2016).

This is in fact the case for radio galaxies with jets or outflows: AGN-powered radio galaxies have relativistic jets extended and collimated for tens to hundreds of kpc. Outflows are observed to be populated by lower energy electrons which produce irregularly shaped radio lobes that typically extend for smaller distances from the galaxy core.

In order to search for and study new extended emissions around radio galaxies, we will use publicly available data from the Faint Images of the Radio Sky at Twenty-cm (FIRST; Becker et al. 1995), GaLactic and Extragalactic All-sky MWA (GLEAM; Hurley-Walker et al. 2017), LOFAR Two-metre Sky Survey (LoTSS; Shimwell et al. 2017), NRAO Very Large Array (VLA) Sky Survey (NVSS; Condon et al. 1998), TIFR GMRT Sky Survey (TGSS; Intema et al. 2017), and VLA Sky Survey (VLASS; Lacy et al. 2020), and VLA Low-frequency Sky Survey Redux (VLSSr; Lane et al. 2014) surveys.

In this study, we propose the usage of ML techniques to search and analyse the diffuse emission from extended radio sources in deep radio surveys at scale. Our study will use the improved sensitivity of the available radio surveys (in particular LoTSS) and their ability to better resolve radio galaxies for studies of magnetic fields, cosmic ray diffusion and particle acceleration mechanisms in their jets/lobes.

Chapter 2

Background Theory

Astrophysical sources produce observable emission across the continuous electromagnetic spectrum, from high-energy charged particles (observed at gamma and X-ray wavelengths) to lower-energy charged particles (observed at radio wavelengths). The wide variation in particle energies between these observable wavelength limits, allows us to observe different types of electromagnetic radiation (produced by charged particles emitting radiation at an observed frequency) which can be grouped into different wavelength/ frequency bands (see Fig. 2.1).

The origin of the observed emission is driven by different aspects of structure formation and evolution, of the Large Scale Structures (LSS), in the universe (Colafrancesco, 2010). Contained within the LSS are tracers of the history and composition of cosmic structures, with galaxy clusters being a representative system of the distribution of the LSS in the universe (Colafrancesco, 2010). Galaxy clusters - the largest bound structures in the universe - accumulate (over cosmic time) a plethora of matter, such as galaxies, hot thermal baryonic plasma, non-thermal and relativistic plasma, cosmic rays (CRs), Dark Matter (DM), magnetic fields and Black Holes (BHs).

To probe the underlying matter composition of an astrophysical source (using the nature of the observed emission; and its role in structure formation and evolution), the dominant emission mechanism signals within each wavelength band are characterised. The characteristic emission signals are then used to model possible interactions that best describe the observed (unique) phenomena in each wavelength band. Some of the possible astrophysical emission mechanisms are from radiating particles emitting gamma-rays (π^0 decay, relativistic bremsstrahlung and ICS), X-rays (non-thermal bremsstrahlung and ICS), microwaves (Sunyaev-Zel'dovich effect from the ICS of CMB photons) and radio emission (synchrotron emission; Colafrancesco 2010). These variations necessitate specialised probes to study the physics and physical dynamics driving the observed emission features, in each wavelength band. Thus, forming one of the main motivations behind multi-wavelength astrophysics, where a collection of telescopes (located in space and on the surface of Earth, shown in Fig. 2.2) are used to probe distinct features of astronomical objects across the observable wavelengths.

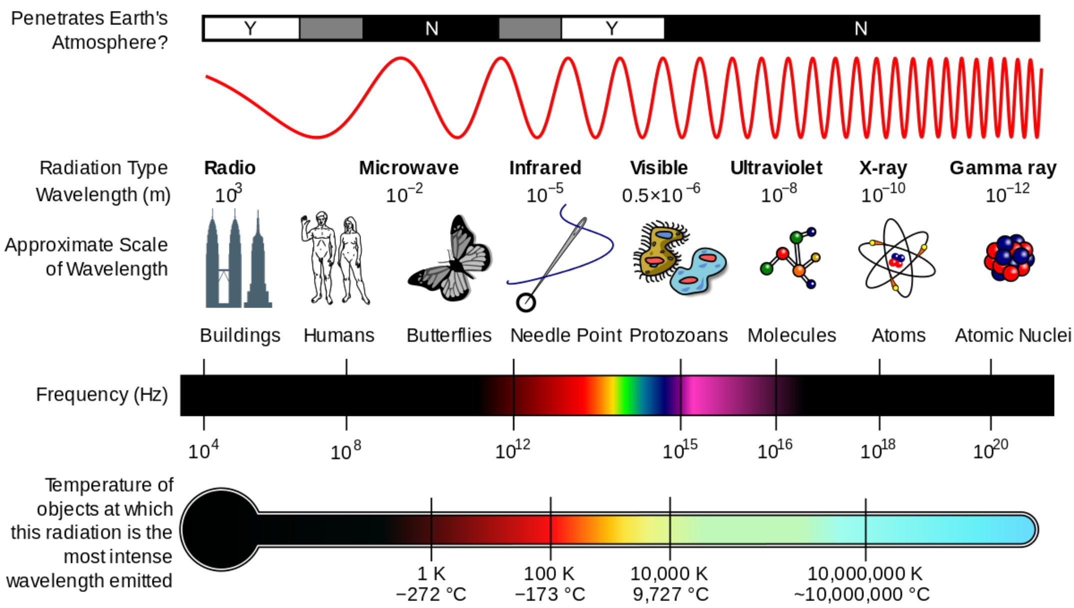


FIGURE 2.1: An illustration of the large variance in frequency found in the electromagnetic spectrum, along with intuitive scales of the different wavelength bands, the temperature in these bands and the possibility of observing these signal from Earth's surface (Inductiveload, 2007).

The work proposed in this thesis is centred around radio astronomy as the main observational probe. We will use data from other wavelengths to complement the characterisation of the extended radio sources that are discussed/ discovered within the different works, contained in the thesis chapters.

The observables that we probe are produced when the energy from a radio source is turned into radiation (radio telescopes measure the intensity of this radiation), and the three main emission mechanisms we can observe in radio astronomy are; spectral line emission, thermal and non-thermal emission processes. We will now look at radio galaxies, their morphology and a few key features characterising the different emission mechanisms seen in these sources.

2.1 Radio galaxies

Galaxies are systems of stars, dust and gas that are gravitationally bound and dynamically interacting with each other through the exchange of matter and energy. These dynamic interactions produce structures that determine the energetics of photons that we observe from these systems. Most galaxies produce some radio emission (from the process of star formation; Thompson,

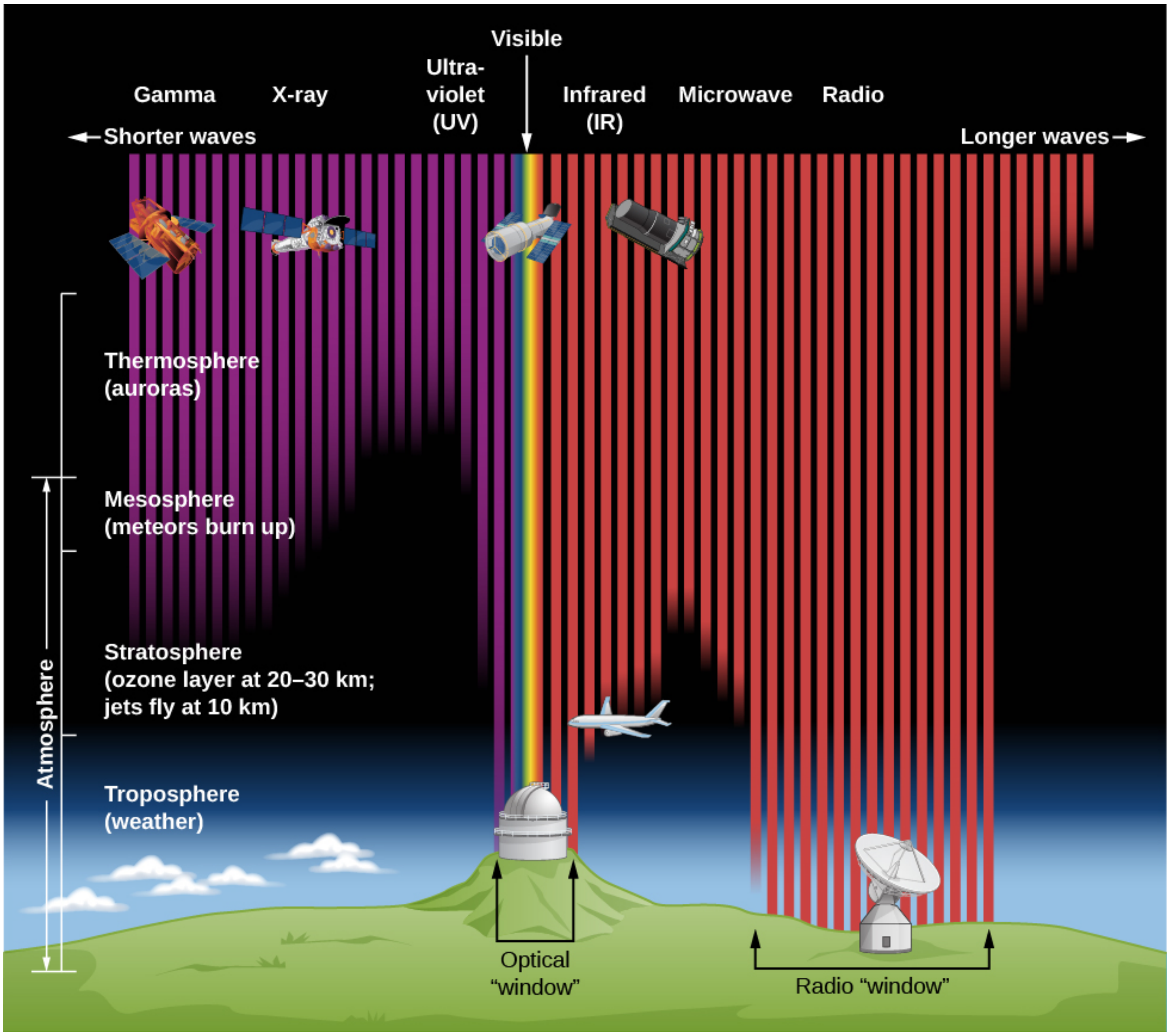
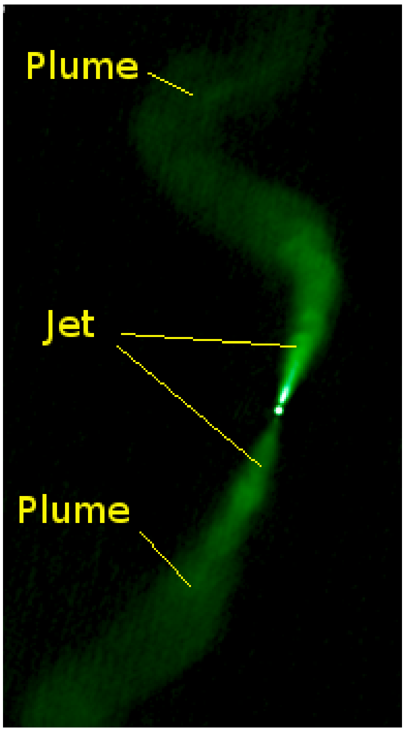
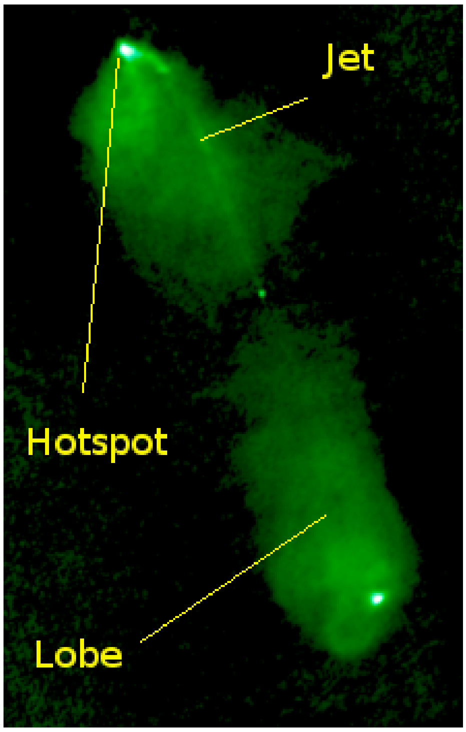


FIGURE 2.2: An illustration of the different telescopes used as probes in astrophysics and the "windows" (wavelength regions of the electromagnetic spectrum) we can probe from the surface of Earth (LibreTexts, 2016).



(A) Radio galaxy 3C31 (FR-I type galaxy) (Mhardcastle, 2006)



(B) Radio galaxy 3C98 (FR-II type galaxy) (Mhardcastle, 2006)

FIGURE 2.3: A side by side view of the two radio morphology classes devised by Fanaroff & Riley based on the large-scale radio emission of extragalactic radio sources.

Quataert, & Murray 2009), leading to star-forming regions. For radio galaxies, this radio emission is produced by non-thermal emission processes associated with a confined area in the centre of a galaxy, known as Active Galactic Nuclei (AGN), that releases a substantial amount of energy throughout the electromagnetic spectrum (Padovani et al., 2017). This emission is bright compared to the optical/infrared emission (radio galaxies are often classified using the radio-to-optical or radio-to-infrared luminosity ratio; Padovani et al. 2017).

The morphology of radio sources can be broken down into different constituents (as seen in Fig. 2.3): namely, the core, normally the source of the observed emission and powered by a black hole; jets, collimated regions where lighter charged relativistic particles, electrons and positrons, are accelerated by the presence of a magnetic field; lobes, extended "elliptical" structures dominated by low surface brightness diffuse emission; plumes, similar to lobes except the approximated shape is not elliptical and they normally show no signs of hotspots; and hotspots, termination points for jets and a possible site for shock waves resulting from interactions between the jet and the intracluster medium (Blandford, Meier & Readhead, 2019).

In 1974 Fanaroff & Riley (Fanaroff & Riley, 1974) devised a classification scheme for extragalactic radio sources based on their morphology. In their presented work, they suggested a division of extragalactic radio sources based on the luminosity of the sources and noticed that sources with radio luminosities below and above $\approx 10^{25}$ W Hz^{-1} at a frequency of 178 MHz, have distinct morphologies. They then proposed grouping extragalactic radio sources into two classes based on the morphology of the large-scale radio emission. Sources which lie below $\approx 10^{25}$ W Hz^{-1} are observed to have no hotspots, are less energetic and have wide differences in morphology of the extended emission (see plumes in Fig. 2.3a). These sources were classified as Class I and are now commonly know as FR-I type galaxies. Sources which lie above $\approx 10^{25}$ W Hz^{-1} are observed to have hotspots, are more energetic and characterised by having a double lobed morphology with their extended emission terminated by hotspots (see Fig. 2.3b). These sources were classified as Class II and are now commonly know as FR-II type galaxies.

This gives us a macroscopic view of extended radio sources. We need to understand the underlying microscopic mechanisms that shape these morphologies in order to better understand how the observed large-scale radio emission emerges. We now shift our attention to the key features characterising the different emission mechanisms we observe in radio sources.

2.2 Spectral line emission

Spectral lines are defined to be narrow ($\delta\nu \ll \nu$) emission or absorption features in the spectra of gaseous and ionized sources (Ransom, 2016). The background to these spectral features is produced by the recombination of electrons with atomic nuclei to form atoms that are usually in an excited state. As atoms cascade down to the ground state, they emit a photon with an energy equal to the difference between the initial and final atomic state, and such spectral lines are called recombination lines.

The emitted photon then travels through different clouds of gas. Because of the difference in the energetics, the gas clouds imprint different absorption and emission features along the line of sight of the photon. This information can then be used to probe gas clouds that lie along the line of sight of the observed spectra.

Examples of radio spectral lines, which are intrinsically quantum phenomena, include recombination lines of ionized hydrogen (Fig. 2.4) and heavier atoms, astrophysical masers, 21 cm hyperfine line of interstellar H_I, and rotational lines of polar molecules such as carbon monoxide (CO) (Ransom, 2016). In our characterisation, we will use the SDSS spectra to obtain information about the galaxy type and redshift of the extended radio sources, provided the SDSS optical data catalogue has spectra on that particular source.

2.3 Continuum emission

From the above discussion, we have seen that spectral line emissions are a result of atomic processes that have specific quantised energies caused by differences in the initial and final atomic states of the cascading atom, rotational and vibrational transitions within molecules (e.g., CO in the sub-mm/radio), and most relevantly, spin flip transitions in neutral atoms (e.g., 21cm emission in Hydrogen).

Continuum emission on the other hand, results from the acceleration of charged particles. Since continuum emission results from processes where the energy exchange is not quantised, the emitted photons may have a continuous energy distribution (Griem, 1997). We now give a brief introduction to the two processes that result in continuum emission: thermal and non-thermal emission processes.

2.3.1 Thermal emission processes

Astronomical objects with a relativistic thermal plasma are optically thick, meaning that subatomic particles are readily absorbed or scattered within the plasma, and have a relativistic Maxwellian electron distribution (Wardzinski & Zdziarski, 2000). An object is said to be in local thermodynamic equilibrium (LTE), if the emitting/absorbing object is in thermal equilibrium at a well-defined temperature T (Ransom, 2016). Thermal emission is produced by a source whose emitting particles are in LTE; otherwise non-thermal emission is produced (Ransom, 2016). Because of this property thermal emission processes are dependent only on the temperature of the emitting object.

The key characteristic of sources with emitting particles that are in LTE, is that they emit energy across the entire electromagnetic spectrum. Objects at higher temperatures emit more radiation and peak at shorter wavelengths, as shown in Fig. 2.5. In this study, we do not take an in depth look at thermal emission processes since, we are interested in studying low and mid-frequency radio continuum observations of extended (radio) galaxies. Radio continuum observations at these frequencies are sensitive to the optically thin non-thermal plasma, which follows a power-law electron distribution.

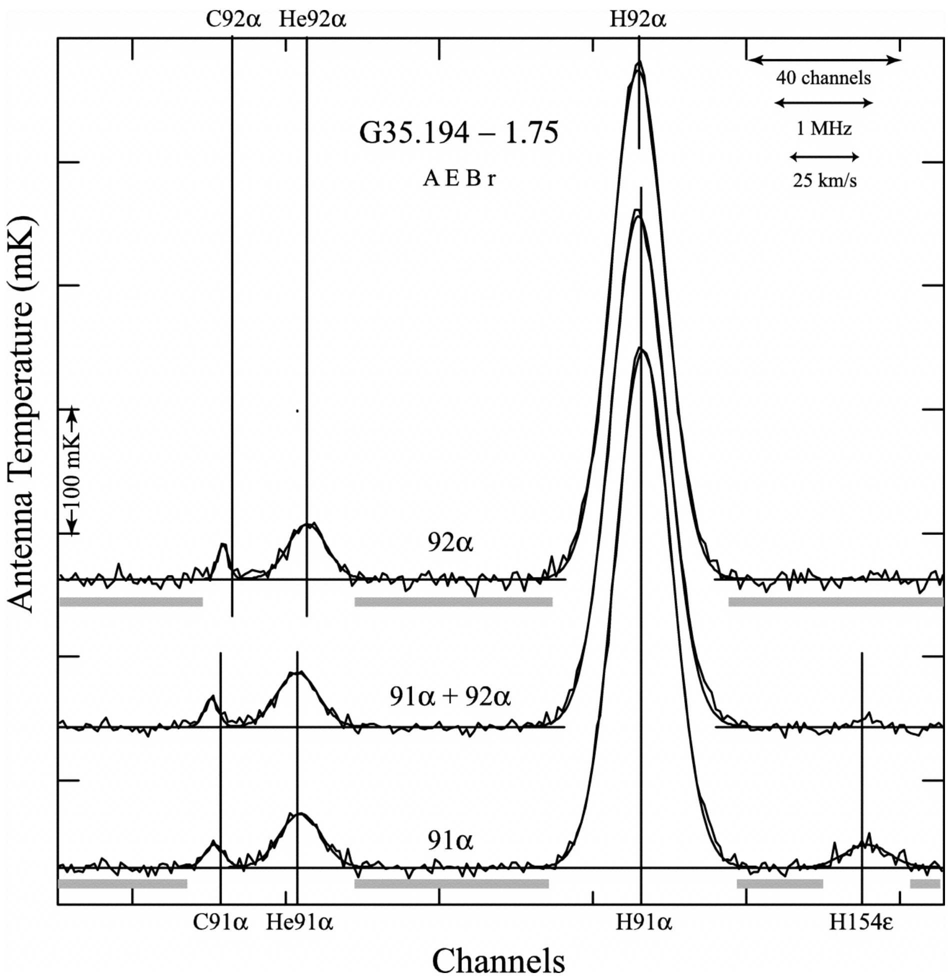


FIGURE 2.4: An example of observed recombination-line spectra from the 91α and 92α transitions of hydrogen, helium, and carbon observed in an H_{II} region (Ransom, 2016).

2.3.2 Non-thermal emission processes

An ensemble of radiating electrons with a decreasing power-law energy spectrum is common in astrophysical systems (e.g., the cosmic ray flux; Bradt 2010). The emission produced by a power-law distribution of an ensemble of radiating electrons, along a line of sight in the direction $\vec{k}$, can be approximated by

$$N_e(E,\, \vec{k})dE = K_e(\vec{k})E^{-\gamma}dE\ ,$$

where $\vec{k}$ is the wave vector, $N_e(E,\, \vec{k})$ the number of electrons on a line of sight per unit solid angle with energies E to E+dE, $K_e(\vec{k})$ the coefficient of the power-law amplitude for the energy spectrum of electrons along a line of sight, and γ is the index in the differential energy spectrum of particles

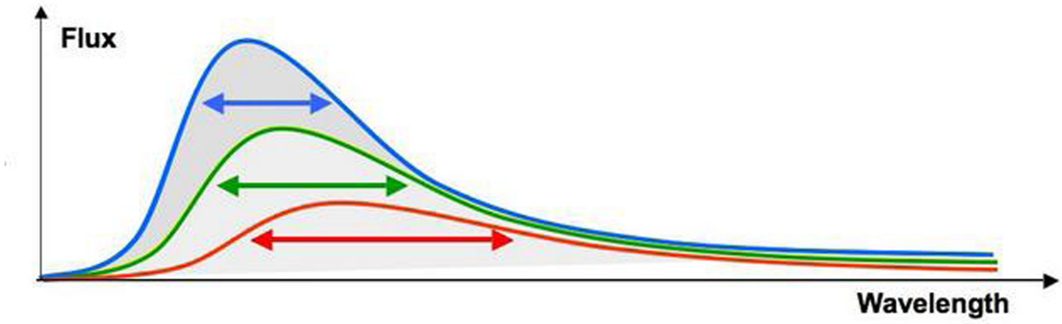


FIGURE 2.5: A plot of the flux against wavelength of a blackbody emitter, showing that such an object will emit a spectrum of radiation which peaks at a wavelength (or frequency) dependent only on the objects temperature (with the blue and red line representing hotter and colder objects, respectively). As the source temperature increases the peak wavelength decreases, as seen by the red. This is why hot stars appear blue and cool stars appear red (Swinburne University of Technology, 2010).

(Ginzburg & Syrovatskii, 1964). Considering a homogeneous and isotropic ensemble of radiating electrons, with a decreasing power-law energy spectrum, this approximation reduces to the equation

$$N(E)dE = K_e E^{-\gamma} dE \ ,$$

where $N(E)dE$ is the number of electrons per unit volume with arbitrary directions of motion and energies in the range E to E+dE, and K_e is the coefficient in the energy spectrum of electrons (Ginzburg & Syrovatskii, 1964; Pollack & Fazio, 1965). The optically thin emission spectrum can be described as

$$I_\nu \propto \nu^\alpha \ ,$$

where α is the spectral (emission) index of the exponential spectrum of emitting electrons (and $\alpha = \gamma - 1/2$).

Such an ensemble in the presence of a magnetic field will emit non-thermal emission (in the form of synchrotron emission) with a descending power law (Bradt, 2010). This then leads us to the distinction that most non-thermal emission processes occur in ionized plasmas, with free electrons being driven by an accompanying magnetic field.

Two of the processes that give rise to non-thermal emission include synchrotron radiation and inverse Compton scattering, also have similar loss rates:

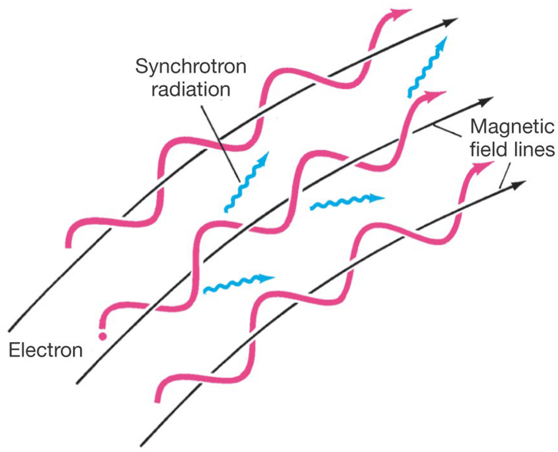


FIGURE 2.6: A simple illustration of a relativistic electron being accelerated by a magnetic field and the synchrotron radiation that is emitted as the charged particle gets boosted by the magnetic field (Brau, 2010). This is the source of the diffuse emission we wish to study from the extended radio sources we are interested in this study.

$$-\left(\frac{\mathrm{d}E}{\mathrm{d}t}\right)_{\mathrm{sync}} = \frac{4}{3}\sigma_{\mathrm{T}} c U_{\mathrm{mag}} \left(\frac{v^2}{c^2}\right) \gamma^2$$

and

$$-\left(\frac{\mathrm{d}E}{\mathrm{d}t}\right)_{\mathrm{IC}} = \frac{4}{3}\sigma_{\mathrm{T}} c U_{\mathrm{rad}} \left(\frac{v^2}{c^2}\right) \gamma^2$$

where σ_{T} is the Thomson scattering strength, U_{mag} and U_{rad} are the energy density of the magnetic and radiation field, respectively (Beckmann, 2006; Hu, 2021).

The extended radio sources we are interested in, in this thesis, produce synchrotron emission as the dominant emission mechanism (with an illustration in Fig. 2.6), and therefore, we mainly focus on this non-thermal emission process. Synchrotron emission is the acceleration of lighter charged ultra-relativistic particles (electrons and positrons) in the presence of a magnetic field. The relativistic electrons in nearly all synchrotron sources have power-law energy distributions, so they are not in LTE (Ransom, 2016). However, a synchrotron source with a relativistic Maxwellian electron-energy distribution would be a thermal source, so "synchrotron" and "non-thermal" are not completely synonymous (Wardzinski & Zdziarski, 2000; Ransom, 2016).

The lack of publicly available high frequency data in the soft and hard X-ray ranges does not allow us to test the inverse Compton (on synchrotron) branch of the core emission from the extended radio sources. X-ray data could also be used to identify possible inverse Compton scattering X-ray emission from the different radio galaxy lobes.

2.3.3 Diffuse emission

As cosmic ray electrons from extended radio sources undergo adiabatic loses, the non-thermal emission from extended structures diffuse to the surrounding medium/environment. The diffusion in time (assuming adiabatic cooling only) smooths observed structures, smudging the initial localisation signals, which trace the origins of the observed emission. This smoothed emission is duly named diffuse emission. If no other sources interact with this ambient emission, it cools down below the sensitivity threshold required for detection with radio surveys, thus, forming a reservoir of mildly relativistic cosmic ray electrons that are a seed population of low-energy particles.

Diffuse emission in different astrophysical environments (from a single isolated radio galaxy [see Chap. 4 for in depth discussions on diffuse emission at galaxy scales] to galaxy clusters [see Chap. 5 for in depth discussions on diffuse emission at galaxy cluster scales]) tends to be outshined by radio emission from more recent activity, and modelling its role/contribution in astrophysical environments is an active research area, especially in galaxy clusters. Traditionally (based on the location, size, morphology and cluster dynamics) diffuse radio emission in clusters is categorised into three categories, namely: i) radio halos (RH) & radio relics (RR) (Bonafede et al., 2021), ii) mini-radio halos (MRH; Gendron-Marsolais et al. 2020) and iii) revived fossil plasma (gently re-energized tails [GReET]; de Gasperin et al. 2017 and radio phoenixes; Mandal et al. 2020).

A key question in re-acceleration models is the origin of the turbulence responsible for re-accelerating the electrons. Gitti, Brunetti & Setti (2002) originally considered magnetohydrodynamic (MHD) turbulence amplified by compression in the cool cores, where the necessary energetics to power radio mini-halos is supplied by the cooling flow process itself. This model predicts a direct connection between the synchrotron emitted power in mini-halos and the energetics of (or energy available into) the cooling flow. Remarkably a trend between the radio power and the cooling flow power is observed in mini-halo clusters (Gitti et al., 2004, 2007; Gitti, 2013), where the strongest radio mini–halos are associated with the most powerful cooling flows, thus hinting at a direct connection between the thermal properties of cooling flows and the non-thermal emission of radio mini-halos.

Although these observations support a connection between synchrotron and (gas-)cooling power, the physical mechanisms responsible for the acceleration of the emitting electrons are still poorly constrained. In particular, in the re-acceleration model the origin of the turbulence necessary to trigger the electron re-acceleration is still debated; one possibility is that the turbulence is powered by AGN activity. We refer the reader to Chapters 4 and 5 for detailed introductions, and discussions, on diffuse emission at radio galaxy and galaxy cluster scales, respectively. We shift our focus to the difficulties associated with detecting this non-thermal emission.

2.4 Observational aspects

Radio interferometry involves combining multiple radio telescopes to function as a single radio telescope, with the goal of creating high-resolution images. An array of radio telescopes that functions in this manner is referred to as an interferometer. Thompson, Moran, & Swenson (2017), present a comprehensive explanation of the ideas, theories, and practical methods used in radio interferometric observations. We particularly recommend this work, as it extensively covers radio interferometry concepts such as baseline designs, beams, and calibration techniques for the different surveys that are used in this thesis. We focus on how various designs affect search and characterisation strategies for detecting diffuse radio emission.

Survey configuration, frequency, sensitivity and angular resolution are important factors to consider when in search of diffuse emission. "Traditionally"[1] a trade off between the type of emission and field-of-view (FOV), the emission is dominant in, influenced the choice of survey configuration. That is, to probe the core of a radio source (and thermal emission processes at play, as described in Chapter 2.3.1) high angular resolution images are necessary. To obtain high angular resolution images, a longer baseline is require to resolve source dynamics at smaller scales. Thus, the ideal survey configuration to probe the core of a source requires the longest possible baseline in order to resolve radio emission to the smallest possible scale.

Since non-thermal processes in radio sources follow a power-law, we expect an increase in the observed radiation, from synchrotron processes, with a decrease in observing frequencies. We also note that the observable radio frequency band spans the largest frequency range MHz to GHz (see Fig. 2.2). The dynamical state images (from continuum radio surveys) provide a means of localising emission within and around sources, depending on the sensitivity and resolution of the survey used. And since low-frequency radio

[1]Modern interferometers, such as the MeerKAT radio telescope, have a mixture of long and short baselines which allow for optimal radio emission detection at multiple scales.

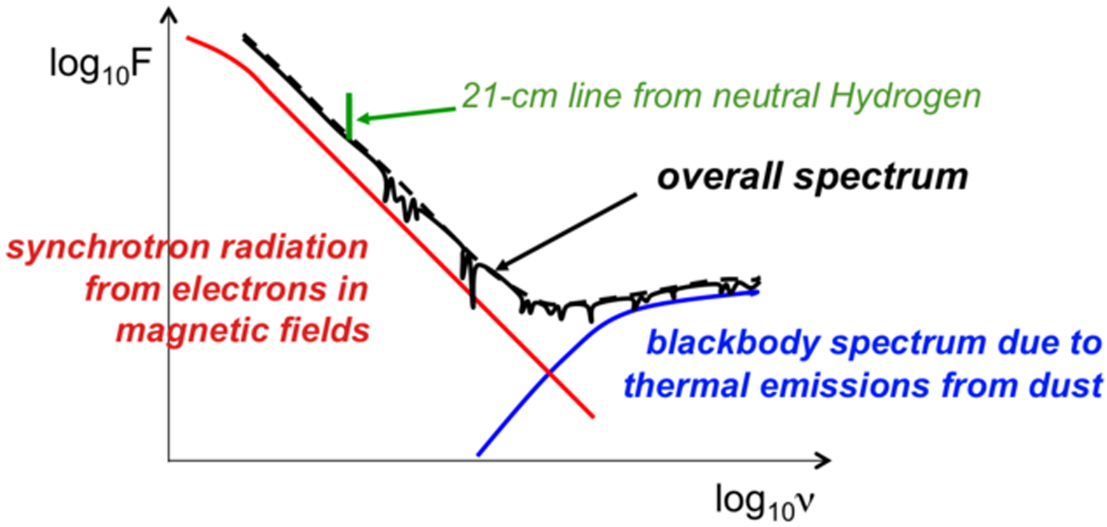


FIGURE 2.7: A log-log scale plot of a composite spectrum for a typical galaxy observed over a wide range of frequencies, from radio to visible (Swinburne University of Technology, 2010).

surveys are sensitive to longer wavelengths, they are ideal probes for this dominant physical phenomena at lower energy scales.

The analyses of the core, coupled with survey (dynamical state) images creates a complimentary framework; where, the flux density spectra from the core of radio sources, at different frequencies, provide clues about the contribution of the thermal emission mechanism at play. While the observed survey images capture the dynamical state of the source. Using the spectral index map and source characterisation, we can then map the radio emission associated with radio sources and diffuse radio emission (see Chapters 4, 5, and the references therein, for detailed discussions on the characterisation of diffuse emission at radio source and galaxy cluster scales).

We can now use Fig. 2.7 to sum up the typical features we are likely to observe from spectral line emission, thermal and non-thermal emission processes. The observed spectral line emission is shown in black, with the absorption and emission features appearing as deviations from the the overall spectrum. The expected descending power law distribution from synchrotron radiation is shown in red, highlighting the fact that at low-frequencies we have a higher contribution from low-energy photons compared to high-energy photons. As the frequency increases we expect thermal emission processes from free-free absorption and thermal dust to become dominant. The case of a nearby starburst galaxy M82 is also presented in Fig. 2.8 as an example. Now that we have covered some of the basic concepts that are necessary to characterise extended radio sources, we proceed to provide motivation as to why it is necessary to carry out this study.

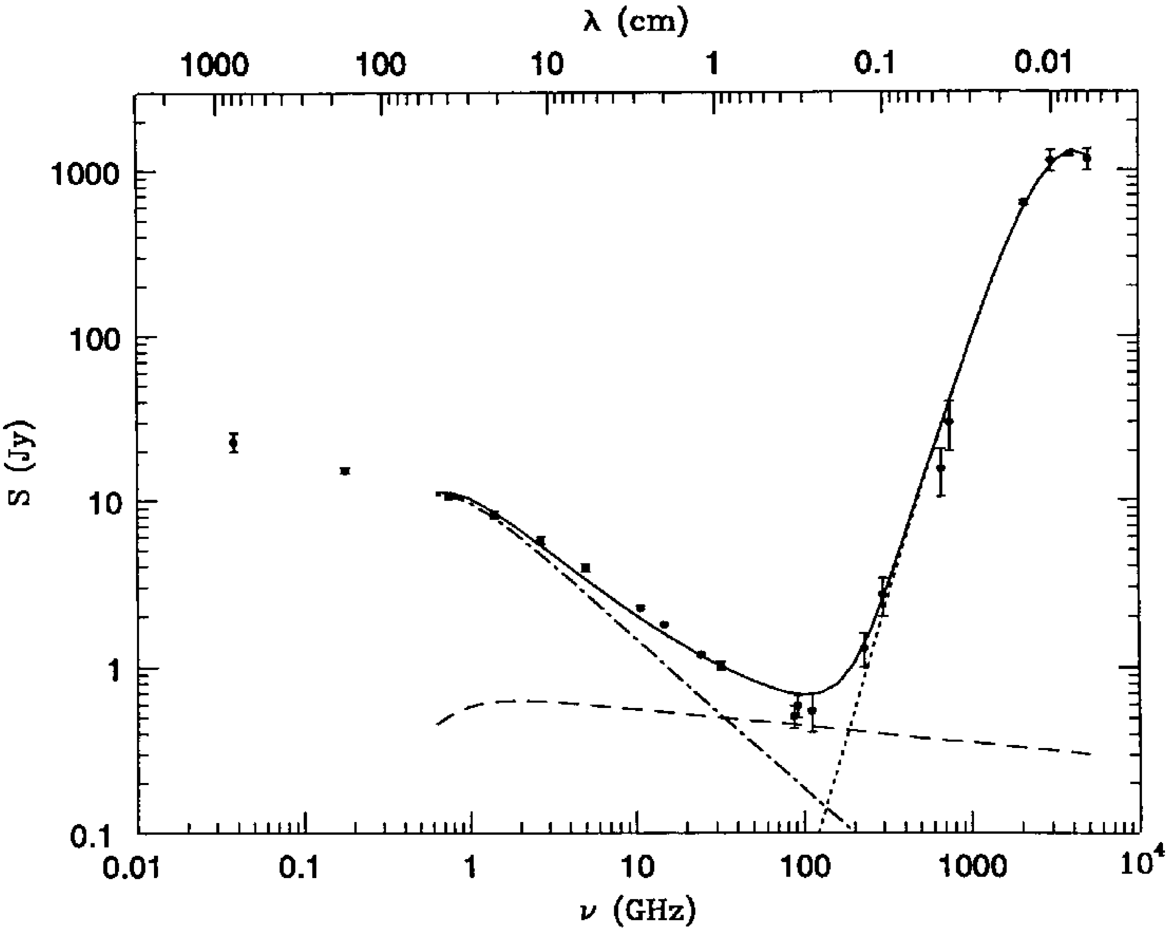


FIGURE 2.8: The radio and far-infrared spectrum of the nearby starburst galaxy M82 (Condon, 1992). The contribution of free-free emission is indicated by the nearly horizontal dashed line. Synchrotron radiation (dot-dash line) and thermal dust emission (dots) dominate at low and high frequencies, respectively. Free-free absorption from H_{II} regions distributed throughout the galaxy absorbs some of the synchrotron radiation and flattens the overall spectrum at the lowest frequencies (Ransom, 2016).

2.5 Motivation

Low surface brightness diffuse emission can be used as a probe of the (non-trivial) interplay between the non-thermal processes in radio sources and their environment, at different scales and frequencies. Though this emission contains a wealth of astrophysical and cosmological signals (such as jets and lobes, radio halos, relics, CR diffusion, the impact on clusters and feedback on LSS), the nature and role of this emission in galaxy feedback and formation is not well understood in the literature (since it is hard to detect, due to the dim nature of low surface brightness diffuse objects and their origins). Unsupervised machine learning techniques can assist in understanding and unmasking structures contained within data of these dim objects.

We aim to develop advanced methods for differential data analysis (DDA) procedures of the diffuse emission from extended radio sources at low to mid-frequency ($\sim$70-1400 MHz; and in general of other cosmological backgrounds that are observable at low radio frequencies) that will function at scale, using ML techniques. Our study will benefit from the improved sensitivity of the available radio surveys (in particular LoTSS) and the ability to better resolve radio galaxies for studies of magnetic fields, cosmic ray diffusion, particle acceleration mechanisms in their jets/lobes. We will focus on the properties and the nature of specific extended radio galaxies which show variance from the expected FR-II morphology, for example, galaxies which show bubble-like symmetric radio lobe structures and with a presence of a prominent radio core.

We hypothesize that it is possible to use ML techniques for searches of previously undiscovered extended radio sources, and that we can develop advanced DDA procedures to characterise the diffuse emission of extended radio sources at scale.

To test this hypotheses, we plan on accessing the different imaging surveys and using the open data across multiple frequencies, in order to identify and characterise extended radio sources. Specifically, for our analyses, we have identified publicly available data from the FIRST, GLEAM, LoTSS, NVSS, TGSS, VLASS, VLSSr, and WENSS surveys (described in Chaps. 3.2 and 4.2) as the ideal surveys to identify and characterise continuum emission from radio sources. Follow up multi-wavelength observations are crucial for source localisation and disentangling astrophysical dynamics, through source characterisation, in other wavelengths. We will use the WISE, DSS-DR2, and SDSS surveys, as complimentary surveys, to produce our multi-frequency characterisation of the sources detected in the radio analysis.

We do note that the differences in resolutions and sensitivities across these surveys necessitates a careful selection criteria for deciding which surveys to use in our proposed studies. This selection criteria will be guided by the versatility of Python (a high-level programming language) and how we can integrate it with existing radio data analysis software (such as CASA, MIRIAD, etc.) to mitigate these differences.

ML will be, mainly, used in a two phase manner (with the second phase being a secondary goal and dependent on the success of the first). The first phase entails using an unsupervised learning method, self-organising map (SOM; Kohonen 1997; Polsterer, Gieseke & Igel 2015), to feed in the downloaded images from the different surveys. The product of this analysis will be abstractions (for each survey) of what the algorithm learns as similar representations of the input images, known as the Kohonen map (see Fig. 1.1 for an example of a Kohonen map). From the Kohonen map of each survey, we will proceed to the second phase where, we will group all of the correlated representations from the different surveys into unified feature classes that best represent the abstracted features. The abstracted pixel intensity features, from the SOM, are a proxy classification of radio morphologies based on the observed astroparticle dynamics of the radio source.

We will then use, in the second phase, convolutional neural network (CNNs, a class of deep neural networks designed for analysing image data) to segment out new features at different field-of-views, to extract information at different resolutions. The lower resolution images (zoomed out field-of-views) give insight on cosmological features (such as the large scale structure of the universe), whilst the high resolution images (zoomed in field-of-views) map out the astrophysical nature of the sources.

Therefore, using the fact that radio continuum images are tracers of dynamical states of radio sources at multi-frequencies, the specific questions we would like to answer (in Chap. 6) are:

- Using advanced analytics, can we curate and probe astrophysical sources at scale?
- Can we perform knowledge discovery in databases?
- Can we use clustering techniques to identify dense environments and is there a (physical) causal link to the observed morphology of radio galaxies?

The structure of the following self contained works is as follows: We detail the methodology used in our heuristic based search (the visual based search we carried out as a precursor study) for radio sources with low surface brightness diffuse emission in Chap. 3. 17 potential sources were identified as possible candidates with interesting features to follow up for characterisation. These were later narrowed down to 12 sources, with Chap. 4 being the report on our findings; on a heuristic based search for radio galaxies with low surface brightness diffuse emission, as guided by the LoTSS pre-print paper data products. PGC2285791, a star-forming giant radio galaxy with low surface brightness diffuse emission, was identified as the focal source to report on from the thresholding technique discussed in Chap. 3. From a parallel search, on galaxy clusters only visible to surveys in the southern hemisphere. We present in Chap. 5 the characterisation of the sources in Abell 1689, its diffuse radio emission and astrophysical environment; in a study which was co-authored with N. F. Khanye.

Both studies, as reported in Chaps. 4 and 5, were key drivers towards designing a multi-frequency and multi-field-of-view, unsupervised based search of non-thermal diffuse emission in extended radio sources, at scale. The insights gained from characterising the sources in these chapters highlighted a need to design an architecture which included pre-trained networks (for abstracting features from data cubes) and considered multiple FOVs (as a mechanism of focusing on diffuse radio emission at different scales). The initial proposed study only considered looking at radio galaxy scale diffuse emission. We then look at how we can transfer the gained insights, from the characterisation of the sources in Chaps. 4 and 5, to use pre-trained networks as abstractors of fundamental features contained within multi-frequency and multi-scale radio continuum data cubes, in Chap. 6. The key results from these three chapters are summed up in the final Chap. 7.

Chapter 3

Methodology and Data Analyses Techniques

This Chapter outlines our methodology for conducting the required preparatory research for an unsupervised search of non-thermal diffuse emission in extended sources. We describe how the radio sources were selected, downloaded, and curated for the analyses of the non-thermal diffuse emission studies discussed in Chapters 4 and 5, especially noting that the selected sources were identified using a heuristics-based approach. The preliminary research carried out in Chapters 4 and 5 are essential case studies involving observations of radio sources at multiple frequencies within a single field of view, focusing on diffuse emission from radio galaxies and galaxy clusters. To do this, we will manually detail the procedure followed in studying and characterising the core, the extended structures, and the environment of the selected radio sources.

3.1 Introduction

When analysing continuum radio images for the detection of low-surface brightness diffuse emission, the presence of compact, bright radio sources introduces significant obstacles. They create a large difference between the highest and lowest measurable luminosity intensities. This high difference between maximum and minimum measurable light intensities makes it less obvious to observe the "dim" diffuse emission we are interested in.

Within continuum radio fields, one usually anticipates images with prominent bright dots representing emission from compact radio sources, a fluctuating background indicating noise, and occasional shadows hinting at the extended radio emission of interest. It is essential to carefully consider these competing characteristics when characterising the different components that make up a radio source.

Radio continuum surveys are an interesting dataset to consider since they are designed with varying scientific objectives and optimised observational strategies (see Chap. 3.2 for detailed descriptions of the legacy radio continuum surveys used in the different studies and their optimised observational strategies). Optimised observational survey designs are necessary for studying the characteristics of radio sources and their extent of interaction with the surrounding environment. I.e., to perform source extraction and characterisation, radio astronomers have to look at a combination of surveys optimised for the detection of the core (these are typically high angular resolution surveys with a wide baseline; e.g., FIRST Becker et al. 1995 and VLASS Lacy et al. 2020) and extended structures (typically lower angular resolution surveys with a compact core; e.g., NVSS Condon et al. 1998 and LoTSS Shimwell et al. 2017). The latest generation of radio telescope arrays, such as MeerKAT (Jonas & MeerKAT Team 2016; the precursor to the SKA; Dewdney et al. 2009; Weltman et al. 2020), are optimised to simultaneously detect radio emission from both the core and extended structures of a source by designing an array with both compact and wide baselines.

In both cases, whether considering radio continuum surveys optimised for a specific angular resolution range or arrays with a wide angular resolution range, there is a need to look at the non-trivial interplay between datasets at different angular resolutions. We also have to consider the changes in dominant physical mechanisms at play in the radio frequency band. This is the largest observable frequency range in astronomy, with multiple probes from 10-240 MHz (LOFAR; van Haarlem et al. 2013) to 3 GHz (VLASS; Lacy et al. 2020), for publicly available legacy datasets of extragalactic radio source, and can easily be extended to 8-15 GHz when considering the MOJAVE multi wavelength dataset (Lister et al., 2018).

Though the photons observed by radio probes might originate from the same radio source(s), the dynamical state captured at very high, high, mid and low-frequencies give different clues about the interactions which shape the overall source(s). That is, observations at very high frequencies (8-15 GHz; VLBA) observations assist in resolving and modelling individual components of the optically thick core regions of radio sources; high-frequency (3-5 GHz; VLA) observations assist in source identification, spectra and modelling the overall dynamics of the core regions of radio sources; differing from the optical thin mid (800-1670 MHz) and low-frequencies (70-325 MHz), which capture the cosmic ray electron population, ambient medium, and radio plasma dynamics.

Studying the electrons from the optically thin low-mid frequencies, using the observed dynamical states captured by radio continuum images at these frequencies, we can analyse the spectral indices calculated from spectral index maps of jets to study the kinematic features tracing jet evolution and interaction with the surrounding ambient medium. Thus, the angular resolution and frequency dependent nature of radio emission is what makes it challenging to draw generic algorithmic rules for source extraction and characterisation, which are open problems in radio astronomy. It is therefore challenging to draw up simple, low-dimensional prescriptions of how the non-trivially interrelated datasets should be mapped to another space. We utilise several radio continuum surveys to study how interfacing humans with machine learning can enhance scientific discovery.

3.2 Legacy Radio Continuum Surveys

We used archival data from the following radio surveys to study extended radio continuum sources. These studies ranged from in-depth multi-frequency analyses of radio emission from single sources in Chapter 4 and galaxy clusters in Chapter 5 to new ways of grouping and classifying radio sources with unsupervised machine learning in Chapter 6. The data was derived from publicly available large radio surveys with the following descriptions:

3.2.1 VLA Low frequency Sky Survey

The VLA Low frequency Sky Survey (VLSS; Cohen et al. 2007) imaged the sky north of declination -30°, at 74 MHz (bandwidth of $\Delta\nu \sim 1.56$ MHz) and a resolution of $80'' \times 80''$. The observations were made between 2001 to 2007 in the VLA B (for a declination in the range: $-10^{\circ} < \delta < 80^{\circ}$) and BnA (for $\delta < -10^{\circ}$ and $\delta > 80^{\circ}$) configurations. The survey has an intergrated time of ≈ 75 minutes per pointing achieving an rms noise level of 100 mJy/beam. For our studies we used the re-reduced VLSS data, VLSS Redux (VLSSr; Lane et al. 2014). The final VLSSr catalogue contains over 92 000 radio sources with a resolution of $75'' \times 75''$.

3.2.2 TIFR GMRT Sky Survey alternative data release 1

The TIFR GMRT Sky Survey alternative data release 1 (TGSS ADR1; Intema et al. 2017) is a publicly accessible catalogue based on the GMRT observations of the radio sky at 150 MHz ($\Delta\nu \sim 16.7$ MHz) and an angular resolution of $25''\times 25''$. The survey achieved over 2000 hours of observations with each pointing being carried for about 15 minutes between the declination of -20° and +60°. The data was processed using the GMRT's automated radio data pipeline, SPAM (Source Peeling and Atmospheric Modeling; Intema 2014) which is a python based program that uses AIPS (the Astronomical Image Processing Pipeline; Intema et al. 2009, 2017) to reduce low frequency, high resolution interferometric radio data.

3.2.3 GaLactic and Extragalactic All-Sky MWA Survey

The GaLactic and Extragalactic All-Sky MWA (GLEAM; Wayth et al. 2015) survey is one of the Square Kilometre Array (SKA) pathfinder telescopes. The survey maps the entire sky, south of the declination δ = +25°, at radio frequencies ranging from 72 to 231 MHz ($\Delta\nu \sim 7.68$ MHz), grouped into four central frequencies: 88; 118; 155 and 200 MHz. The observations were made from August 2013 and lasted for 2 years. We used the data released from the first year of observation. The data reduction process was carried out by the GLEAM team (Hurley-Walker et al., 2017).

3.2.4 Westerbork Northern Sky Survey

The Westerbork Northern Sky Survey (WENSS; Rengelink et al. 1997) is a low-frequency radio survey observed at a frequency of 325 MHz and 609 MHz, to produce continuum images with an angular resolution of $54'' \times 54''$ cosec δ and $28'' \times 28''$ cosec δ, respectively, typical noise levels that are less than 5 mJy/beam and a limiting flux density of about 18 mJy. We used the 325MHz products from this survey.

3.2.5 NRAO VLA Sky Survey

The NRAO VLA Sky survey (NVSS; Condon et al. 1998) is a publicly accessible archive containing more than 1.8 million radio sources with flux densities stronger than $S_{1.4\,\mathrm{GHz}} > 2.5$ mJy. The NVSS mapped the sky at 1400 MHz ($\Delta\nu \sim 50$ MHz) north of declination $\delta = -40°$, using the VLA D and DnC configurations between September 1993 and October 1996 for a total of 2700 hrs. The NVSS catalogue completely overlaps with the TGSS ADR1 catalogue.

3.2.6 Faint Images of the Radio Sky at Twenty centimeters Survey

The Faint Images of the Radio Sky at Twenty centimeters survey (FIRST; Becker et al. 1995) was carried out using the NRAO VLA at 1400 MHz in the B-array between 1993 and 2002. The survey is the radio equivalent of the Sloan Digital Sky Survey (SDSS) covering $\sim 10,000$ square degrees of the North and South Galactic Cap. The archive consists of 3-minute snapshots which cover a hexagonal grid using fourteen 3 MHz frequency channels. The survey's 5σ flux density limit is 1 mJy, with a restoring beam FWHM of 5.4 arcsec and an rms of 0.15 mJy/beam. The final catalog contains $\sim 9 \times 10^5$ radio sources, roughly 40% of which are resolved on scales from 2-30 arcsec.

3.2.7 VLASS

The Very Large Array Sky Survey (VLASS; Lacy et al. 2020) is a synoptic, all-sky radio survey being carried out with the Karl G. Jansky VLA in B-configuration and S-band, between the period of September 2014 and October 2024. The survey has a central frequency of 3000 MHz ($\Delta\nu \sim$ 2-4 GHz) with an angular resolution of $\sim 3''$ and a sensitivity goal of 70 μJy at 1σ. VLASS will use approximately 5500 hours and observe the whole sky north of of declination $\delta = -40°$ covering a total of 33,885 square degrees.

We now describe the heuristics-based data exploration procedure followed to look for extended radio sources with low-surface brightness features. This is based on applying a simple thresholding technique to make it easier to find extended sources, using mosaics from the LoTSS catalogue as illustrative examples. This exploration is suitable for array designs optimised for extended radio sources like TGSS, WENSS, and NVSS, but not for arrays optimised for compact radio sources such as VLSSr, FIRST, and VLASS. The data analyses and sample spectral energy distribution (SED) calculations are then presented in the final section.

3.3 Data exploration procedure

As emphasised in the introduction, in mining radio astronomy images for extended radio sources, one of the major challenges we face is finding bright radio point sources. The challenge they introduce was made apparent when we initially surveyed (by visual inspection) the mosaics from the LoTSS catalogue. We noted that:

- most of the bright point sources emit well above 200 mJy/pixel,
- the diffuse emission we are interested in studying is typically radiated below 80 mJy/pixel, and
- most images (from the different radio surveys) have noise levels that are typically at 5 mJy/pixel and even lower for modern telescopes, reaching levels of 0.5 mJy/pixel.

To reduce the large difference between the maximum and minimum measurable light intensities, we applied a mask to bright sources [1] (see Figs. 3.1 and 3.2, smaller cut outs from a larger mosaic). Once the mask was applied, we began to notice an increased number of sources that were extended in nature. Furthermore, these sources were now among the brightest, making it easier to detect sources with large diffuse emission.

[1] In this instance we are loosely categorising bright sources as sources that lie above the 160 mJy/pixel intensity threshold.

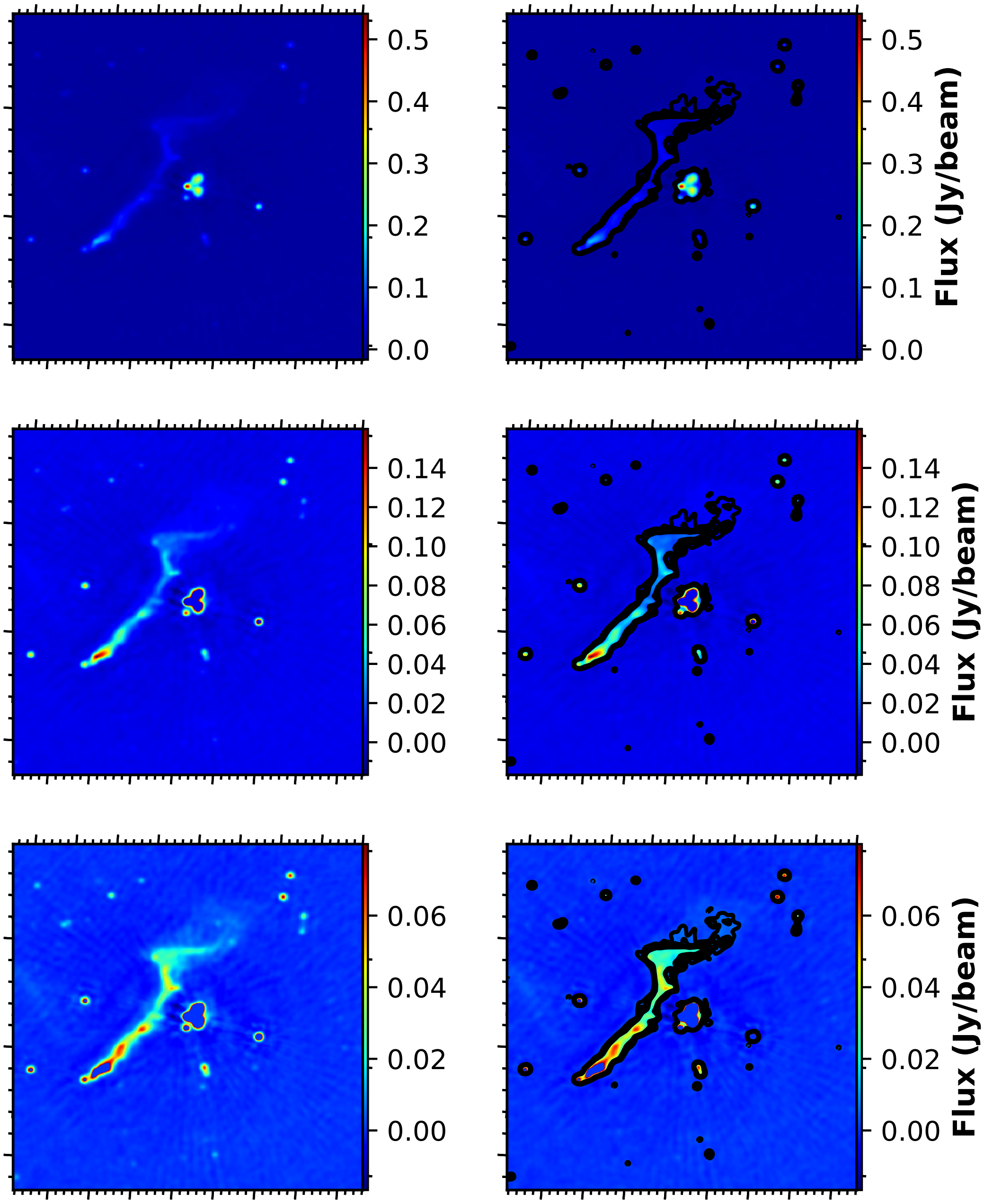


FIGURE 3.1: First row: the LoTSS 141 MHz continuum image of IC 711, shown without contours on the left and overlaid with black LoTSS contours on the right, starting at at 3σ (σ = 1.31 mJy/beam) with 2σ increments and ending at 14.41 mJy/beam. Second row: same images as the First row, with a mask applied to IC 711 at the 160 mJy/pixel intensity threshold to remove "bright" radio point sources and highlight the obscured diffuse emission we analyse in this study. Third row: same images as the First row, with a mask applied to IC 711 at the 80 mJy/pixel intensity threshold for the same purpose mentioned in the second row.

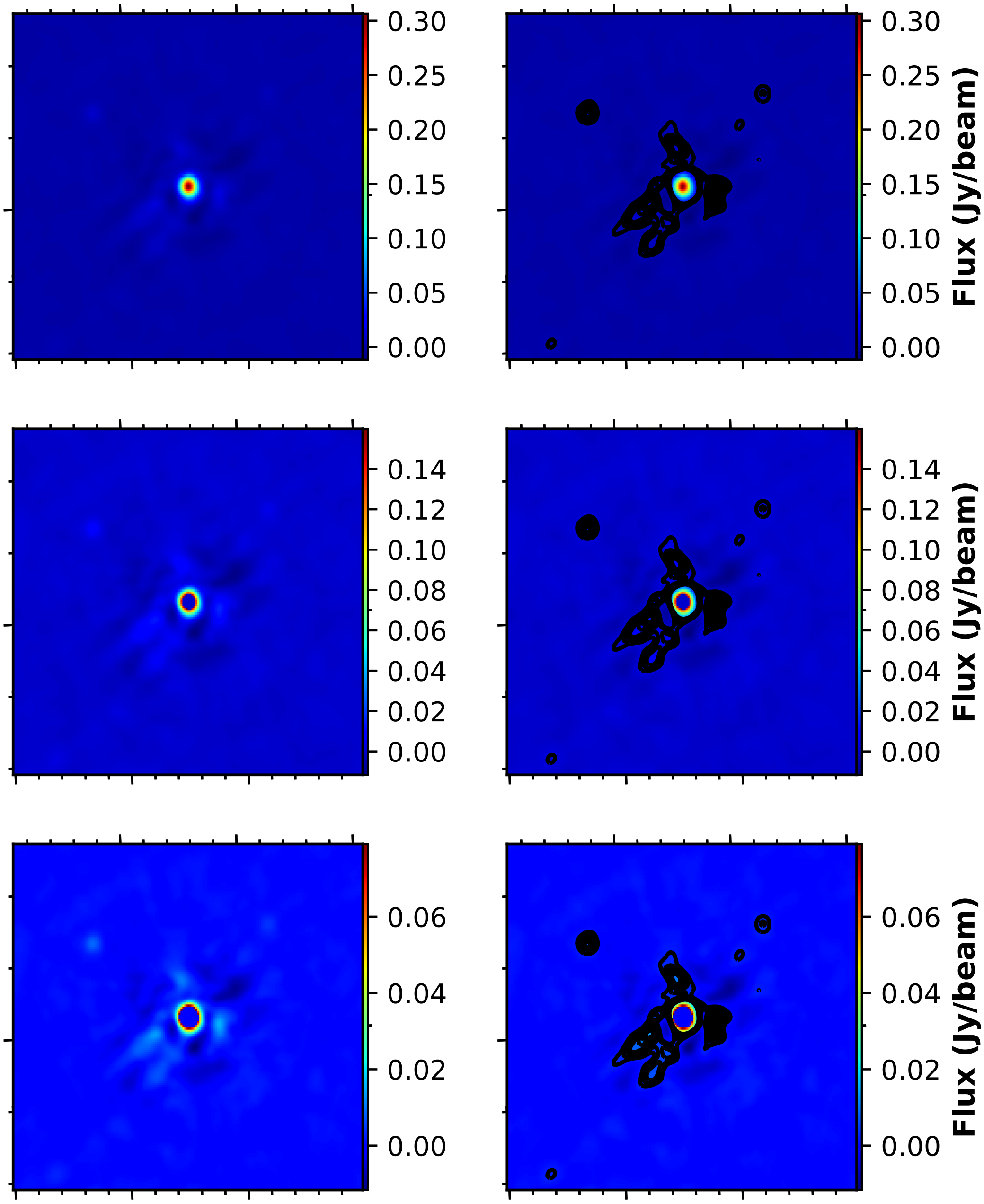


FIGURE 3.2: First row: The LoTSS 141 MHz continuum image of source 13, shown without contours on the left and overlaid with black LoTSS contours on the right, starting at 3σ (σ = 0.81 mJy/beam) with 2σ increments and ending at 15.39 mJy/beam. Second row: same images as the first row, with a mask applied to source 13 at the 160 mJy/pixel intensity threshold to remove "bright" radio point sources and highlight the obscured diffuse emission we analyse in this study. Third row: same images as the first row, with a mask applied to source 13 at the 80 mJy/pixel intensity threshold for the same purpose mentioned in the second row.

We then found that even at the 160 mJy/pixel intensity threshold, there were still a sizeable number of point radio sources, so we further applied masking at different intervals until we got to the 80 mJy/pixel intensity threshold. At the 80 mJy/pixel intensity threshold most of the images were dominated by diffuse emission from the extended radio sources. We then looked for interesting sources in different mosaics and found that the P7-Mosaic (from LoTSS) was populated by a large variety of different and interesting extended radio sources. We then identified 15 extended radio sources from this mosaic for analysis, with the addition of PSZ2 G133.60+69.04[2] and A1304+4649A (which were taken from the P27-Mosaic, imaged at a frequency of 125.31 MHz).

To clarify which sources were considered as original sources, we state that source 01 to 10 and IC 711 were the first sources considered for this analysis. Whilst carrying out some preliminary analysis on these initial sources, the remaining sources were found and added on to the list. Initially, source 10 was grouped with source 11 and 12 as a cluster. This grouping was separated based on the observed differences in the spectral index analysis and spatial separation of the sources. Source 13 and B31059+460-00 were added after the second iteration of the data exploration process as missed sources. The final two sources PSZ2 G133.60+69.04 and A1304+4649A are from the well known clusters A1550 and A1682, respectively.

3.4 Data analyses

The 17 identified extended radio sources that were selected using the heuristic based data exploration procedure outlined earlier, are given in Table 3.1. The source names (in cases where the core of the source had no designated name, we named the source incrementally starting from Source 01) and positions of the extended radio sources along with the spectral indices calculated from the different surveys are also included in these tables.

To calculate the integrated flux density uncertainties we used the relation:

$$\sigma_{S_\nu} = \sqrt{(\sigma_{\rm cal} S_\nu)^2 + \left(\sigma_{\rm rms}\sqrt{N_{\rm beam}}\right)^2},$$

where $\sigma_{\rm cal}$ is the uncertainty in the calibration of the flux scale, S_ν is the integrated flux density at the observed frequency ν, $\sigma_{\rm rms}$ is the rms of the image and $N_{\rm beam}$ is the number of independent beams in the source (Cantwell et al., 2016). Here, $\sigma_{\rm cal}$ is taken to be 10% for the LoTSS (Shimwell et al., 2017), NVSS (Condon et al., 1998) and TGSS (Intema et al., 2017) surveys, and 15% for the VLSSr (Lane et al., 2014) survey.

[2]Extended radio source in the galaxy cluster A1550.

TABLE 3.1: Data on the selected extended sources detected by *LoTSS*. Column (1): Radio source names; Column (2): Right Ascension (J2000); Column (3): Declination (J2000)

Source Name (1)	R.A. (J2000) (2)	Dec (J2000) (3)
Source 01	10 51 33	+50 12 49
Source 02	11 00 41	+45 51 00
Source 03	11 03 39	+45 50 07
PGC2285791	11 04 33	+46 42 24
Source 05	11 08 05	+51 20 53
Source 06	11 08 15	+51 29 53
Source 07	11 09 41	+44 58 28
PGC2396883	11 11 57	+51 36 22
Source 09	11 27 10	+51 12 25
Source 10	11 32 50	+50 57 05
Source 11	11 32 47	+51 01 19
Source 12	11 33 04	+51 04 26
Source 13	11 01 29	+46 00 26
B31059+460	11 02 49	+45 44 43
PSZ2 G133.60+69.04	12 28 53	+47 36 49
A1304+4649A	13 06 45	+46 33 33
IC 711	11 34 46	+48 57 24

The $\sigma_{\rm rms}$ was calculated for each source, using continuum images with a field-of-view of 10 arcmin for each survey. In cases where sources had weak detection signals, $\sigma_{\rm rms}$ was used to compute the 3σ upper limit for that respective survey.

Once we have obtained the integrated flux density of a radio source, we can calculate the radio source's spectral luminosity (provided we have the radio source's luminosity distance). For a ΛCDM cosmological model with $\Omega_{\rm k} = 0$, the luminosity distance, D_L, at redshift z is given by

$$D_L = \frac{c(1+z)}{H_0} \int \frac{dz}{\sqrt{(1+z)\zeta}}$$

with

$$\zeta = \sqrt{1 + \Omega_{\rm m} z + \Omega_\Lambda [1 - (1+z)^{-2}]}\,,$$

where H_0 is the Hubble constant and $\Omega_{\rm m}$, Ω_Λ are the density parameters for matter and the cosmological constant, respectively (Zhou, 2002). The equation does not include the k-correction term for a power law SED (i.e. $(1+z)^{(1+\rm alpha)}$, where alpha = -0.7) due to the sources being at very low redshift, hence the k-corrections being negligible.

In most cases where a source has an observed optical counterpart, we can obtain the source's spectroscopic information from a multi-spectral imaging and spectroscopic redshift survey such as the Sloan Digital Sky Survey (SDSS; Blanton et al. 2017). We can then use the redshift from such surveys in the above equation to obtain D_L. Alternatively, we can query the NASA/IPAC Extragalactic Database (NED) and retrieve the luminosity distance from this database. This is the easier route to follow because one retrieves the D_L calculated value along with the parameters for distances and the adopted cosmology used, which are stated as $H_0 = 73.0$, $\Omega_{\rm m} = 0.27$, and $\Omega_\Lambda = 0.73$ with the derived quantities using a redshift corrected to a reference frame defined by the 3K CMB. This is the cosmology we adopt in this analysis.

The spectral luminosity L_ν of a source is defined as the total power per unit bandwidth radiated by the source at frequency ν, its mks units are W Hz^{-1}. The inverse-square law relation between spectral luminosity and flux density in free space is

$$L_\nu = 4\pi D_L^2 S_\nu \ ,$$

where the distance (luminosity distance) D_L between the source and the observer is much larger than the dimensions of the source itself (Ransom, 2016).

From the mks units of the spectral luminosity (W Hz^{-1}), we can easily see that the spectral luminosity gives us a measurement of how much energy (transferred per unit time) is produced by the source at the observed frequency. We would, therefore, expect different spectral luminosities at different frequencies. This allows us to plot the spectral energy distribution (SED) of a source at different frequencies. SEDs are energy plotted against some measure of the photon – frequency or wavelength. In our case we are interested in plotting νS_ν against ν.

Now to carry out an example of how to implement the above procedure, we will use the extended radio galaxy PGC2285791 and present our calculations for the core radio flux density spectrum of this extended radio galaxy (shown in Fig. 4.2), which is used to create the SED of the source (shown in Fig. 4.4).

Considering the case where we are integrating over the area equal to the size of the LoTSS beam, we find that:

$$\begin{aligned} S_{\rm int\_core(141MHz)} &= 7.34 \pm 0.73 \ {\rm mJy} \\ S_{\rm int\_core(1.4GHz)} &= 1.04 \pm 0.10 \ {\rm mJy} \ , \end{aligned}$$

where we are not considering the calculated 3σ upper limit, for both the TGSS and VLSSr surveys.

PGC2285791 has an optical spectrum, found in the SDSS DR13, where it is classified as a star-forming galaxy at redshift $z = 0.14126 \pm 0.00001$. Using the source's luminosity distance of 645 Mpc in the adopted cosmology, the radio luminosity of the LoTSS and NVSS surveys is given by

$$\begin{aligned} L_\nu &= 4\pi D_L^2 S_\nu \\ L_{141\text{MHz}} &= 4\pi D_L^2 S_{\text{int\_core}(141\text{MHz})} \\ L_{1.4\text{GHz}} &= 4\pi D_L^2 S_{\text{int\_core}(1.4\text{GHz})} \,. \end{aligned}$$

By definition, we know that the mks units of flux density, $\text{Wm}^{-2}\text{Hz}^{-1}$, are much too big for practical astronomical use, so smaller values are then defined as

$$1\text{Jansky} = 1\text{Jy} \equiv 10^{-26}\text{W m}^{-2}\text{Hz}^{-1} \equiv 10^{-23}\text{erg s}^{-1}\text{ cm}^{-2}\text{Hz}^{-1}$$

in terms of the mks and cgi units, respectively. We, therefore, have

$$\begin{aligned} S_{\text{int\_core}(141\text{MHz})} &= 7.34 \pm 0.73 \times 10^{-3}\text{Jy} \equiv 7.34 \pm 0.73 \times 10^{-29}\text{ Wm}^{-2}\text{Hz}^{-1} \\ &\equiv 7.34 \pm 0.73 \times 10^{-26}\text{ erg s}^{-1}\text{cm}^{-2}\text{Hz}^{-1} \end{aligned}$$

$$\begin{aligned} S_{\text{int\_core}(1.4\text{GHz})} &= 1.04 \pm 0.10 \times 10^{-3}\text{Jy} \equiv 1.04 \pm 0.10 \times 10^{-29}\text{ Wm}^{-2}\text{Hz}^{-1} \\ &\equiv 1.04 \pm 0.10 \times 10^{-26}\text{ erg s}^{-1}\text{cm}^{-2}\text{Hz}^{-1}. \end{aligned}$$

Thus, the radio luminosity of the LoTSS and NVSS survey is

$$\begin{aligned} L_{141\text{MHz}} &= 3.65 \times 10^{23}\text{ W Hz}^{-1} \\ L_{1.4\text{GHz}} &= 0.52 \times 10^{23}\text{ W Hz}^{-1} \,. \end{aligned}$$

We then use $\nu \times S_\nu$ to obtain the SED of the source at the different frequencies and plot νS_ν against ν. The overall multi-ν SED of the radio galaxy core, derived from the combination of archival data and of the radio data obtained from this analysis, is shown in Fig. 4.4.

In this Chapter, we examined the issues related to detecting diffuse radio emission in continuum radio images and the data sources used to characterise and explore radio sources with this emission at various scales, as elaborated in subsequent Chapters. We demonstrated how a basic thresholding method might enhance the detectability of diffuse radio emission in extended sources, resulting in the initial set of candidate sources for further investigations. We performed calculations for the radio spectrum and spectral energy distribution (SED) of the galaxy. The following Chapters provide systematic methods for characterising diffuse radio emission and its interaction with the radio source at radio galaxy and galaxy cluster scales. We then propose a novel method for scaling beyond heuristics-based techniques, exploring ways to improve the interaction between humans and machine learning to increase scientific discovery. This is followed by the concluding Chapter.

Chapter 4

The study of extended emission in a radio galaxy detected in the LOFAR Two-metre Sky Survey

This Chapter of the thesis focuses on the multi-frequency radio analyses of the radio galaxy PGC2285791 and its diffuse radio emission. The novelty of the work was in identifying 11 extended radio sources (using the heuristic based search from the previous Chapter) that had low-brightness extended features which were previously not studied at low frequencies. It ties into the global structure of the thesis, by providing a particular example of how a discovered radio galaxy with diffuse emission can be characterised. PGC2285791 was of particular interest since it had previously only appeared in source catalogues without characterisation. We then characterised it to be an extended star-forming radio galaxy exhibiting bubble-like (and previously undetected diffuse) radio lobe structures, with an angular extent of giant radio galaxies.

Note: The Chapter is adapted from the published work by: S. A. Thwala, N. Shafi, S. Colafrancesco, F. Govoni and M. Murgia, titled "*The study of extended emission in a radio galaxy detected in the LOFAR Two-Metre Sky Survey* ". It was submitted for publication to the Monthly Notices of Royal Astronomy journal and appears on page 1938 of volume 485, published in 2019.

For this work, S. A. Thwala searched and identified the extended radio sources of interest. He authored the paper, produced all of the images and carried out suggested revisions by the co-authors (and referees). N. Shafi co-supervised the project, edited drafts, carrying out the stellar population properties analysis and contributed to the Discussion and Conclusions. S. Colafrancesco supervised, suggested interesting sources for the project (narrowed the count from 20+ sources to the final 11), motivated the study, and in consultation with F. Govoni and M. Murgia contributed to the Multi-ν, Spectral index map, Discussion and Conclusions.

Abstract

We performed a search for extended radio sources in the LOFAR Two Meter Survey HETDEX Spring Field, and we were able to identify 11 low-brightness extended sources. In this Chapter we focus on the poorly studied radio galaxy PGC2285791, presenting a detailed analysis of its core, its two extended radio lobes, a spectral index map between 141 and 1400 MHz and the multi-frequency association with its IR and optical counterparts. Our results led to the identification of this source with a $z = 0.14$ extended radio galaxy exhibiting bubble-like radio lobe structures ($\sim$ 6 arcmin in the S-N direction).

4.1 Introduction

Low-frequency radio continuum observations of (radio) galaxies are very important because the associated synchrotron emission probes the aged and low-energy electrons, which are less affected by energy losses, and therefore, can propagate further away from their site of origin (assuming that the diffusion coefficient does not vary with particle energy) (Mulcahy et al., 2014). Depending on the magnetic field strength and spatial extension, one would expect to see a large radio synchrotron emission in extended sources (Mulcahy et al. 2014; see also the case of the $\sim$ 700 kpc remnant radio galaxy J1828+49 studied with the Low-Frequency Array (LOFAR) Brienza et al. 2016; van Haarlem et al. 2013).

This is in fact the case for radio galaxies with jets or outflows: AGN-powered radio galaxies have relativistic jets extended and collimated for tens to hundreds of kpc. Outflows are observed to be populated by lower energy electrons which produce irregularly shaped radio lobes that typically extend for smaller distances from the galaxy core.

With the mechanism of particle (re)acceleration and physical processes that produce radio outflows still open to speculation, these sources are ideal candidates for probing the relativistic and magnetized plasma dynamics of the cosmic ray electrons with the host galaxy's interstellar medium (ISM).This interaction, between the core of the source and the ISM, can also probe the energy associated with large scale gas outflows driven by high energy phenomenon in an AGN (e.g. Fanti & Fanti, 1994; O'Dea et al., 2002; Holt et al., 2006; Holt, Tadhunter & Morganti, 2008). Therefore, radio galaxies that exhibit overlapping features between AGN and outflow properties are crucial to filling in the gaps in our understanding of galaxy formations.

Studying the morphologies of radio lobes can give insights into the interaction between the beamed plasma emanating from an AGN and the surrounding ambient medium (Subrahmanyan, Saripalli & Hunstead, 1996). Since synchrotron emission follows a power-law, it is believed that the ambient medium has a higher contribution at lower frequencies. The morphology of radio sources at different frequencies can then be a tracer of the plasma's dynamical states. Using spectral index analysis we can further probe the acceleration and cooling mechanisms that operate on the cosmic ray electrons of this plasma and/with the ambient medium.

Thus, the improved sensitivity and resolution of modern low-frequency radio continuum observations can provide powerful tools to study the evolution and relation between the core of radio continuum sources, the cosmic ray electrons driving the lobes and the ambient medium that surrounds the source. They also can provide detailed insights into the processes that regulate these different regions and their evolutionary time-scales.

In order to search and study new extended emissions around radio galaxies, we used the LOFAR Two-metre Sky Survey (LoTSS; Shimwell et al. 2017) in the still poorly explored low-frequency region of the electromagnetic spectrum.

Exploring the HETDEX Spring Field (see Chapter 4.2 for details), we were able to identify 11 extended sources for our study in a region of $9.5^{\circ} \times 8.7^{\circ}$, an area of the sky comprised between α_{J2000}: $10^{h}37^{m}00^{s}$ to $11^{h}48^{m}00^{s}$ and δ_{J2000}: $+44^{\circ}19'00''$ to $+52^{\circ}55'00''$.

In this Chapter, we present the analysis of the diffuse emission from these sources at low-frequency, and focus on the properties and the nature of a specific radio galaxy, PGC2285791, which shows a bubble-like symmetric radio lobe structure with a prominent radio core. Our study will use the improved sensitivity of the available radio surveys (in particular the LoTSS) and their ability to better resolve radio galaxies for studies of magnetic fields, cosmic ray diffusion and particle acceleration mechanisms in their jets/lobes.

The structure of this Chapter is as follows: we describe in Chapter 4.2 the different data sets we use in our analyses and in Chapter 4.3 we describe the data analysis techniques. We discuss in Chapter 4.4 the identification of the extended radio sources extracted from the LoTSS and we study in particular the nature of the extended radio galaxy PGC2285791. We then discuss our results and draw our conclusion in the final Chapter 4.5.

We adopt a flat, vacuum-dominated cosmology with H_0 = 73 km/s/Mpc, $\Omega_{\rm matter}$ = 0.27 and $\Omega_{\rm vacuum}$ = 0.73 .

4.2 Data

We searched for extended sources in the LoTSS in order to characterise the morphology, integrated spectra and spectral index maps of poorly known or unidentified extended radio sources. We then looked at the different available imaging surveys and used the open data across multi-frequency in order to identify and characterise these sources. Specifically, in our analysis we used publicly available data from the Faint Images of the Radio Sky at Twenty-cm (FIRST; Becker et al. 1995), LoTSS, NRAO Very Large Array (VLA) Sky Survey (NVSS; Condon et al. 1998), TIFR GMRT Sky Survey (TGSS ADR1 (First Alternative Data Release); Intema et al. 2017), VLA Low-frequency Sky Survey Redux (VLSSr; Lane et al. 2014), Wide-field Infrared Survey Explorer (WISE; Wright et al. 2010), ESO Online Digitized Sky Survey (ESO-DSS2-red) and Sloan Digital Sky Survey (SDSS; Blanton et al. 2017) surveys.

The differences in resolutions and sensitivities across these surveys necessitated a careful selection criteria for deciding which surveys to use in this study. For instance, the resolution of the TGSS ADR1 is an ideal match to the LoTSS survey, however, the integration time per pointing for TGSS ADR1 is 15 minutes while LoTSS has 8 hours, thus reducing the sensitivity of TGSS ADR1 to faint extended emission features.

These differences make it challenging to correctly characterise the morphology of the extended radio sources and hinder the creation of reliable spectral index maps (since some of the structures which are present in the LoTSS are not detected to a similar extent in the TGSS ADR1). The difference in sensitivities also had to be factored in, since most of the sources have brightnesses below 0.1 Jy/beam, which is the typical noise level for the VLSSr.

With these restrictions in mind, we opted to use the LoTSS survey in conjunction with the NVSS survey to characterise the morphology of the extended sources and create spectral index maps, seeing that both surveys can resolve extended sources well. However, we noted that the resolutions of the two surveys do not match perfectly, with the NVSS survey having a lower angular resolution of $45''$ compared to the higher angular resolution of $25''$ for the LoTSS survey. To mitigate this difference we smoothed the LoTSS images to match the angular resolution of the NVSS images. In the following we will briefly describe each one of the data sets used in our analysis.

The LoTSS is a deep imaging survey that has been obtained using a compact core and long baselines at the LOFAR observatory observed in the frequency range of 120-168 MHz (Shimwell et al., 2017). This design allows for observations with excellent sensitivity to both extended and compact emissions, producing images that are significantly more sensitive than those produced by other existing large-area low-frequency surveys (Shimwell et al., 2017).

TABLE 4.1: Summary of the continuum images of PGC2285791 used in this work.

Survey	Frequency	Synthesized beam size	$\sigma_{\rm rms}$	Obs. date
	(MHz)	($''$)	(mJy/beam)	
(1)	(2)	(3)	(4)	(5)
VLSSr	73.79	75	71.15	2003 Sep. 20
LoTSS	141.5	25	0.58	2014 Jun. 13
TGSS ADR1	147.5	25	2.99	2016 Mar. 15
NVSS	1400	45	0.57	1993 Nov. 15
FIRST	1400	5	0.18	1997 Mar. 25

The products we used from the survey are the preliminary released images, which are created using a fully-automated but direction-independent calibration strategy (Thompson, Moran, & Swenson, 2017), observed at 141 MHz with a resolution of 25$''$ and typical noise levels of less than 0.5 mJy/beam (Shimwell et al., 2017). Most of the following investigations were largely guided by observations from this survey, and all the other data were used as follow ups/complements to this imaging survey.

The survey covers an area over 350 square degrees, in the region of the HETDEX Spring Field ($\alpha_{\rm J2000}$: $10^{\rm h}45^{\rm m}00^{\rm s}$ to $15^{\rm h}30^{\rm m}00^{\rm s}$ and $\delta_{\rm J2000}$: $+45°00'00''$ to $+57°00'00''$) (Shimwell et al., 2017). Within this area, our search was focused on the P7-Mosaic (with matching coordinates of the sky area provided in Chapter 4.1). The P7-Mosaic was partitioned into smaller regions that cover 3 square degrees to make it easier to survey the region for interesting sources. We had a particular interest in the region limited by $\alpha_{\rm J2000}$: $10^{\rm h}50^{\rm m}00^{\rm s}$ to $11^{\rm h}35^{\rm m}00^{\rm s}$ and $\delta_{\rm J2000}$: $+45°00'00''$ to $+52°00'00''$, since we initially noted the existence of the extended radio source PGC2285791 that is not characterised well previously. We searched for low surface brightness extended radio sources in this region and were able to identify the LoTSS extended radio sources listed in Table 4.2.

The FIRST, NVSS, and VLSSr radio surveys were conducted using the 27 reconfigurable radio antennas of the VLA. The FIRST survey was carried out at a frequency of 1400 MHz using the B configuration of the VLA, producing images with an angular resolution of 5$''$, typical noise levels that are less than 0.13 mJy/beam and a survey flux density threshold of about 1 mJy (Becker et al., 1995). We used the products from this survey to locate the source positions of the optical counterparts and study the integrated spectra analysis for the radio sources selected from the LoTSS.

The NVSS survey was conducted at a frequency of 1400 MHz using the compact D and DnC configurations of the VLA to produce continuum and linear-polarization images with angular resolution of $45''$, typical noise levels that are less than 0.5 mJy/beam and a point source flux density limit of about 2.5 mJy (Condon et al., 1998). We used the products from this survey to study the extent of the diffuse emission, perform integrated spectra analysis, characterise the morphology and create spectral index maps of the extended radio sources.

The VLSSr is an imaging survey which used the VLSS catalogue as a sky model to correct for the ionospheric distortions in the data and then created a new set of sky maps and a corresponding catalogue at 73.8 MHz (Lane et al., 2014). This survey used the B and BnA configurations of the VLA to produce images with angular resolution of $75''$ and typical noise levels of less than 0.1 Jy/beam (Lane et al., 2014). In this work we used the products from this survey to study the integrated spectra for the radio sources selected from the LoTSS.

The TGSS ADR1 is an imaging survey of the 150 MHz radio sky, using the GMRT, aimed at producing a reliable compact source survey (Intema et al., 2017). The products used from the survey have a resolution of $25''$ and typical noise levels of less than 5 mJy/beam (Intema et al., 2017). These were then used to perform integrated spectra analysis, because the short integration time could not map out the diffuse emission sufficiently well.

4.3 Data Analysis

Radio continuum images were downloaded for each of the surveys considered here [1] [2] [3] [4] [5].

Once the sources were identified as likely astrophysical sources, we used the Python Blob Detector and Source Finder (PyBDSF), a tool designed to decompose radio interferometry images into sources and make available their properties for further use (Mohan & Rafferty, 2015). Apart from the source finding purposes, we also used this tool to calculate the rms values, perform basic error data analysis and calculate the integrated flux density for each individual data cube. The rms value obtained for each continuum image was then used as the sigma clipping levels for that corresponding image in our analysis (see discussion in Chapter 4.4 below).

[1] https://third.ucllnl.org/cgi-bin/firstcutout
[2] http://lofar.strw.leidenuniv.nl
[3] http://www.cv.nrao.edu/nvss
[4] http://tgssadr.strw.leidenuniv.nl
[5] http://www.cv.nrao.edu/vlss/VLSSpostage.shtml

To regrid our image datasets to the same grid, we used MIRIAD a data reduction and analysis software (Sault, Teuben & Wright, 1995). We further used this data reduction package to smooth the LoTSS beam to the NVSS beam, to better match the angular resolution of the two surveys. Both of the above mentioned steps were carried out in preparation for the derivation of the spectral index map (see Chapter 4.4.1.5).

4.3.1 Images

Most of the image calibration, cleaning, systematic checks (e.g., RFI flagging) and direction-dependent calibration (DDC) were carried out by the respective survey teams. The only exception to this treatment is that the direction-dependent calibration step still needs to be carried out for the LoTSS, and this will be part of a proposed future data release (Shimwell et al., 2017).

However, looking at the noise levels and image resolution of the different surveys considered here, we decided to use the LoTSS without performing the direction-dependent calibration because as the sigma clipping levels are increased most of the observed artefacts are clipped. This procedure however, does not completely fix the errors in the shape of the sources and thus introduces an error at the edges of our spectral index maps (see our discussion of the spectral index map shown in Fig. 4.5).

We note that for regions closer to the core of the sources analysed, the results are more reliable because the region is detected from both surveys, i.e., the NVSS and LoTSS surveys in this instance. However, as the emission moves towards the edges of the extended emission region, the results serve as a guide since some of the telescopes are not sensitive to faint emissions and some surveys do not resolve the sources well enough.

In this respect, our study can be considered as a path-finder study for the search and characterisation of extended radio sources where we kept in mind that some of the details of the peripheral extension of the radio sources we are interested in may be affected by DDC effects.

To create a more complete astrophysical characterisation of the selected sources, we also looked at the counterparts of the sources in the infrared and optical wavelengths using the WISE, ESO-DSS2-red and SDSS catalogues, respectively. This analysis was then extended to perform cross-correlations at multi-wavelengths to check the consistency of our results with the information available in the literature.

Finally, in order to obtain the stellar population properties of the specific radio galaxy PGC2285791, we applied a full spectral fitting code, and derived

TABLE 4.2: Data on the studied extended sources detected by *LoTSS*. Column (1): Radio source names; Column (2)-(3): Right Ascension (J2000) and Declination (J2000); Column (4)-(8): Integrated flux densities of the extended radio sources; Column (9): Two-frequency radio spectral index between 141 (LoTSS) and 1400 MHz (NVSS).

			Integrated flux density (mJy)					Spectral index [a]
Source Name	R.A. (J2000)	Dec (J2000)	VLLSr 73.8 MHz	LoTSS 141 MHz	TGSS ADR1 150 MHz	NVSS 1400 MHz	FIRST 1400 MHz	α_{141}^{1400}
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Source 01	10 51 33.43	+50 12 53.84	305.41 ± 117.0	247.71 ± 24.77	91.84 ± 9.19	46.35 ± 4.63	24.24 ± 2.43	-0.73 ± 0.06
Source 02	11 00 48.28	+45 51 41.83	$300.00^b \pm 45.00$	154.21 ± 15.42	15.52 ± 1.56	10.91 ± 1.09	1.49 ± 0.15	-1.16 ± 0.06
Source 03	11 01 29.19	+46 00 27.50	701.96 ± 70.20	385.87 ± 38.59	310.34 ± 3.57	5.13 ± 1.64	8.39 ± 0.84	-1.88 ± 0.15
B31059+460	11 02 50.47	+45 44 38.40	394.47 ± 39.45	454.10 ± 45.41	380.72 ± 38.07	76.04 ± 7.60	81.16 ± 8.12	-0.78 ± 0.06
Source 05	11 03 39.37	+45 50 08.58	$300.00^b \pm 45.00$	233.14 ± 23.31	78.96 ± 7.90	16.87 ± 1.69	9.45 ± 0.95	-1.15 ± 0.06
PGC2285791	11 04 32.99	+46 42 24.83	$300.00^b \pm 45.00$	301.20 ± 30.12	28.40 ± 10.56	45.93 ± 4.59	2.10 ± 0.25	-0.82 ± 0.06
Source 07	11 07 57.85	+51 20 23.88	$300.00^b \pm 45.00$	315.70 ± 20.50	172.35 ± 17.23	4.31 ± 0.43	3.03 ± 0.30	-1.87 ± 0.05
Source 08	11 09 41.77	+44 58 27.48	$300.00^b \pm 45.00$	204.99 ± 20.50	$15.00^b \pm 1.50$	32.95 ± 3.29	30.22 ± 3.02	-0.80 ± 0.06
PGC2396883	11 11 57.61	+51 36 21.30	$300.00^b \pm 45.00$	175.54 ± 17.55	48.63 ± 4.87	35.70 ± 3.57	12.87 ± 1.28	-0.69 ± 0.06
Source 10	11 27 13.17	+51 13 25.91	$300.00^b \pm 45.00$	140.00 ± 14.00	35.61 ± 3.04	16.43 ± 1.20	3.82 ± 0.38	-0.93 ± 0.05
Source 11	11 32 50.56	+50 57 04.25	$300.00^b \pm 45.00$	99.28 ± 9.93	39.01 ± 3.35	14.65 ± 1.46	9.08 ± 0.91	-0.83 ± 0.06

[a] The convention, $S \propto \nu^{\alpha}$, was used to calculate the spectral index, where S is the radio flux density, ν is the frequency, and α is the spectral index.

[b] The calculated 3σ upper limit, from the respective survey, was used for instances where the radio survey provided a weak detection.

age, metallicity, star formation rate (SFR) and black hole mass (see Chapter 4.4.1.4).

4.4 Extended Sources

Our strategy for the deep search was to look for faint extended sources (as detailed in Chapter 3.3), since they would have not been picked up by previous studies. The sources found in our search are listed in Table 4.2. There are in total three sources which have an identification in the literature, namely: PGC2285791, PGC2396883 and B31059+460. These sources, as well as the eight remaining sources, have only little information available in the literature, which we detail below.

- Source 01 has been classified as a flat-spectrum source (Massaro et al., 2014), while it also appears in the SPECFIND V2.0 catalogue tool (Vollmer et al., 2010) and the 5C2 survey (Pooley G. G. & Kenderdine S., 1968).
- We found that the sources 02 and 11 have no prior information available in the literature.
- Source 03 only appears in (Vollmer et al., 2010)(The SPECFIND V2.0 catalogue of radio cross-identifications and spectra. SPECFIND meets the Virtual Observatory).
- The source B31059+460 appears in various catalogues described in (Ficarra, Grueff & Tomassetti, 1985; Hales, Baldwin & Warner, 1988; Becker,

White & Edwards, 1991; Gregory P. C. & Condon J. J., 1991; Green & Riley, 1995; Gregory et al., 1996; McMahon & White, 2002; Cohen et al., 2007; Vollmer et al., 2010; Farnes, Gaensler & Carretti, 2014).

- Both of the studies (West et al., 2008, 2011) associated with the source 05 are from the SDSS spectroscopic catalogue, which classify this source as an M-Dwarf star.

- The source 06 is associated with the galaxy PGC2285791 (Best & Heckman, 2012; Paturel et al., 2003), and due to its interesting extended radio emission it is the main subject of our study, and discussed in detail in Chapter 4.4.1.

- Source 07 and 10 have been identified as optical galaxy clusters, using an adaptive matched filter technique described in (Szabo et al., 2011).

- Source 08 is a $z = 0.31$ radio galaxy that seems to be in the same field of view as a well studied QSO source at $z = 2.22$.

- The source 09 is associated with the galaxy PGC2396883 which has been part of a few statistical studies (Barber, Meiksin & Murphy, 2007; Lin et al., 2010) and also appears in the following catalogues (Maslowski et al., 1984; Paturel et al., 2003; Best et al., 2005; Best & Heckman, 2012; Farnes, Gaensler & Carretti, 2014).

The integrated flux density uncertainties were computed using the relation:

$$\sigma_{S_\nu} = \sqrt{(\sigma_{\rm cal} S_\nu)^2 + \left(\sigma_{\rm rms}\sqrt{N_{\rm beam}}\right)^2}\,, \tag{4.1}$$

where $\sigma_{\rm cal}$ is the uncertainty in the calibration of the flux scale, S_ν is the integrated flux density at the observed frequency ν, $\sigma_{\rm rms}$ is the rms of the image and $N_{\rm beam}$ is the number of independent beams in the source (Cantwell et al., 2016). $\sigma_{\rm cal}$ is taken to be 10% for the LoTSS, NVSS and TGSS ADR1 surveys (Shimwell et al., 2017; Condon et al., 1998; Intema et al., 2017) and 15% for the VLSSr survey (Lane et al., 2014). The $\sigma_{\rm rms}$ was calculated for each source, using continuum images with a field of view of 10 arcmin for each survey. In cases where sources had weak detection signals, $\sigma_{\rm rms}$ was used to compute the calculated 3σ upper limit for that respective survey (see Table 4.1). We notice that most of the extended sources (8 out of 11) previously unknown, were not previously studied in detail.

We will discuss in this Chapter the nature of the specific source PGC2285791 which is an extended radio galaxy that is poorly known, while we will present our analysis of the remaining radio sources in a forthcoming paper (Colafrancesco et al. in preparation).

4.4.1 PGC2285791

The source PGC2285791 (α_{J2000}: $11^{h}04^{m}32^{s}.99$; δ_{J2000}: $+46^{\circ}42^{'}24^{''}.83$) is seen as an extended radio source in the LoTSS (see Fig. 4.1). This source is poorly characterised in the literature and it is listed as a radio galaxy source in Best & Heckman (2012); Paturel et al. (2003), while it is also identified as a low-redshift source with ultra-steep spectrum in Bornancini et al. (2010). This is the only available information for this source in the literature.

It is justified, therefore, to perform an extensive analysis of this extended radio source based on the LoTSS data.

4.4.1.1 Morphology

The radio images of PGC2285791 derived from the LoTSS at 141 MHz and the NVSS at 1400 MHz are shown in Fig. 4.1.

This radio galaxy has two diffuse bright lobes forming a double structure with LoTSS mean surface brightnesses of 3.96 mJy/beam and 2.40 mJy/beam for the northern and southern lobes, respectively. These lobes are quite irregular, at variance with typical lobes of powerful radio galaxies connected by collimated jets, spanning an overall angular extension of $\sim$ 6 arcmin which corresponds to a projected linear size of $\sim$ 0.86 Mpc.

The source is also seen to have a similar morphology at 1400 MHz, as shown in Fig. 4.1. However, we notice that in the LoTSS image the core of the source is brighter and two additional compact sources (or knots) appear in the northern lobe. These two knots have peak LoTSS brightness levels of 8.19 $\pm$ 1.00 mJy/pixel and 15.26 $\pm$ 1.63 mJy/pixel for the top and bottom knots, respectively, where the pixel size is 11.83 arcsec^2. The region of the northern lobe containing these radio knots has quite steep spectral index (see Fig. 4.5 and discussion in Chapter 4.4.1.5 below).

4.4.1.2 Integrated spectra

The integrated flux emitted by the core of the radio galaxy is $S_{core(141MHz)} = 8.74 \pm 1.11$ mJy and $S_{core(1400MHz)} = 3.40 \pm 1.07$ mJy. Its radio luminosity is, therefore, $L_{141MHz} \sim 4.35 \times 10^{23}$ W Hz^{-1} and $L_{1400MHz} \sim 1.69 \times 10^{23}$ W Hz^{-1} using the source's luminosity distance of 645 Mpc in the adopted cosmology.

The core radio flux density spectrum of the radio galaxy is shown in Fig. 4.2. We notice, however, that the spatial resolution of the five radio surveys from which we derived the core flux are quite different, with the VLSSr having the lowest resolution 75$^{''}$ (the beam size is three times larger than the LoTSS beam), the NVSS having a moderate resolution at 45$^{''}$ (the beam size is 1.8 times larger than the LoTSS beam) and the FIRST has the highest resolution

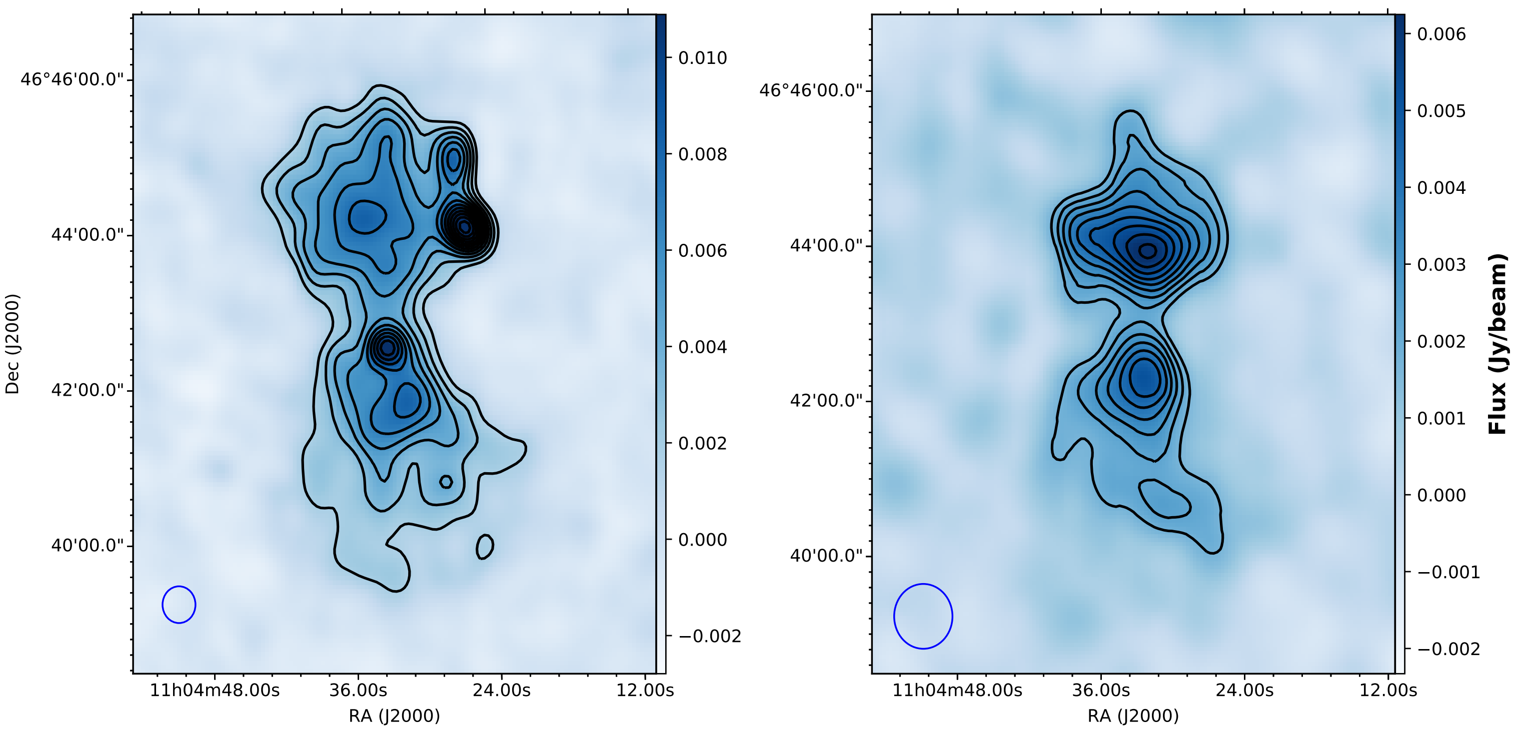


FIGURE 4.1: Left: The LoTSS 141 MHz continuum image of PGC2285791 overlaid with black LoTSS contours starting at 3σ (σ = 0.58 mJy/beam) with 2σ increments. The beam size of the LoTSS is shown in the lower left corner. Right: The NVSS 1400 MHz continuum image of PGC2285791 overlaid with black NVSS contours starting at 3σ (σ = 0.57 mJy/beam) with 1σ increments. The beam size of the NVSS is shown in the lower left corner.

at $5''$ (the beam size is 5 times smaller than that of the LoTSS beam).

The two knots have an integrated flux of $S_{\mathrm{nk(141MHz)}} = 11.26 \pm 1.51$ mJy for the northern knot, and $S_{\mathrm{sk(141MHz)}} = 20.53 \pm 1.42$ mJy for the southern knot.

4.4.1.3 Multi-ν analysis

The core of the radio source PGC2285791 is coincident with an optical and IR source in the ESO-DSS2-red and WISE images (see Fig. 4.3): this optical source is extended and has the appearance of a galaxy (green square in Fig. 4.3). The galaxy nature of this source is confirmed by the image and the spectrum of the source as found in the SDSS DR14 (Abolfathi et al., 2018), where it is classified as a star-forming galaxy at redshift $z = 0.14126 \pm 0.00001$. Its optical spectrum is characterised by a number of prominent emission lines, such as, H_{ff}, [NII], [SII], as well as the forbidden oxygen lines [OII] and [OIII]. The source is also associated with a radio galaxy in the QORG catalogue (Flesch

TABLE 4.3: List of flux densities and their respective errors of PGC2285791 for the core and source at different frequencies.

Frequency	Flux density	
	$S_{\rm core}$	$S_{\rm source}$
(MHz)	(mJy)	
(1)	(2)	(3)
73.79	$213.45^{6} \pm 32.02$	213.45 ± 32.02
141.5	8.74 ± 1.11	301.20 ± 30.12
147.5	13.02 ± 3.04	28.40 ± 10.56
1400	3.40 ± 1.07	45.93 ± 4.59
1400	2.10 ± 0.25	-

et al., 2012) retrieved from the SSDC Sky Explorer tool [7] denoting its radio galaxy nature.

It is also interesting to notice the presence of two optical sources that lie within the region of the northern lobe (highlighted by yellow and cyan squares in Fig. 4.3). They have confirmed spectra from SDSS DR14, with the top source being identified as a galaxy at $z = 0.11452 \pm 0.00003$ (yellow square) and the bottom source identified as PGC2286197 (cyan square), an AGN galaxy at $z = 0.11451 \pm 0.00003$ with a FIRST radio source in the SSDC Sky Explorer. Even though these two sources are at a comparable redshift to PGC2285791, they appear to be projections that are not interacting with PGC2285791. No optical or IR sources were found at the position of the two radio knots. We will discuss the spectral behaviour of these two knots in Chapter 4.4.1.5 below.

The overall multi-ν spectral energy distribution (SED) of the radio galaxy core, obtained from the combination of archival data and of the radio data obtained from this study, is shown in Fig. 4.4.

The SED of the core is clearly dominated by non-thermal emission at low frequencies while it is dominated by the integrated galaxy emission in the IR and Optical frequency range peaking at $\sim 5 \times 10^{14}$ Hz, and it is quite typical of radio galaxy spectra (Colafrancesco et al., 2016). The lack of high frequency data in the soft and hard X-ray ranges does not yet allow us to test the inverse Compton (on synchrotron) branch of the core emission. X-ray data would also be crucial in identifying possible inverse Compton scattering X-ray emission from the radio galaxy lobes.

[7] https://tools.asdc.asi.it/index.jsp

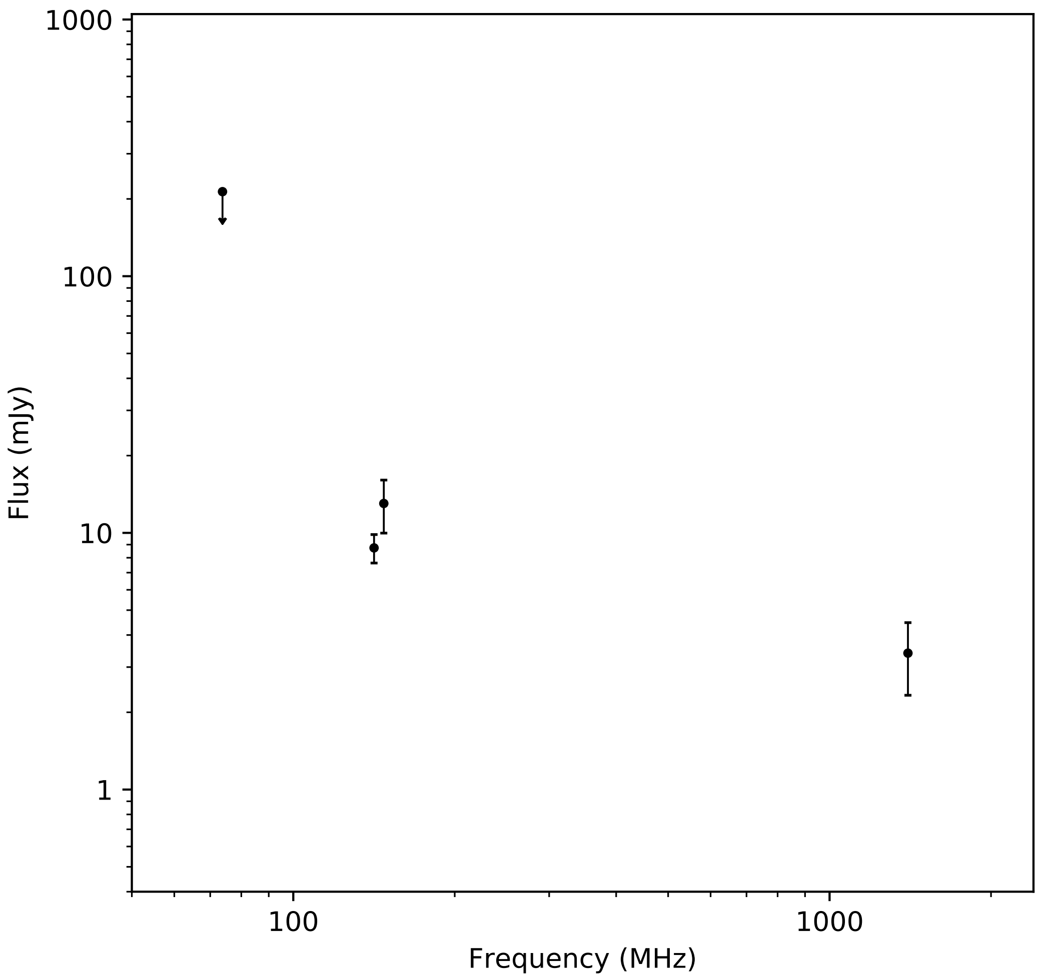


FIGURE 4.2: The spectrum of the core of the radio galaxy PGC2285791 is shown for different frequencies, with the frequency-flux pairs (Frequency, S_{core}) given in Table 4.3.

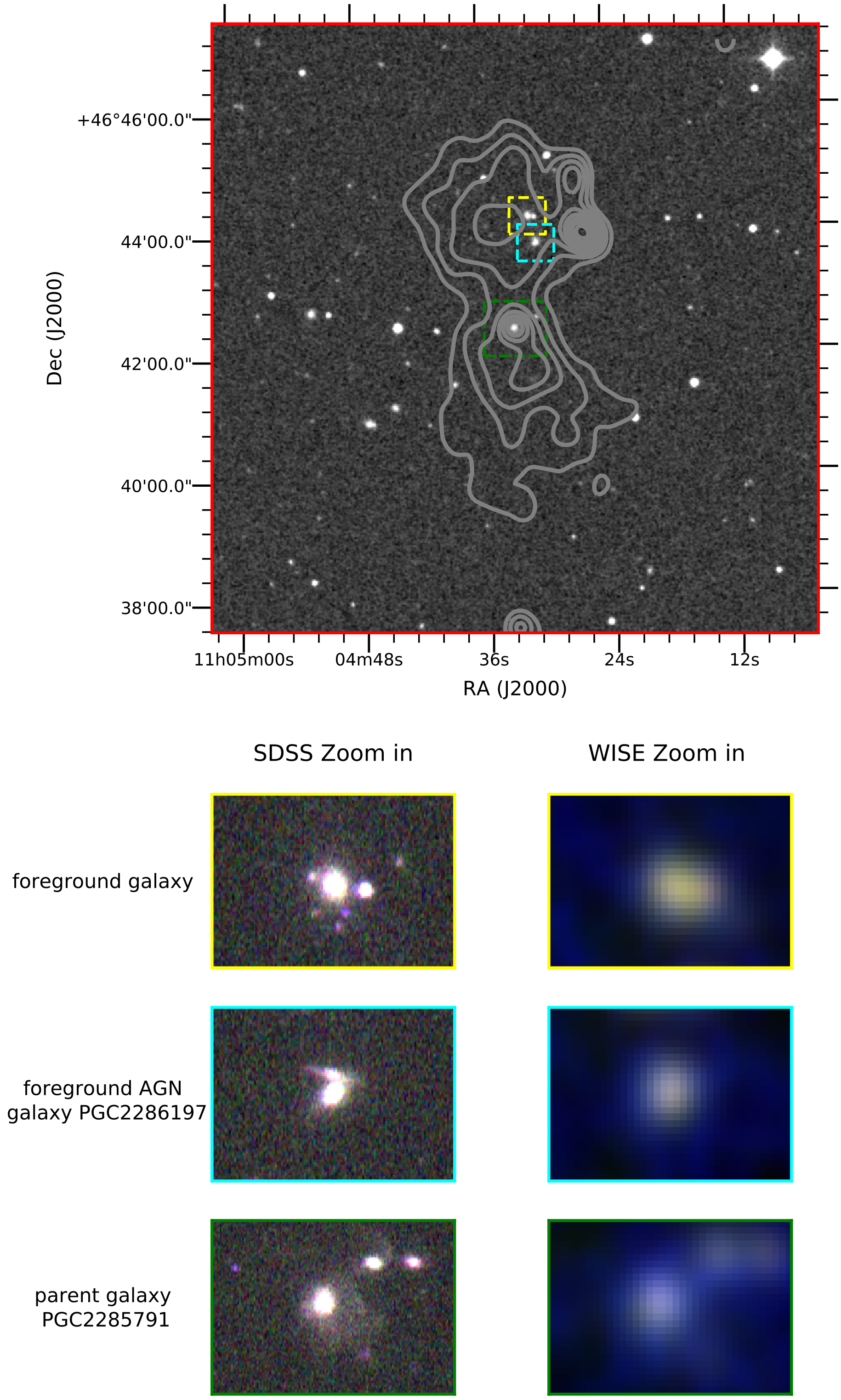


FIGURE 4.3: Main: background ESO-DSS2-red image of PGC2285791 overlaid with gray LoTSS contours at [1.74, 3.48, 5.22, 6.96, 8.7, 10.44, 12.18, 13.92, 15.66] mJy/beam. The green, yellow and cyan dashed squares are the regions of interests (to be) highlighted by the SDSS i, r, g-band images with corresponding frame colours (below on the left) and the WISE images from the W1, W2 and W3 bands (on the right). Yellow squares: SDSS and WISE zoom-in views of the foreground galaxy. Cyan squares: SDSS and WISE zoom-in views of the foreground AGN galaxy PGC2286197. Green squares: SDSS and WISE zoom-in views of the parent galaxy PGC2285791.

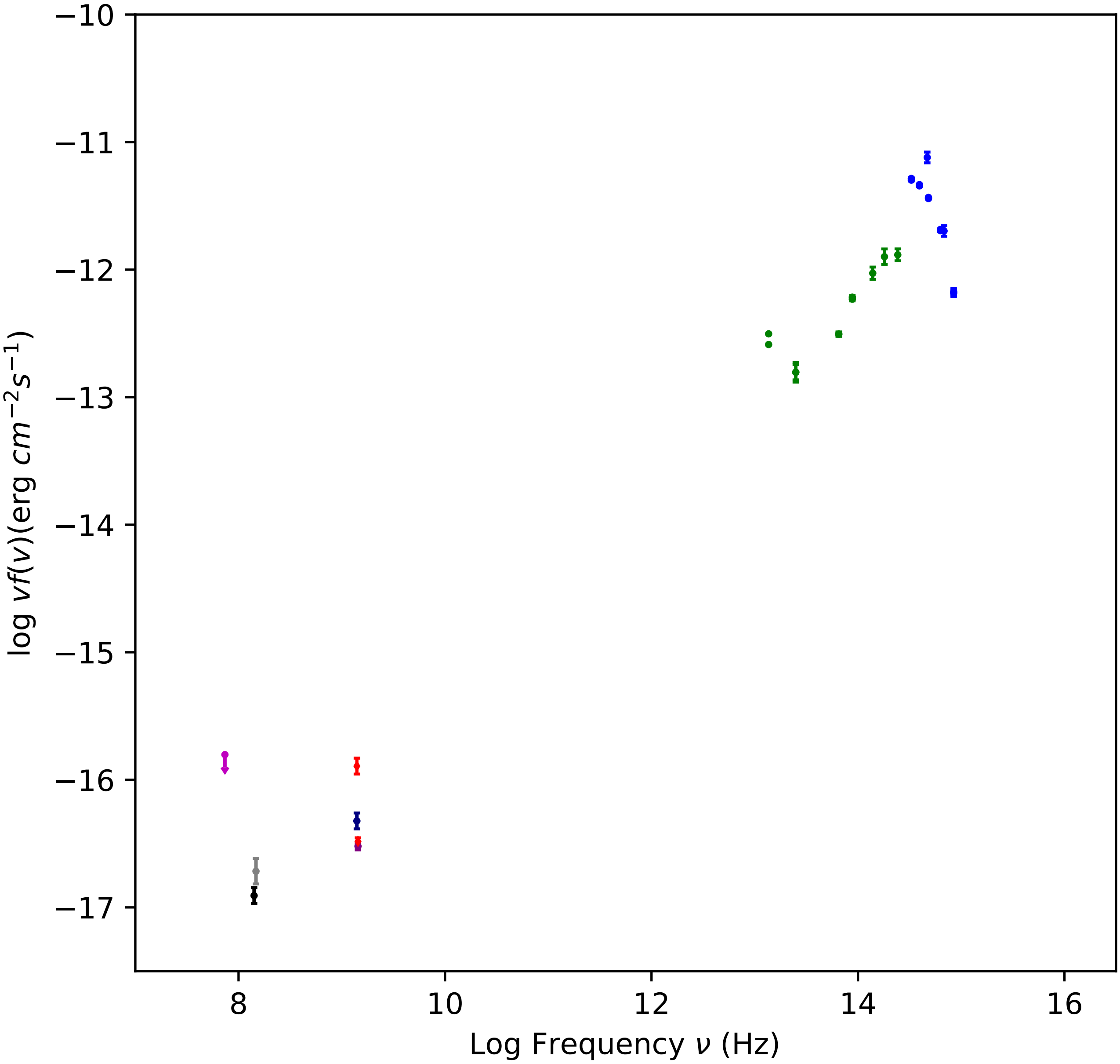


FIGURE 4.4: The SED of the core of the radio galaxy PGC2285791 with previously known optical (blue), infrared (green) and radio (red) data points plotted with the addition of the VLSSr upper limit (magenta), LoTSS (black), TGSS ADR1 (grey overlaps with black), NVSS (navy) and FIRST (purple overlaps with red archival data) data points. There are four points in the radio domain that have the same value for the frequency. Starting from the top; the red point is the archival NVSS data point, navy is the NVSS value from this work and the FIRST purple point from this study overlaps with the FIRST archival data point.

4.4.1.4 Stellar population properties

We modelled the stellar population of PGC2285791 spectrum obtained from the SDSS DR12 using the full spectral fitting code FIREFLY (Wilkinson et al., 2017) and masking all the emission lines. Full details about the modelling process are given in several papers (Wilkinson et al., 2015, 2017; Goddard et al., 2017a,b; Comparat et al., 2017). In short, FIREFLY is a chi-squared minimisation fitting code that, for a given input SED, fits combinations of simple stellar populations based on Maraston & Strömbäck (2011) with various input spectral libraries and Initial Mass Functions (IMFs). For our analysis here, we only show stellar population properties (such as age, metallicity, SFR and black hole mass) derived based on MILES (Sánchez-Blázquez et al., 2006) stellar library with a Charbier (Chabrier, 2003) IMF, since it is the best fit in this case. However, we note that a full analysis of the stellar population components is beyond the scope of this work.

Nonetheless, the derived light-weighted average age and metallicity of PGC 2285791 are 8.7 Gyr and -0.5 in [Z/H], respectively, with smaller contribution from 10 Myr old and solar metallicity stars. The total and the black hole masses are $\sim 5 \times 10^{10}$ $M_{\odot}$ and 2.7×10^{8} $M_{\odot}$, respectively. We have obtained a SFR of ~ 2 $M_{\odot}$ yr^{-1}.

4.4.1.5 Spectral index map

In order to produce spectral index maps, we aligned the images using MIRIAD (Sault, Teuben & Wright, 1995). The higher resolution LoTSS images were smoothed to the lower resolution NVSS images using a Gaussian beam with beam parameters (major="45arcsec", minor="45arcsec") in MIRIAD.

We then derived the flux densities from each matching pixel coordinate obtained from the two images at different frequencies ($S_{141\mathrm{MHz}}$ and $S_{1400\mathrm{MHz}}$) in order to calculate the spectral index at the corresponding point. The following power-law form of the spectral index formula was used, $S \propto \nu^{\alpha}$, where S is the radio flux density, ν is the frequency, and α is the spectral index. The uncertainties were computed using the relation:

$$\sigma_{\alpha} = \frac{\sqrt{\left(\frac{\sigma_1}{S_1}\right)^2 + \left(\frac{\sigma_2}{S_2}\right)^2}}{\ln\left(\frac{\nu_1}{\nu_2}\right)}, \tag{4.2}$$

where σ_1, σ_2 are the uncertainties of the LoTSS and NVSS flux densities, S_1, S_2 are the integrated flux densities at the LoTSS and NVSS frequency bands (ν_1, ν_2), respectively.

The resulting image array was then imaged in python using Matplotlib and APLpy (see Fig. 4.5 for the resulting maps). As we previously cautioned, the

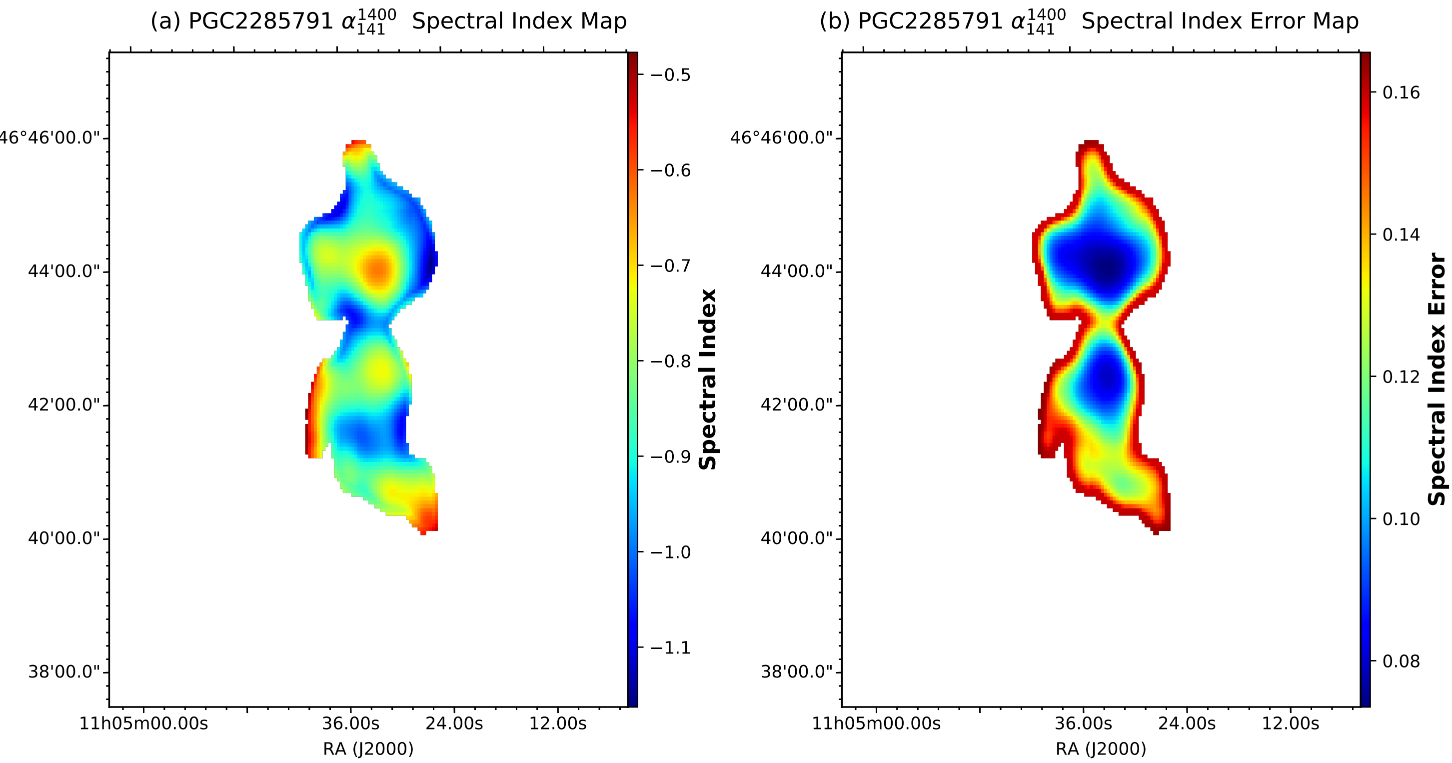


FIGURE 4.5: The spectral index (a) and spectral index error (b) maps of PGC2285791 obtained between 141 and 1400 MHz.

lack of access to raw data is a limiting factor and hence our resulting images have not been corrected for direction-dependent effects, which we will leave to the LoTSS survey team.

A two-frequency radio spectral index image was computed from the maps at 141 MHz and 1400 MHz. The distribution of the radio spectral index α is shown in Fig. 4.5 (a) along with the spectral index error map in Fig. 4.5 (b).

The spectral index of the extended radio source PGC2285791 varies from $\alpha = -1.1$ to $\alpha = -0.5$, while the southern radio lobe shows flat spectra with an average value of $\alpha = -0.7$. The emission of the northern lobe has an island of spectral index values that flattens up to $\alpha = -0.6 \pm 0.06$ surrounded by a steeper spectral region including the two radio knots $\alpha = -1.1 \pm 0.06$ towards the north-western boundary. The derived spectral index values are $\alpha_{\rm nk} = -1.07 \pm 0.06$ and $\alpha_{\rm sk} = -1.12 \pm 0.06$ for the northern and southern knots, respectively. The spectral index map of the northern lobe indicates the presence of external regions with steep spectra ($\alpha \sim -1.1 - -0.9$) which are encasing an inner region with much flatter spectra distribution ($\alpha \sim -0.8 - -0.6$).

The southern lobe seems to have a similar structure: the steep spectral index region ($\alpha \sim -1.0 - -0.9$) is intermediate between the radio galaxy core and the flatter spectral indices region ($\alpha \sim -0.8 - -0.7$) in the outer parts of the lobe. The spectral index distribution of the radio core is relatively steep ($\alpha \sim -0.8 - -0.7$), justifying the decrease of its flux observed from 141 MHz to 1400 MHz.

We note that there are uncertainties in the estimate of the spectral index due to differences in the (u, v) coverage between the two observations; as we can see in the spectral index error map Fig. 4.5 (b) the errors increase slightly towards the edges of the image but are still at a level $< 18\%$.

4.5 Discussion and Conclusions

We have used high resolution LOFAR images to search for low-surface brightness extended sources at low frequencies. This has enabled us to uncover several either unknown or poorly studied low-surface brightness extended radio sources. Here we presented the specific case of PGC2285791. Our search revealed a very interesting radio morphology and spectral distribution of PGC2285791, enabling us to better characterise the source. We produced spectral index distribution maps by combining the LoTSS image at 141 MHz and NVSS image at 1400 MHz. The spectral index of the extended radio source PGC2285791 varies from $\alpha = -1.1$ to $\alpha = -0.5$. An island of a flat spectral region ($\alpha = -0.6$) and hot spots ($\alpha = -1.0$) are observed. Possible interpretations of the spectral-index variation could be density and magnetic field enhancement due to shock compression by external plasma.

The source exhibits two diffuse bubble-like structures (lobes) on both sides of the core with a projected linear size of 0.84 Mpc. Bubble-like radio structures have also been observed in a few radio galaxies (e.g., Centaurus A (Quillen et al., 2006)) and also in a number of Seyfert and starburst galaxies without obvious jets (e.g., NGC 6764 (Croston et al., 2008), Markarian 6 (Mingo et al., 2011), NGC 3079 (Duric & Seaquist , 1988; Shafi et al., 2015)). Several mechanisms have been suggested for the formation of bubble-like structures in these galaxies. A jet-driven outflow (Quillen et al., 2006) and interaction of starburst wind with ISM are among them (Veilleux, Cecil & Bland-Hawthorn, 2005). It is believed that radio galaxies harbour a supermassive black hole ($10^8 - 10^9$ $M_\odot$) in their galactic centre which is responsible for launching large-scale radio jets and lobes (Begelman, Blandford & Rees, 1984). However, how these jets and lobes grow to their large-scale sizes (>100 kpc in cases such as this), and their impact on the surrounding medium is still not fully understood. The radio jet close to the nucleus of PGC2285791 could possibly give rise to the bubbles of non-thermal radio emission that would interact with the external environment (shock) in the host galaxy. However, wide frequency observations are necessary for detailed spectral index maps and to gain further insight into the physical environment of PGC2285791.

From our analysis of the SDSS spectrum, we see that PGC2285791 is approximated by an 8.7 Gyr old stellar population with lower metallicity. This could be because MILES-based results, though they give the best fit, they tend to have their minimum χ^2 values at higher age, but a lower metallicity (Wilkinson et al., 2017). However, if correctly represented, the low metallicity could be reflective of gas outflow and the interplay between the galaxy's ISM and the IGM. On the other hand, star formation may have been triggered by galaxy mergers or interaction, though we did not see any sign of interaction or disturbance.

In addition, in a simplified emission line diagnostics based on different line ratio diagrams in Kewley et al. (2006) classification (see Kewley et al. (2006) and references for details), PGC2285791 seems to fall under composite sources. While the lack of noticeable line-broadening in the optical spectrum suggests a type 2 AGN, the prominent emission line features could be from nuclear starbursts that are a few Myr old. However, unless we have spatially resolved observation, it is often difficult to disentangle the contributions of the two to the emission line spectrum. We note that a better understanding of the star forming and AGN nature of the source and how it affects the physical properties of the source requires a more in depth analysis of the ionization properties as well as the gas kinematics. The next generation radio telescopes such as ASKAP and MeerKAT would be very useful to investigate the physical conditions and kinematics of the gas in low surface brightness radio galaxies.

Finally, our study shows the value of the high-resolution and high sensitivity of LOFAR maps for studies of low surface brightness sources, and in particular radio galaxies with extended lobes.

Chapter 5

Archival radio observations of the Abell 1689 cluster

This chapter of the thesis focuses on the multi-frequency radio analyses of the Abell 1689 cluster and its diffuse radio emission. The radio source was initially subject to dark matter studies and identified (by co-author Khanye N.F.) to be lacking radio continuum analyses. This links into the global structure of the thesis, by focusing on the use of publicly available data to characterise diffuse emission in extended radio sources at galaxy cluster scales.

Note: This chapter is adapted from the ongoing work by: Khanye N. F., Thwala S. A., and Marchegiani P., tittled "*Archival radio observations of the Abell 1689 cluster* ". The work is currently being prepared for submission to the Monthly Notices of Royal Astronomy journal.

The candidate is the co-author of this work and has been responsible for revising the drafts, imaging, co-analyses of the data and spectral index analysis. This work has been guided by the supervision of Dr. Marchegiani and Dr. Beck.

Abstract

Over the course of recent years, our knowledge on the evolution and dynamics of galaxy clusters has improved. This is due to the progress in radio and X-ray surveys, and in the development of theoretical models which try to explain the intra-cluster medium (ICM). These advances in understanding have raised more questions about the origin of cluster magnetic fields, and the origin of large areas of diffuse synchrotron emission, such as radio halos. In this Chapter, we present new radio analyses of the Abell 1689 galaxy cluster, using publicly available archival data from the VLA, GMRT, MWA, SDSS, and WISE surveys. We use the radio observations to help us understand and map out the morphology, and spectral properties of the radio emission in the cluster. The optical and infrared observations are then used to cross match the observed features with their possible optical and infrared counterparts. We calculated the integrated flux density of the radio halo to be 8.7 mJy at 1400

MHz and 202.9 mJy at 147.5 MHz from the NVSS and TGSS observations, respectively. We also, for the first time, present the spectral index image of the cluster using observations from the NVSS and TGSS observations, and find a power-law correlation in the radio spectrum of the diffuse emission, with a spectral index of α= -1.34 (where S$\propto \nu^{\alpha}$).

5.1 Introduction

The radio emission observed in galaxy clusters is usually attributed to the effects of shocks and turbulence that produce (re-)acceleration of relativistic particles in magnetic fields which permeate the cluster atmospheres (Feretti et al., 2012). Clusters contain different sources of radio emissions ranging from active galactic nuclei (AGN) to enormous regions of diffuse radio emission which lack optical counterparts. The large regions of diffuse radio emission are thought to result from the re-acceleration of relativistic particles and can be classified into three classes: (i) radio halos, (ii) radio mini-halos and (iii) radio relics. Radio halos are large low-surface brightness diffuse radio sources which are generally located at the cluster centre with a very steep spectrum ($\alpha \leq$ -1, where S$\propto \nu^{\alpha}$). They have a regular spherical morphology and are mostly found to be unpolarized. Examples of galaxy clusters which host radio halos are Abell 2163 in (Feretti et al., 2001) and Abell 1351 (Giovannini et al., 2009). Radio mini-halos are similar to radio halos, with the differentiating factors being that, they are usually found in relaxed, cool-core clusters, coupled with the extent of the radio emission (around the central radio galaxy) being confined to smaller scales (∼500 kpc, see e.g., Ignesti et al. (2020) and references therein). Examples of galaxy clusters that host mini-halos are Abell 2029 and the Ophiuchus cluster (Govoni et al., 2009). Radio relics are also large low surface brightness regions, however, unlike radio halos they are mostly located in the peripheral regions of the cluster and have an elongated, irregular morphology. Radio relics are usually strongly polarized e.g the Toothbrush cluster (Ogrean et al., 2013) and the double radio relic found in CIZA J2242.8+5301 (van Weeren et al., 2010, 2011).

The origin of diffuse radio emissions in galaxy clusters is still up for debate. This is due to the difficulty in explaining their large sizes and the short life span of radiating electrons, which are vital for their occurrence. A number of models have been introduced to try and explain their puzzling origin. This includes: (i) the widely accepted primary acceleration model, which attributes the occurrence of radio halos to the diffusion of relativistic electrons into the intra-cluster medium (ICM) by different events occurring in the cluster, such as shocks and turbulence from cluster merger events (Rephaeli (1979); Brunetti et al. (2004)) and AGN activities (Jaffe (1977); Ferrari et al. (2008)); (ii) the secondary electron model, which predicts that relativistic electrons may result from the inelastic hadronic collisions between non-thermal and thermal protons from the hot gas in the ICM (Blasi & Colafrancesco, 1999); and (iii) the secondary electron production from dark matter (DM) annihilation model, which predicts the generation of gamma-rays, electrons and positrons during the annihilation of DM particles such as weakly interacting massive particles (WIMPs) (Colafrancesco & Melé (2001); Bertone, Hooper & Silk (2005); Colafrancesco et al. (2006)).

The occurrence of radio relics in galaxy clusters is also highly associated with the presence of merger shocks within the cluster's X-ray gas, which are caused by highly energetic merger events that are capable of re-accelerating cosmic rays. The merger shocks are expected to lead to the amplification of magnetic fields in the cluster and enable the occurrence of large areas of diffuse emission on the outskirts of the clusters, when they come in contact with relativistic electrons. This phenomena can also lead to the formation of shock fronts, which are sharp contact discontinuities between regions of the ICM gas with different entropies. Shock fronts exhibit a sharp gas density edge and a temperature jump (prerequisites for determining the gas velocity and density jumps in the ICM). Shock fronts have been observed in a few galaxy clusters that host diffuse radio emission (Markevitch et al., 2005; Markevitch, 2005).

In this work, we present the radio observations of Abell 1689 (A1689 from here-on) at four different radio frequencies, while also paying attention to the properties of the diffuse radio emission found in the cluster, whose presence was first reported by Vacca et al. (2011). We also provide the spectral analyses of the cluster using an integrated radio spectrum and spectral index images of the cluster, to develop a better understanding of the energy spectrum of the relativistic electrons.

Throughout the Chapter we have employed a flat, vacuum dominated ΛCDM cosmological model with H_0= 69.6 km $s^{-1}Mpc^{-1}$, $\Omega_m = 0.286$, $\Omega_\Lambda = 0.714$. With these parameters the luminosity distance of A1689 at $z = 0.183$ is D_L = 894.6 Mpc and 1 arcmin corresponds to 185.94 kpc.

5.2 Abell A1689

A1689 is a massive optically rich galaxy cluster at redshift $z = 0.183$ (Struble & Rood, 1999). The cluster is well used for strong and weak gravitational lensing studies due to its display of gravitational arcs, arclets and multiple image systems as seen in the optical regime (Broadhurst et al., 2005a). The cluster also hosts the largest Einstein radius known so far (Tyson, Valdes & Wenk, 1990; Taylor & Dye, 1998; Oguri & Blandford, 2009) and further strong lensing studies have also revealed several substructures with different concentrations in the central and north eastern regions of the cluster. These optical properties and the discrepancies in mass, ~ 10, between the X-ray and gravitational lensing observations (Tyson & Fischer, 1995; Shan et al., 2010;

Parameter	value	ref
RA (J2000)	13 11 30.2	-
DEC (J2000)	-01 20 10.6	-
z	0.183	Struble & Rood (1999)
$L_x^{0.1-2.4}$	1.2×10^{45}	Vacca et al. (2011)
$M_{\rm cool}$ ($M_\odot {\rm yr}^{-1}$)	$683\,^{239}_{128}$	Chen et al. (2007)
$t_{\rm cool}$ (Gyr)	4.7 ± 0.4	Chen et al. (2007)
M_{200} ($M_\odot h^{-1}$)	$1.25 \pm 0.3 \times 10^{14}$	Leonard et al. (2011)
$S^{\rm RH}_{\rm 1200\,MHz}$ (mJy)	11.7 ± 3.4	Vacca et al. (2011)
LLS (kpc)	730	Vacca et al. (2011)
$P^{\rm RH}_{\rm 1200\,MHz}$ ($\times 10^{24}$ WHz^{-1})	1.08	Vacca et al. (2011)

TABLE 5.1: Global properties of the A1689 cluster and its radio halo (written as RH) at 1200 MHz.

Serano et al., 2013) are what makes A1689 an interesting laboratory for indirect studies of dark matter (Serano & Umetsu, 2011). XMM-Newton observations compared with ASCA, ROSAT and Chandra data suggest that the cluster is in a post merger scenario, with the core of the cluster not being fully relaxed (Andersson & Madejski (2004); see Vacca et al. (2011) for a comprehensive review of the clusters, X-ray, dynamics and history).

Vacca et al. (2011) reported on the detection of diffuse radio emission in the cluster, at 1200 and 1400 MHz, using the Very Large Array (VLA) in different configurations. The diffuse emission spreads over the central regions of the cluster and is mostly outshined by several radio galaxies found throughout the cluster. The authors estimated the largest linear size (LLS) of the diffuse emission to be $\sim$ 730 kpc and classified the emission as a small radio halo with a central radio brightness of $\simeq 1.7$ μJy/arcsec^2 at 1200 MHz and estimated a flux density of 9.9 mJy at 1400 MHz, for the radio halo. They also investigated the radio halo–X-ray correlation of the cluster, using Chandra X-ray observations, and found that there is a non-linear relation due to the radio radial profile being much flatter than the X-ray profile. The global properties of the cluster are listed in Table 5.1.

5.3 Observations and imaging

To perform the study on the extended radio emission in the galaxy cluster, we used archival radio continuum observations from six radio surveys (Chapter 3.2 provides a comprehensive explanation of the surveys' specifications and data reduction procedures): VLSSr (Lane et al., 2014), TGSS (Intema et al., 2017), GLEAM (Hurley-Walker et al., 2017), NVSS (Condon et al., 1998), FIRST (Becker et al., 1995), and VLASS (Lacy et al., 2020) at their respective central frequencies of 74, 150, 200, 1400, 1400, and 3000 MHz, by retrieving images of the galaxy cluster from their image cutout services. The cutout images were already cleaned and calibrated by the relevant survey teams before

Survey	ν(MHz)	Resolution	$\sigma_{\rm rms}$(mJy/b)	Obs. date
VLSSr	73.5	$75'' \times 75''$	109.75	2003/09/20
TGSS DR1	147.5	$27'' \times 25''$	3.22	2016/03/15
GLEAM-4	200	$142'' \times 132''$	18.97	2014/03/10
NVSS	1400	$45'' \times 45''$	0.53	1995/02/27
FIRST	1400	$5'' \times 5''$	0.15	1998/08/-
VLASS	3000	$3'' \times 2''$	0.17	2019/04/21

TABLE 5.2: Properties of the A1689 radio continuum images used in this study. Col. 1: survey name, col. 2: observation frequency, col. 3 synthesized beam size, col. 4: noise value of each image and col. 5 observation date.

being publicly released, and thus, no calibration (step) was necessary for this study. A summary of all the radio continuum images used in this study is provided in Table 5.2.

5.3.1 Imaging and Analysis

The differences in survey configurations, resolution and sensitivity across the publicly available archival image products requires a detailed systematic approach to matching complementary surveys for extended emission analyses. We made informed decisions about which surveys to match, from the considered data products, by considering the nature of the extended radio and diffuse emission found in proximity of the galaxy cluster.

In light of the fact that diffuse emission has a low-surface brightness, which is typically resolved out in high angular resolution images, we opt for the usage of lower angular resolution surveys (with survey configurations) that are optimised for extended emission, namely TGSS and NVSS in this study, to map out the spectral properties of the relativistic electrons characterising this emission. We note that there are still issues with the differences in integration times, affecting the observed depth and extent of the sources (see Thwala et al., 2019 and Chapter 4.2) which cannot be trivially addressed.

The high angular resolution images with increased sensitivity, namely FIRST and VLASS in this study, map and probe the radio emission at deeper depths, allowing us to better distinguish candidate sources of the observed extended emission. We used products from these two surveys to perform multi-wavelength cross matching, with infrared and optical counterparts and produce a spectral index map of the cluster sources at mid-high frequencies. Thus, given the two survey combinations of TGSS with NVSS, and FIRST with VLASS; we can provide spectra analyses of the cluster's diffuse emission and candidate radio sources, respectively, using the computed spectral index maps between 150-1400 MHz and 1400-3000 MHz (see Figure 5.7). We now detail the image packages used to produce these results and the integrated flux values.

The imaging in this Chapter was done using the Common Astronomy Software Application (or CASA in short, Jaeger (2008)) and python, using Astropy (and its affiliated packages; Price-Whelan et al. (2018)). We used PyBDSF (Python Blob Detector and Source Finder, Mohan & Rafferty (2015)) to perform source finding, subtraction (see Chapter 5.4.3.1) and calculate the integrated flux density, as well as the rms values of the radio continuum images used. To match the image area and resolution, in preparation for the computation of the spectral index and error maps, we utilised Miriad (Sault, Teuben & Wright, 1995) for re-gridding and smoothing the previously mentioned survey combinations (where we smoothed the higher angular resolution image, to match the resolution of the lower angular resolution image, see Chapter 5.4.4 for a full description of how the specific surveys were matched). The spectral index maps and their error maps were then computed using Astropy.

5.4 Results

5.4.1 Radio emission and cluster sources

Figure 5.1 and 5.2 show the radio maps of the cluster as observed by VLSSr, TGSS, and NVSS at 74, 150, and 1400 MHz. At these frequencies, the central region of the cluster is dominated by a very bright emission, resulting from multiple radio galaxies (RGs, labelled A; B; BiC and C in Figure 5.3). The FIRST and VLASS radio contours resolve one of these sources, labelled A in Figure 5.3, to be a double jet extended radio galaxy, coincident with an optical and infrared counterpart (see Figure 5.3, and discussion in 5.4.2). Sources B, BiC, and C appear to be point sources in both surveys.

Mid-high frequency observations also confirm the appearance of additional RGs (D and E in Fig. 5.3)) towards east of the cluster as observed by Vacca et al. (2011). These seem to dim with a decrease in observing frequency, with a detection of the northern source in TGSS and no detection for both sources, in the VLSSr image. This is likely to be caused by the sensitivity of the low frequency surveys as well as the low-surface brightness of the sources.

The lower resolution images, in Figure 5.1, reveal the presence of diffuse radio emission that permeates the cluster, surrounding the central RGs, and extends beyond the centre of the cluster.

5.4.2 Multi-ν counterparts

In order to learn more about the bright emission that dominates the cluster's centre, we turned to observations from the FIRST and VLASS surveys at 1400

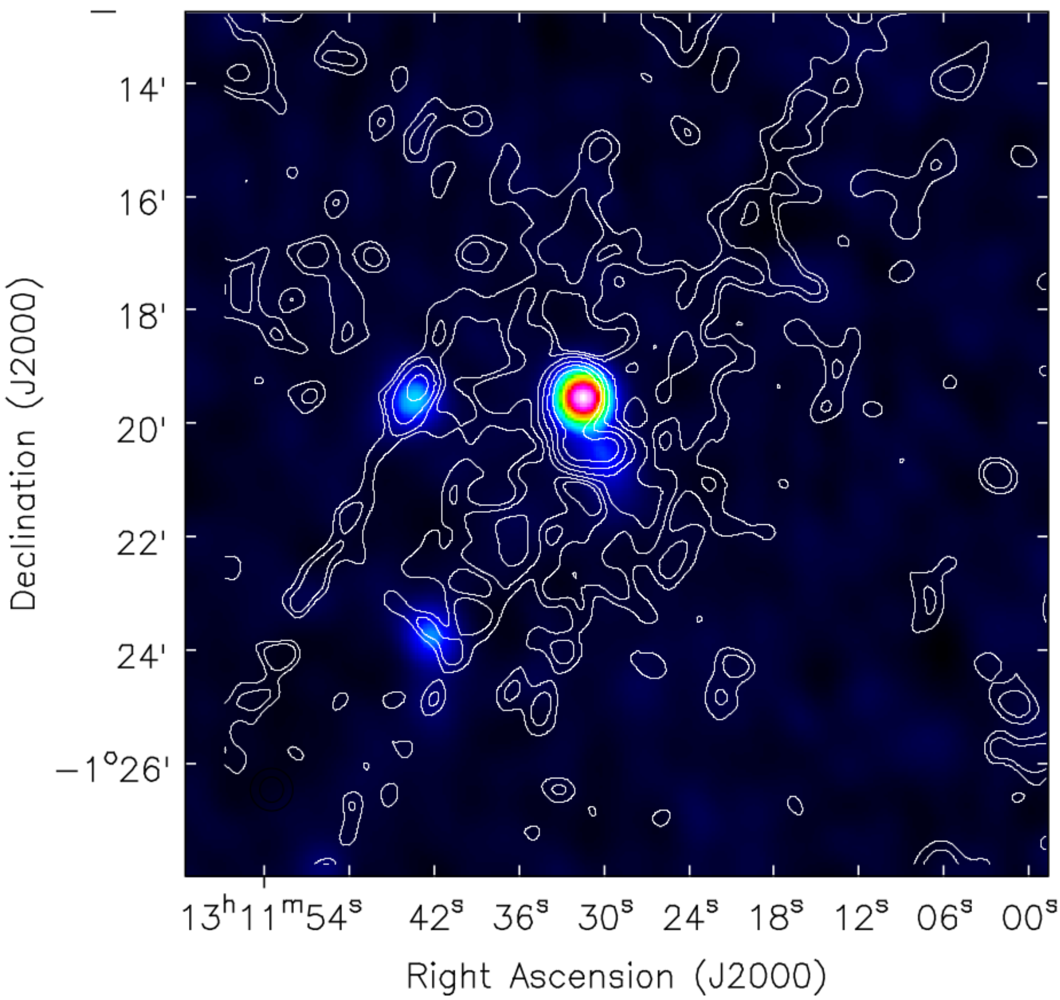


FIGURE 5.1: The NVSS 1400 MHz radio image of the cluster shown in colour with a resolution of $45'' \times 45''$ and rms noise of 0.5 mJy/b, overlaid with the GMRT 147.5 MHz radio contours. The contour lines are drawn at [1,2,4,8,16,...] $\times\ 2.5\sigma_{rms}$, where σ_{rms}=3.22 mJy/b.

MHz and 3000 MHz, with a resolution of $5'' \times 5''$ and $3'' \times 2''$, respectively. We overlapped the FIRST and VLASS observations with infrared (IR) observations from the Wide-field Infrared Survey Explorer (WISE, Wright et al. (2010)) telescope. We do this in order to understand the positions of the RGs in relations to each other and confirm the number of existing point sources in the cluster centre and along the cluster's field-of-view in the radio frequencies. The resulting image is show in Figure 5.3.

The IR-Radio observations reveal that the centre of the cluster has four radio emitting galaxies, three of which were reported in detail by Vacca et al. (2011), these galaxies are labelled as galaxies A, B and C in Figure 5.3. Galaxy A is the cluster's dominant radio-galaxy, located 64.25 arcsec away from the cluster centre and associated with the optical source, TGN259Z069, which

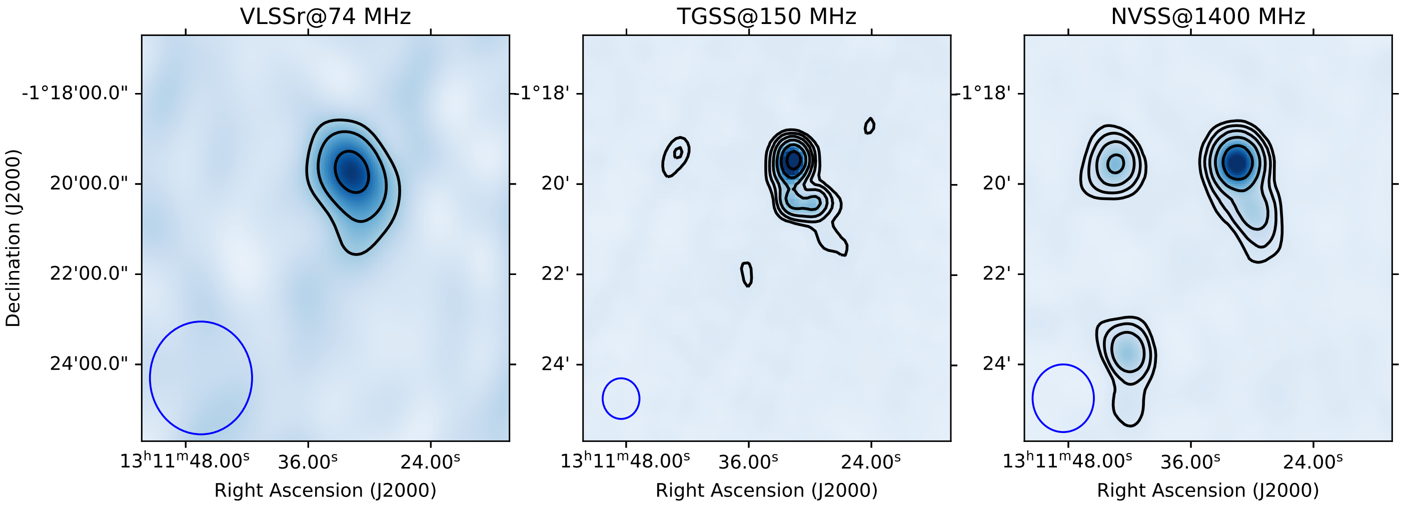


FIGURE 5.2: Left to right: Radio continuum images of each titled survey (and observed frequency) with their respective contours levels drawn at $[1, 2, 4, 8, 16, 32, 64] \times 5\sigma_{\text{rms}}$, where $\sigma_{\text{rms}_{\text{VLLSr}}}, \sigma_{\text{rms}_{\text{TGSS}}}, \sigma_{\text{rms}_{\text{NVSS}}} = [109.75, 3.22, 0.53]$ mJy/beam. The beam sizes are shown on the lower left corner, with a resolution of $75'' \times 75''$, $27'' \times 25''$ and $45'' \times 45''$ for VLSSr, TGSS and NVSS, respectively.

has a spectroscopic redshift of $z = 0.187$. Galaxy B is a radio-galaxy associated with the optical source SDSSJ131130.2-0120, which has a spectroscopic redshift of $z = 0.181$. Vacca et al. (2011) revealed that this point source is located at the cluster's X-ray peak. Galaxy C appears to be the dimmest RG, compared to all the other central RGs, and is associated with an optical source, Slee ACO 1689 12, located at a spectroscopic $z = 0.195$.

The fourth central radio source (associated with the optical source 2MASX J13113265-0119590), which we labelled as BiC in Figure 5.3., is outshined by the dominant radio-galaxy, making it very easy to miss at low frequencies and poor resolutions due to its proximity to the dominant radio-galaxy. This radio-galaxy is located 55.45 arcsec away from the cluster centre and is associated with the brightest galaxy in the cluster (BiC) in the X-ray band, which is located at a spectroscopic redshift of $z = 0.202$. This difference in redshifts amongst the central galaxies can also indicate the possibility of sub-cluster mergers occurring in the cluster.

We have listed the properties of the four central radio sources and the two peripheral sources (D and E) found in the cluster and along its field of view in Table 5.3, where we used Simbad[1] and the Space Science Data Centre (SSDC) Sky Explorer[2] tool to study the redshift, source name and the classification

[1] http://simbad.u-strasbg.fr/simbad/
[2] https://tools.ssdc.asi.it/

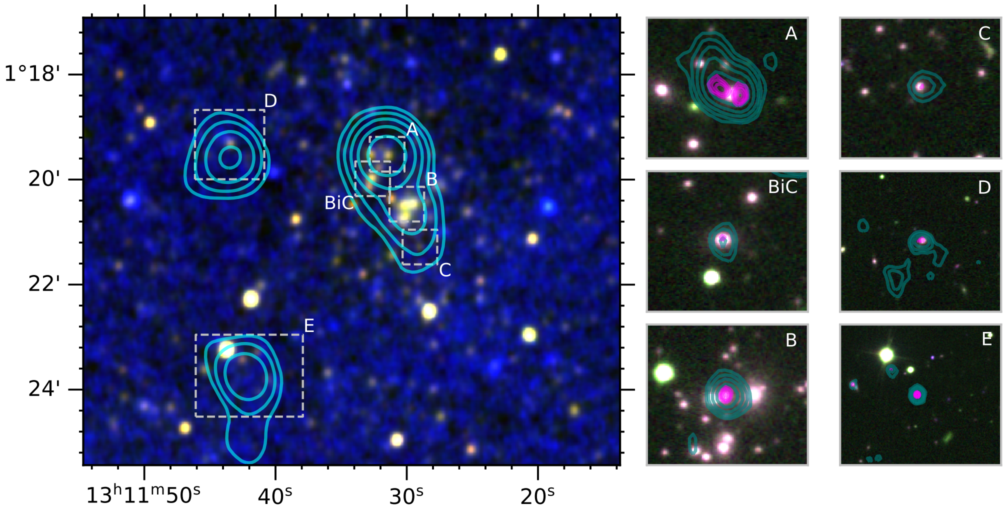


FIGURE 5.3: Left: WISE infrared (W1, W2 and W3 bands) image of the cluster overlaid with the NVSS contour levels at [1,2,4,8,16,32,64]$\times 5\sigma_{\rm rms}$, where $\sigma_{\rm rms} = 0.53$ mJy/beam. Right: SDSS (r, g and i bands) images of the highlighted, silver infrared, regions with FIRST (teal;$\sigma_{\rm rms_{FIRST}} = 0.15$ mJy/beam) and VLASS (magenta;$\sigma_{\rm rms_{VLASS}} = 0.17$ mJy/beam) contours at [1,2,4,8,...]$\times 3\sigma_{\rm rms_{FIRST}}$ and $5\sigma_{\rm rms_{VLASS}} + 3\sigma_{\rm rms_{VLASS}} \times [0, 1, 2, 4, 8, ...]$, respectively.

of the cluster's point sources. We used PyBDSF (Mohan & Rafferty, 2015) to calculate the flux densities of the galaxies at 147.5 MHz with TGSS and 1400 MHz with FIRST. We calculated the flux density values of the point sources, and modelled the point sources as described in Chapter 5.4.3.1. We are unable to obtain an integrated flux density value for the point source associated with the optical source 2MASX J13113265-0119590 at 147.5 MHz as the source cannot be distinctly resolved from Galaxy A by the TGSS.

In Figure 5.1, the diffuse emission permeating the cluster seems to extend beyond the cluster central region at 147.5 MHz, while at 1400 MHz the diffuse emission is almost completely outshined by the the prominent point sources. However, to better understand the extend of the diffuse emission in all the frequencies we have to separate the point sources from the diffuse emission.

5.4.3 Diffuse emission in A1689

5.4.3.1 Flux density estimation

To estimate the flux density of the diffuse radio emission in the cluster, we have to start by subtracting the flux densities of all the compact radio sources (and their extended emission) from the total emission in the cluster. We used

Sources	RA (J2000)	DEC (J2000)	$S_{147.5\,\text{MHz}}$ (mJy)	$S_{1400\,\text{MHz}}$ (mJy)	z	Object type
A^F	13:11:31.5	-01:19:31	443.53 ±29.01	41.95±1.92	0.187	RG
B^G	13:11:30.0	-01:20:28	119.32 ± 12.05	8.44 ±0.73	0.181	RG
C^I	13:11:29.0	-01:21:17	111.5 ± 10.56	2.14 ±0.22	0.195	RG
BiC^H	13:11:32.6	-01:19:59	-	2.04 ±0.19	0.202	BiC
D^J	13:11:43.5	-01:19:20	51.29 ± 4.81	6.22±0.61	0.183	RG
E^K	13:11:42.0	-01:23:44	-	10.09 ±0.92	-	RG

TABLE 5.3: Position, radio and general properties of the radio sources found in A1689 and along its field-of-view. Col. 1 shows the label of the cluster sources, Cols. 2 and 3 are the coordinates of the sources, Cols. 4 and 5 are their integrated flux densities at 147 MHz and 1400 MHz, Col. 6 is the redshift of the source and Col. 7 is the classification of the point source. (RG- radio-galaxy, BiC - Brightest galaxy in the cluster in Simbad). The associated source names (or identifier in Simbad) of each labelled sources are: *F* - TGN259Z069, *G* - SDSSJ131130.2-0120, *H* - 2MASX J13113265-0119590, *I* - A1308-0150D, *J*- TGN259Z044 , *K*- Slee ACO 1689 18 .

PyBDSF to model the point sources, with a Gaussian fitting function and estimated their corresponding integrated flux densities (such as the ones presented in Table 5.3) at the different observed frequencies. The FIRST survey was used to resolve positions of potential point sources within the studied fov and 13 point sources were identified.

To get the integrated flux density of the diffuse radio emission, the total flux density contributions of the embedded point sources (and their extended emission) at that particular frequency are modelled and then subtracted from the total emission in the observed radio image in PyBDSF . The resulting images then produces a low-surface brightness residual map that contain emission which could not be modelled as point sources by the source finder. We then use this emission to calculate the diffuse emission contained within the fov of the continuum radio images.

The errors in the flux density measurements of the diffuse emission are estimated by taking into account the errors that are caused by the flux densities used to calibrate the radio images and the errors resulting from the noise of the image. These two errors are then added together for the flux density error estimation, using the formula:

$$\Delta S_\nu = \sqrt{(\sigma_{\rm cal} S_\nu)^2 + \left(\sigma_{\rm rms}\sqrt{N_{\rm beam}}\right)^2}, \tag{5.1}$$

where $\sigma_{\rm cal}$ is the uncertainty in the calibration, $\sigma_{\rm rms}$ is the noise of the radio image as obtained from from the FITS files and $N_{\rm beam}$ is the number of beams

ν (MHz)	73.5	147.5	200	1200	1400
S_ν (mJy)	497.3 ± 61.3	202.9 ± 19.8	153.3 ± 16.8	11.7 ± 3.4	8.7± 0.86

TABLE 5.4: Flux densities of the diffuse emission found in the galaxy cluster at different radio frequencies, including the 1200 MHz flux density value from Vacca et al. (2011). The values at 147.5 MHz and 200 MHz are taken as upper limits, due to the poor resolution of the VLSSr and GLEAM observations.

present in the extended source. We assumed $\sigma_{\rm cal}$ to be 10 % for the TGSS, NVSS and GLEAM images (Intema et al. (2017); Condon et al. (1998); Hurley-Walker et al. (2017)) and assumed it to be 15% for the VLSSr images (Lane et al., 2014).

Table 5.4 shows the flux densities of the diffuse emission at the different radio frequencies, after the subtraction of all the point sources. We found the flux density of the diffuse emission at 147.5 MHz to be $S_{147.5\,\rm MHz}$ = 202.9±19.8 mJy and to be $S_{1400\,\rm MHz}$ = 8.7±0.86 mJy at 1400 MHz, for the NVSS survey. We have also included the value from Vacca et al. (2011) at 1200 MHz on our table. The VLSSr and GLEAM images had the highest noise levels due to their surveys' low resolutions and poor sensitivities, making them less ideal for detecting low-surface brightness features such as radio halos. Thus, their resulting integrated flux densities of the diffuse emission at 73.5 MHz and 200 MHz are taken only as upper limits. The flux density at 1400 MHz is comparable to the the value estimated by Vacca et al. (2011) at 1400 MHz, which was 9.9 mJy. We didn't find any diffuse emission in the VLASS and FIRST radio images above 3σ.

We use the flux density values stated in Table 5.4. to plot a radio spectrum for the diffuse emission in the cluster. The radio spectrum is shown in Figure 5.4 and it includes the flux density value of the Radio halo at 1200 MHz as calculated by Vacca et al. (2011). The flux density values at 73.5 MHz and 200 MHz are also expressed as upper limits in the radio spectrum. Using the radio spectrum plot, we see that the diffuse emission in the cluster follows a power law relation with a steepness of $\alpha_{\rm diffuse}$= -1.34±0.14 using a linear fit. We estimate this value as the spectral index of the diffuse emission. This spectral index value and the flux density value $S_{1400\,\rm MHz}$ = 8.7 ±0.86 corresponds to a radio power $P_{1400\,\rm MHz}$= 9.81 ± 0.77 × 10^{23} WHz^{-1} at 1400 MHz.

5.4.3.2 Morphology of diffuse emission

Figure 5.5 shows the diffuse emission map in the cluster at 147.5 MHz after the subtraction of the point sources as described in the previous subsection. At 147.5 MHz the diffuse emission extends beyond the central regions of the cluster, this is in agreement with what has been observed at 1200 MHz by Vacca et al. (2011), where the radio halo was mainly concentrated at the

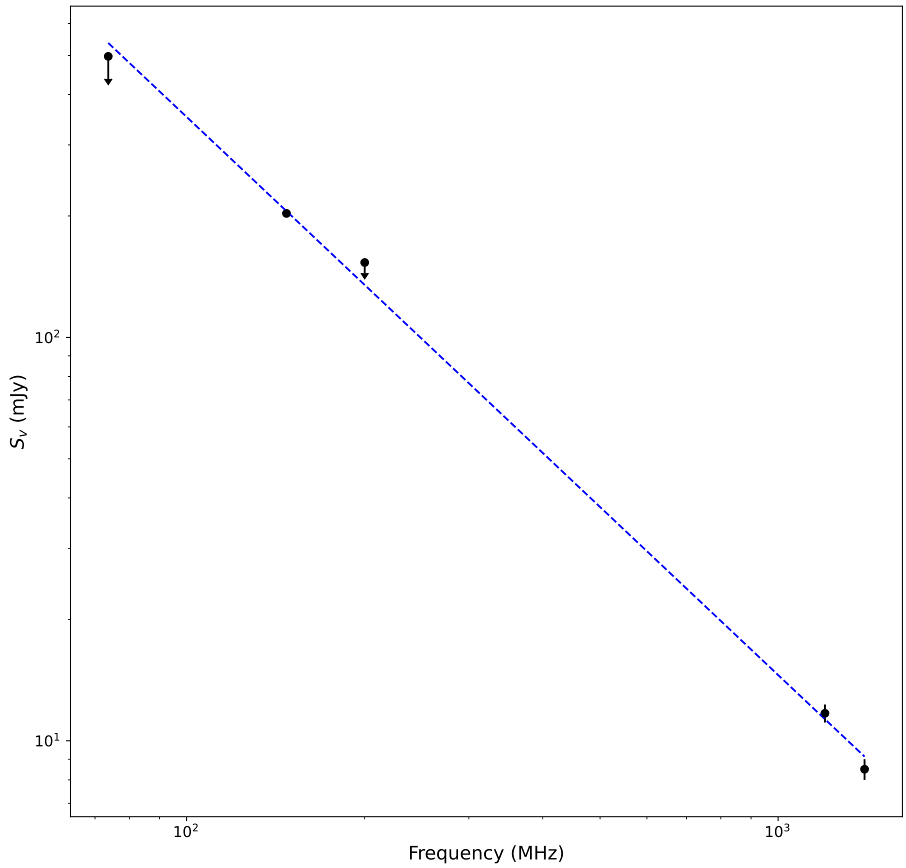


FIGURE 5.4: The spectrum of the diffuse emission found the A1689 cluster at frequencies ranging from 73.5 MHz to 1400 MHz with the errors obtained through equation 1, we have also included the flux density at 1200 MHz as calculated by Vacca et al. (2011). The blue line is the linear fit to the data.

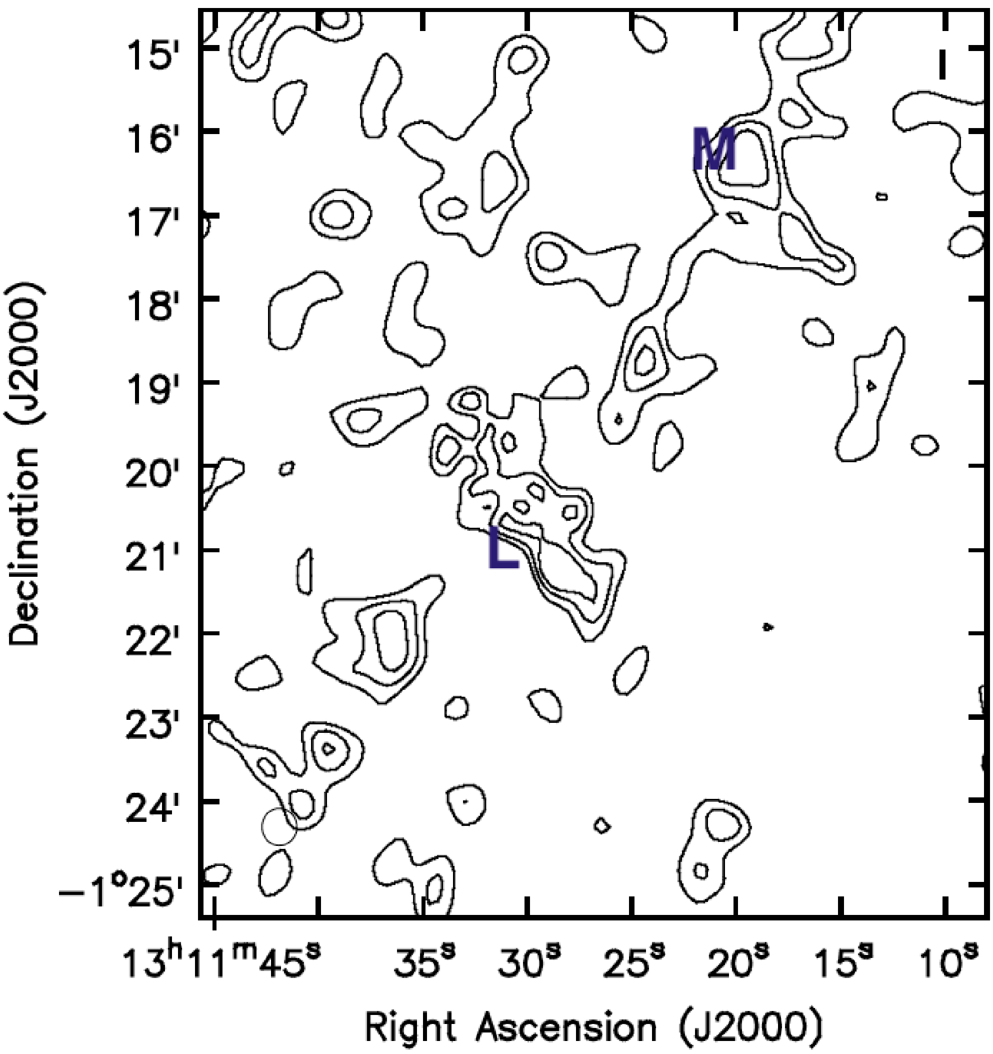


FIGURE 5.5: The 147.5 MHz radio contours of the diffuse emission found in the cluster after the subtraction of point sources using PyBDSF. The resolution of the image is $\sim 27'' \times 25''$. The contour levels are [1,2,4,8,16,...]$\times 3\sigma_{rms}$ where σ_{rms}= 3.22 mJy/b.

cluster centre. The diffuse radio map at 147.5 MHz reveals that the cluster hosts numerous diffuse radio filaments and peaks, with the major peak being found in the cluster centre. The major peak, labelled L, displays an irregular morphology and has sharp edges and occupies the cluster centre. this is the same structure detected at 1200 MHz by Vacca et al. (2011). We found that the longest linear size (LLS) of the radio halo peak located at the centre, peak A, to be LLS ~720 kpc at 147.5 MHz. This value is not too far from the one reported in Vacca et al. (2011), where they found LLS=730 kpc for the radio halo. However, we expected this value to be greater than 730 kpc at 147.5 MHz as radio halos tend to appear bigger in size at lower frequencies, due to lower frequencies being optimal for the detection of low surface brightness radio emission. We may be obtaining a lower LLS value due to the sensitivity limitations and calibration & cleaning techniques of publicly available TGSS data. The short LLS may also hint at the diffuse emission being in a transitional phase from being a mini-halo to a radio halo, with the gas sloshing of the central galaxies and sub-cluster mergers as possible sources of the turbulence producing diffuse radio emission. The central diffuse emission is expected to be incident with the X-ray peak of the cluster as shown in Vacca et al. (2011). In figure 5.6 we compare the LLS of the cluster's radio halo to other well known radio halos at different cluster redshift. From figure 5.6 we observe that the LLS, for the radio halo in A1689, obtained from this study is small to medium sized when compared to radio halos of the same redshift range.

We also observe the presence of another irregular diffuse structure in the north-western region of the cluster, labelled M, this is the second largest diffuse structure after the main one found in the cluster. There is a possibility that the irregular diffuse structure is a radio relic, however, this possibility has to be investigated independently with low frequency radio observations paired with polarization and X-ray studies. The rest of the radio filaments, spread unevenly throughout the cluster, although most of them are found in the northern to south eastern regions of the cluster. Most of the filaments also form part of the diffuse radio emission as they mostly lack optical counterparts. The diffuse filaments can also result from possible smaller scale sub-cluster merger events, which have not been able to destroy the central cool-core of the cluster.

The main diffuse structure is still quite small in size even at 147.5 MHz compared to most radio halos, raising the possibility of a mini-halo instead, although this will be hard to argue at this sensitivity.

5.4.4 Spectral index imaging

Spectral index images can be used to constrain the energy spectrum of relativistic electrons and for studies of magnetic field distribution in clusters

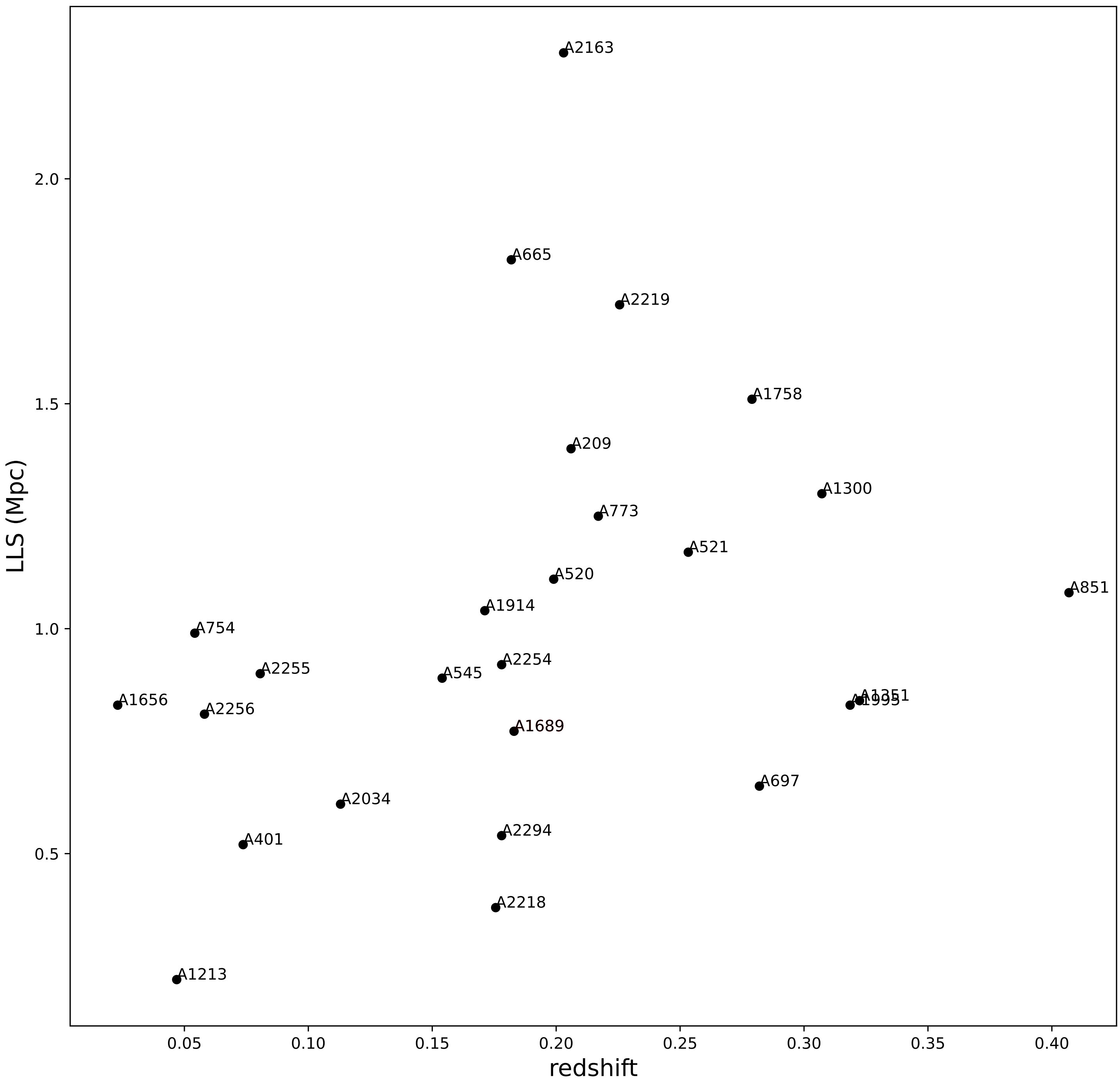


FIGURE 5.6: The largest linear sizes of some well known radio halos compared to the redshift compared to the LLS for A1689 (in red) found in this study. Data points are taken from Feretti et al. (2012)

hosting diffuse radio emission. Thus, providing useful information to understand the formation and evolution of cluster merger processes.

In order to further examine the cluster dynamics of A1689 at radio frequencies, we computed spectral index distribution maps of the cluster using a combination of images from TGSS & NVSS, at low-mid frequencies and images from FIRST & VLASS, at mid-high frequencies, with overlapping coverages.

For the preparation step, the TGSS radio continuum image was re-gridded to match the size of the NVSS image. We then smoothed the TGSS image (using the SMOOTH task in MIRIAD) in order to match it to the resolution of the NVSS image(i.e., $45'' \times 45''$) with a 2-D elliptical Gaussian function. The same process was followed to match the high resolution VLASS image with the FIRST image. The spectral index image was created using ASTROPY and its affiliated packages, where regions with surface brightness values that are less than desired noise level of each image were blanked out. The formula used for the spectral index map was of the form: $S_\nu \propto \nu^\alpha$, where S_ν is the surface brightness at a frequency ν and α is the spectral index value. The uncertainty in the spectral index value was calculated using the formula:

$$\Delta\alpha = \frac{1}{\ln\left(\frac{\nu_1}{\nu_2}\right)} \sqrt{\left(\frac{\sigma_{1,\mathrm{rms}}}{I_1}\right)^2 + \left(\frac{\sigma_{2,\mathrm{rms}}}{I_2}\right)^2} \tag{5.2}$$

where $\sigma_{1,2,\mathrm{rms}}$ are the rms values of the TGSS & NVSS and FIRST & VLASS images. $I_{1,2}$ are the surface brightness values of the respective frequencies at each pixel and $\nu_1 > \nu_2$.

Figure 5.7 shows the resulting spectral index and error map between 147.5 MHz & 1400 MHz, alongside the spectral index images (computed between 1400 & 3000 MHz) of the previously highlighted regions in Figure 5.3. From the computed spectral index maps, we notice that the central region of the cluster does not show enhanced spectral index features, with the spectral index of the extended radio source varying from $\alpha^{1400\mathrm{MHz}}_{147.5\mathrm{MHz}} = -1.8$ to $\alpha^{1400\mathrm{MHz}}_{147.5\mathrm{MHz}} = -0.8$. The central flatter regions have an average value of $\alpha^{1400\mathrm{MHz}}_{147.5\mathrm{MHz}} = -1.0$, surrounded by regions which steepen towards the edges of the source.

The two peripheral regions have flatter spectral index values, with region E exhibiting average inverted spectral index values closer to $\alpha^{1400\mathrm{MHz}}_{147.5\mathrm{MHz}} = 0.2$, and average flatter $\alpha^{1400\mathrm{MHz}}_{147.5\mathrm{MHz}} = -0.35$ for region D. We note that this is likely to be caused by the sensitivity of the TGSS survey, since both region D and E are sites of low-surface brightness emission that might not be properly picked up by the legacy survey. If these features are not from systematics, a possible interpretation of the physical phenomenon could be that; since A1689 is

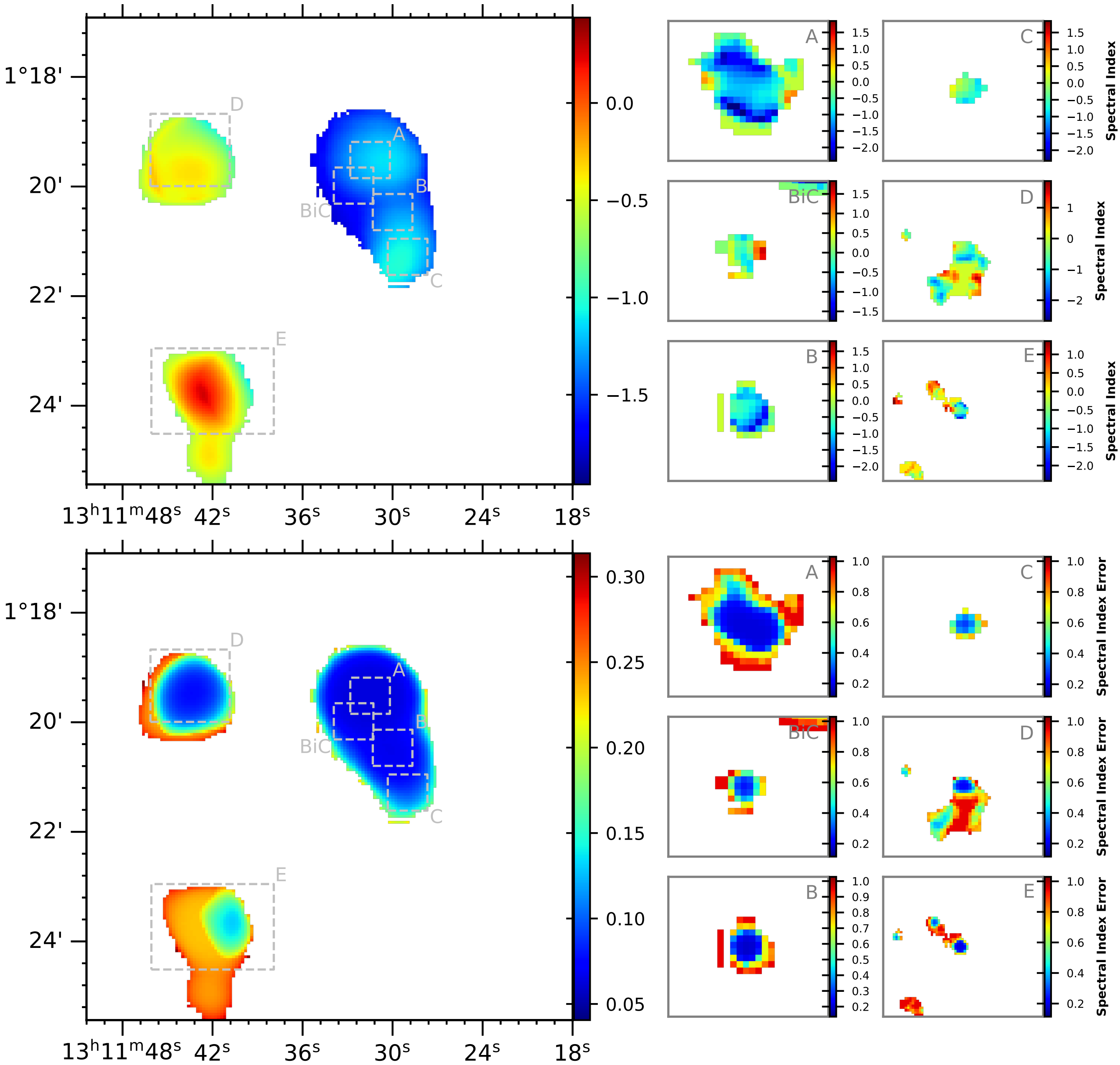


FIGURE 5.7: **Top left panel:** Spectral index map between 147.5 and 1400 MHz at a resolution of $45'' \times 45''$, with a mask that covers the same region as the NVSS continuum image in Fig 5.2. **Top right panel:** Spectral index maps between 1400 and 3000 MHz at a resolution of $5'' \times 5''$, with a mask that covers the same region as the FIRST continuum image for values higher than $5\sigma_{\rm rms}$, highlighting identical regions found in Fig 5.3. **Bottom left and right panels:** The corresponding spectral index error maps of the detailed top panels.

a cool-core cluster, most of the energy radiated by the source is radiated at higher frequencies. This would potentially explain why the extension of the source is decreasing with frequency and why we are not observing a consistent power-law relation that is expected with synchrotron emission.

We also notice that the centre of the cluster is the steepest region, with $\alpha <$-2. This steepness is due to the combination of all the unresolved point sources at the central regions of the cluster and along its field of view. This poor resolution of unrelated sources creates an unevenly steep central region compared to other regions with resolved radio maps in the cluster.

We currently refrain from using these spectral index images to estimate the spectral index of the cluster's diffuse radio emission, since the diffuse emission in the NVSS, 1400 MHz image does not extend as much as the diffuse emission on the 147.5 MHz image. Also the mask used to blank out values in the spectral index image was derived from the FIRST radio map and so it leaves out most of the diffuse radio emission observed at 147.5 MHz .

5.5 Summary and conclusion

In this Chapter we studied the radio properties of A1689 using archival radio observations from numerous radio telescopes. The study confirmed the presence of an irregular shaped radio halo in the central regions of the cluster, as previously reported by Vacca et al. (2011) at 1200 MHz. We found that the diffuse emission in the cluster is outshined by four radio-galaxies found in the cluster's centre, one of which is the central radio-galaxy, located at redshift z = 0.187 and with flux density of $S_{147.5\,\mathrm{MHz}}$ = 443.53 mJy.

At 147.5 MHz we also found the existence of more diffuse radio structures and filaments which spread over the cluster. These diffuse filaments could result from the re-acceleration of cosmic electrons by small scale sub-cluster mergers. We also found the LLS of the main radio halo at 147.5 MHz to be LLS$\sim$ 720 kpc which is within range to the value of LLS = 730 kpc reported by Vacca et al. (2011) at 1200 MHz. We expected the LLS value to be more than 730 kpc at 147.5 MHz, however due to the spatial resolution, calibration and imaging techniques of the TGSS, we obtained a smaller value for the LLS.

Due to the NVSS's short baseline, it is insensitive to extended, low surface brightness emissions larger than 15$'$ and so using PyBDSF we tried our best to estimate the integrated flux density of the diffuse emission at 1400 MHz. We found the flux density of the diffuse emission in the cluster to be $S_{1400\,\mathrm{MHz}}$= 8.7 $\pm$ 0.86 mJy which is not that far from the value of $S_{1400\mathrm{MHz}}$ =9.9 mJy, estimated by Vacca et al. (2011). A spectral analysis of the cluster from the integrated flux densities of the diffuse emission at 5 frequencies shows that the diffuse emission in the cluster follows a power law relation. We estimated the spectral index value of the diffuse emission to be α_{diffuse} = -1.34. $\pm$ 0.14. This spectral index value and the flux density value at 1400 MHz corresponds to the radio power of the diffuse emission to be $P_{1400\,\text{MHz}}$= 9.81 $\pm$ 0.77 $\times$ 10^{23} W Hz^{-1}. In this study we have also presented the first spectral index image of the cluster, computed between 147.5 MHz & 1400 MHz and 1400 MHz & 3000 MHz. The spectral index image reveals that the central region of the cluster is the steepest, most likely due to the combination of different radio-galaxies.

The revelation of the radio halo and more diffuse radio filaments at 147.5 MHz compared to higher observational frequencies shows the importance of combining low radio frequencies and sharp sensitivity in telescopes for the optimal detection of diffuse radio emission. An improvement in the sensitivity and spatial resolution of the observations at lower frequencies would allow us to study the relationship between the diffuse emission and the central radio-galaxies in the cluster better. This study can be possible for the cluster with the launch of the SKA, which will assist in the study of the cluster dynamics which lead to the occurrence of diffuse emission.

Data availability

All the data used in this work is available on the archives of the NVSS[3], TGSS[4], VLSS[5], MWA[6], VLASS[7] and FIRST [8] websites.

[3]https://www.cv.nrao.edu/nvss/
[4]http://tgssadr.strw.leidenuniv.nl/doku.php
[5]https://heasarc.gsfc.nasa.gov/W3Browse/all/vlss.html
[6]https://www.mwatelescope.org/gleam
[7]https://science.nrao.edu/science/surveys/vlass
[8]https://science.nrao.edu/science/surveys/first

Chapter 6

Unsupervised machine learning for the characterisation of radio sources

This chapter aims to connect the astrophysical insights from earlier chapters with the essential machine learning ideas that the authors believe are effective solutions for scaling up prior searches. It focuses on exploring the possibility of using dimensional reduction techniques (pre-trained networks in this instance) as feature extractors. The feature of interest is diffuse radio emission from extended radio sources. The novelty is in considering a multi-frequency and multi-field-of-view treatment of radio sources and curating these sources using an unsupervised prescription. Thus, the only human invention is in the selection of the datasets, and beyond that, the outputs are machine-curated representations of the input data.

Note: This chapter is adapted from the ongoing work by: Thwala S. A., Sefala R., and Beck G., titled "*Unsupervised machine learning for the characterisation of radio sources*". The work is currently being adapted for submission to the Monthly Notices of Royal Astronomy journal.

For this work, Thwala S. A., carried out all of the experiments and reporting. Sefala R., suggested the use of dimensional reduction techniques as an avenue for making sense of high dimensional data and was consulted extensively about the techniques, possible avenues of understanding, and edited the ICLR2020 extended abstract. Beck G., supervised the project and suggested revisions to the drafts.

Abstract

Fundamental science corners its principles in the idea that nature can be modelled, using a set of minimal variables, to trace complex observable behaviour. This presents a challenging obstacle in curating datasets for fundamental science objectives, like searching for diffuse emission in large datasets

from modern radio surveys. The high dimensionality of input data, due to probes having high resolution and sensitivity at multiple scales, coupled with varying physical phenomenon and immense data volumes, requires thorough domain expertise to set up fundamental prescriptions that are explainable and computationally retractable. In this Chapter, we explore the use of convolutional neural networks (CNNs) to reduce the dimensionality of radio astronomical images, and self-organising maps (SOMs) to cluster the reduced radio continuum images into different morphological classes. We discuss how the reduced dimensional space can be thought of as a physical abstracted space. The study uses data from the VLSSr, TGSS, GLEAM, WENSS, NVSS, and FIRST surveys at their respective central frequencies of 74, 150, 200, 325, 1400, and 1400 MHz. For the dimensionality reduction step, we use a combination of four pre-trained CNN architectures, namely: Inception, Inception-Resnet, PNASNET, and Resnet. We then use the results of these architectures to train a SOM, which clusters these reduced features into different morphological classes. This method provides a systematic unsupervised path for curating large radio continuum data cubes (produced by large radio surveys such as the MeeKAT and SKA in the future) and can support radio astronomers with source characterisation (and ultimately, source extraction).

6.1 Introduction

Scientists working with high-dimensional data, such as images captured by astronomical surveys, often face challenges of computational efficiency for data analyses. Coupled with most improvements in sensor or probe technologies is an increase in collected data size and observed features, which introduce a challenge in extracting knowledge from datasets. This challenge is not easily addressed by high-performance computing (HPC), though HPC is a solution for addressing the high-throughput of data being captured by the sensors efficiently. The subtle challenge introduced here is how to efficiently extract novel insights from datasets with non-trivial interrelations.

Data-intensive computing, a class of parallel computing applications that use a data-parallel approach to processing large volumes of data (Jayalakshmi, Srinivasan & Srinivasa, 2016), is designed with the described challenge in mind. Looking at astronomical datasets, we seek to mine high-dimensional data to extract the least number of variables that best describe the physical structure(s) contained within the non-trivially interrelated datasets. Chapter 3.1 presents an in-depth discussion of the complex interrelationship between astronomical datasets and the phenomena they contain.

Machine learning (ML), the study of algorithms and statistical models that extract patterns and learn representations from raw data, offers an avenue for insight extraction from these non-trivially interrelated datasets. An approach of interest would be to assume we are not smart enough to detail prescriptions to computers and let machines learn on their own, some suitable relation(s) of the interplay[1].

Baron (2019) provides a detailed and practical summary of ML in astronomy, including its history, the difference between supervised and unsupervised learning, neural networks, and applications of ML in radio astronomy. Works by Aniyan & Thorat (2017) and Tuccillo et al. (2018) demonstrate the significant impact of using deep learning techniques, specifically convolutional neural networks (CNNs), to address the task of galaxy morphological classification for radio and near-infrared images, respectively. Self-organising maps (SOMs; Kohonen, 1997, 2013) have been a useful, unsupervised learning technique for the morphological analysis of large image datasets in radio astronomy (Polsterer, Gieseke & Igel, 2015; Galvin et al., 2020; Mostert et al., 2021). A common limitation in the aforementioned studies is that the training of inputs only focuses on one field-of-view or scale at a time. This can be addressed by considering approaches that take in n-th layer inputs to generate flattened representations or channels (such as Jin, Dundar & Culurciello, 2015), which can be interfaced with CNNs. We outline the differences between this study and the above studies in the final paragraph of this section.

Although the adoption and implementation of ML in astronomy has been rapid and widespread (see Baron (2019) and references therein), the challenge of making sense of the data without human intervention is still a big hurdle. ML techniques tend to generalise, from input data, further into parameter spaces that are unfeasible or seem unfeasible when applying human reasoning. This has led to semi-supervision being the most viable solution, where the machine models learn what the underlying statistics are, and then a human has to verify if the outputs are feasible. Though we agree with this approach, the claims that tend to follow are not aligned with the sentiment and end up leaning towards the view of machines being able to work without human intervention. In this work, we explore if using pre-trained networks can achieve a machine representation that converges with human representations. We have a particular interest in finding diffuse emission features in radio sources.

[1]We then assume we are smart enough to understand the machine learned representations (through visualisations and continuum radio source characterisation), evaluating if the learned representation are physical.

Works by Pan, Kwok & Yang, 2008 have illustrated the benefit of using transfer learning as a dimensionality reduction technique, to produce latent features. Focusing on the latent feature space and considering inputs from fundamental fields, we can treat the latent feature space as an abstracted, physical space. The intuitive justification is that since the inputs follow physical laws (as extensively discussed in Chapter 2), the dimensionally reduced space should contain structure(s) that resemble physically abstracted representations. We also have that by design, different pre-trained architectures capture and treat data differently. These architectural variations can be used to probe the abstraction of each reduction network and group networks which produce similar abstractions. The grouped abstractions can then be compared with known fundamental principles and, if consensus is reached can, be viewed as physical abstractors.

The physical abstractors approach offers us the ability to do consistency checks with unsupervised models. By fixing the input data and the unsupervised model, we gain control over analysing and understanding the features the abstractors are abstracting. Through the variation of the pre-trained models used and detailed post-processing analysis, insightful groupings can then be formed. Making it possible to systematically create a set of useful measures and tools for identifying underlying structures in fundamental datasets from abstractors. These measures provide complementary tools of analysis where abstractors capture the perceived reality contained within the image and domain experts capture the consensus of the accepted reality contained in the literature. We now sum up all of the above information, restate why this study is of interest, and conclude this section with the practical implications of the study.

To test if we can use the latent feature space from dimensionality reduction techniques as a physically abstracted space, pre-trained CNNs were used as abstractors of radio continuum survey images. These abstractions are useful in addressing the task of source extraction, one of the open problems in radio astronomy. It has been a challenging problem to address since the evolution of radio galaxies is largely shaped by the environment it coevolves in. This coevolution gives clues to the dynamics at play within the source and its interactions with the surrounding environment. The diversity in the observed features makes it challenging to draw up simple rules for source extraction; although these may be easy for domain experts to pick up, the problem has not been solved at scale. Abstractors[2] provide a unique opportunity to find statistical tracers for these environments, and aid astronomers with source extraction.

In this work, we explore using transfer learning ML techniques to assist us

[2]We note that the terms pre-trained networks and abstractors are used interchangeable, in this Chapter.

in reducing the dimensionality of radio continuum images in our dataset, and for abstracting features contained within radio continuum data cubes. We compare four different pre-trained CNN architectures namely: Inception (Szegedy et al., 2016), Inception-Resnet (Szegedy et al., 2017), PNASNET (He et al., 2016), and Resnet (Liu et al., 2018) for the dimensionality reduction step. A SOM was then used to cluster the radio continuum images into different morphological classes. Our study differs from prior works in that it uses as input a multi-dimensional data cube (composed of multiple frequencies and field-of-views) with abstractors to pre-process the input data, producing a reduced abstract representation of the multi-dimensional data cube that is then passed to the SOM for clustering. We then look at the consistency of these classes, as perceived by the different abstractors.

6.2 Treatment of Data

We investigated the feasibility of using pre-trained networks as abstractors of physical features contained within radio continuum images by analysing data from six publicly available deep legacy continuum radio surveys, as detailed in Chapter 3.2. These were VLSSr (Lane et al., 2014), TGSS (Intema et al., 2017), GLEAM (Hurley-Walker et al., 2017), WENSS (Rengelink et al., 1997), NVSS (Condon et al., 1998), and FIRST (Becker et al., 1995), all of which had central frequencies of 74, 150, 200, (88, 118, 155, and 200), 1400, and 1400 MHz, respectively. No calibration step was necessary for this study because the relevant survey teams had already cleaned and calibrated the cutout images before making them available to the public.

To ensure that we trained the models on diverse representations of extended radio sources, we used coordinates from a number of published studies that focused on:

- extended radio sources and galaxy clusters (Schoenmakers et al., 2000; De Breuck et al., 2006; Cavagnolo et al., 2009; Gendre, Best, & Wall, 2010; Nandi & Saikia, 2012; Mann & Ebeling, 2012; Martino et al., 2014; Giacintucci et al., 2014; Proctor, 2016; Capetti, Massaro, & Baldi, 2017a,b; Aniyan & Thorat, 2017; Kuźmicz et al., 2018; Mahatma et al., 2019; van Weeren et al., 2019; Dabhade et al., 2020; Duchesne et al., 2021),
- a combination of both extended and compact radio sources (Vardoulaki et al., 2019), and
- compact radio sources (Baldi, Capetti, & Massaro, 2018).

We further augmented the source list by randomly sampling combinations of the RA and DEC from the known list. This was done to include off centred sources, and noise dominated images into the sample size. The caveats with the pursued bootstrapping approach are, the computational costs it takes to

TABLE 6.1: Retrieved image sample sizes, from deep legacy continuum radio surveys, considered in this study.

Survey	Frequency	Synthesized Beam Size	Sample Size
VLSSr	74 MHz	75″	≈10300
GLEAM-01	88 MHz	≈120″	≈8000
GLEAM-02	118 MHz	≈120″	≈8000
GLEAM-03	155 MHz	≈120″	≈8000
GLEAM-04	200 MHz	≈120″	≈8000
TGSS ADR1	150 MHz	25″	≈10300
WENSS	325 MHz	54″×54″ cosec δ	≈2100
NVSS	1400 MHz	45″	≈10500
FIRST	1400 MHz	5″	≈8500

continuously retrieve the randomly sampled images, and some sources only appear in some surveys[3]. Due to these two issues, we then end up with varying sample sizes of radio continuum images considered (for different analyses) in this study, as presented in Table 6.1. These issues are a factor in the analyses of radio sources abstracted in the multi-frequency component of the study, where we abstracted sources observed across multiple surveys. We address these by only considering sources which are observed by a specified combination of considered surveys.

As a consequence of choosing diverse representations, the wide range and subtleties in differentiating features of considered extended sources allow us to test the abstraction and binning capabilities of the pre-trained networks and SOMs, respectively. The architecture diagram of the developed approach is presented in Fig. 6.1.

The developed framework takes in astronomical coordinates as inputs to SkyView[4], querying and retrieving observed radio continuum images from the image cutout services of the previously mentioned publicly available legacy surveys. Each SkyView query searches for a specified field-of-view (FOV; region of interest), and survey(s).

For each source coordinate, 6 images were downloaded at different scales with a FOV of 10′, 20′, 30′, 45′, 84′, and 120′ corresponding to 0.167°, 0.333°, 0.5°, 0.75°, 1.4°, and 2.0°, respectively. These allowed us to carry out analyses from individual radio galaxy scales to galaxy clusters and cosmological scales. We also specify the mentioned surveys as the data sources to

[3]This is generally true for all the surveys (since radio surveys are geolocated at specific sites) and is not unique to the described bootstrapping approach.

[4]https://skyview.gsfc.nasa.gov/

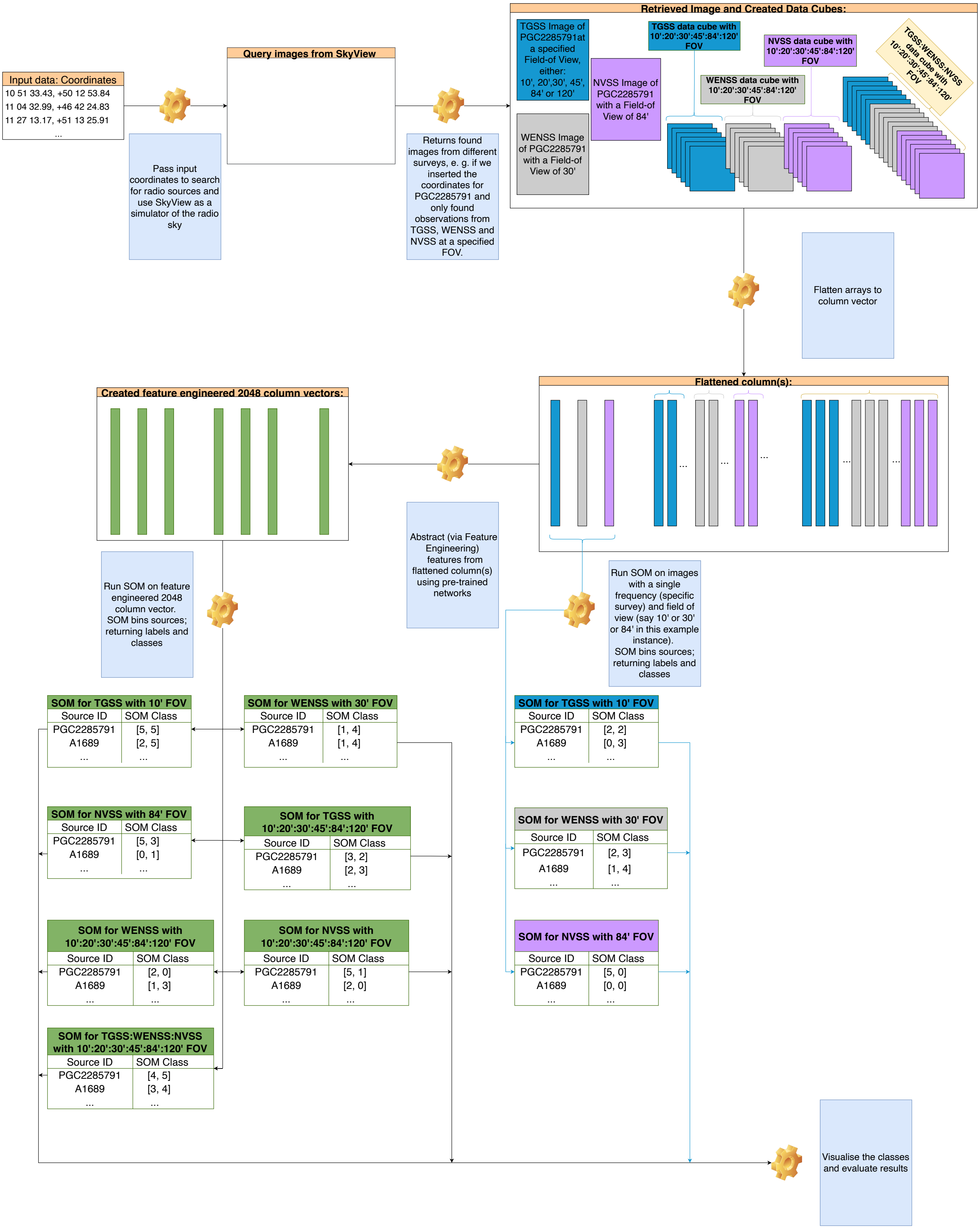


FIGURE 6.1: An architecture diagram of the different concepts used to go from an input of coordinates to a set of visualised SOM classes.

query at different frequencies, with the surveys spanning from low to mid-frequencies, at a coverage of 74 MHz, 88 MHz, 118 MHz, 150 MHz, 155 MHz, 200 MHz, 325 MHz, 1400 MHz and 1400 MHz for the VLSSr, GLEAM-01, GLEAM-02, TGSS ADR1, GLEAM-03, GLEAM-04, WENSS, NVSS and FIRST surveys, respectively.

The returned images are then given a unique source name (based on its coordinates, e.g. "source 09h59m33.288s +02d08m32.7264s") to function as a global identifier, and a derived source key (e.g. "source_149.88869999999997_2.142424_gleam_02_0.75_118") to function as an image specific identifier. These two identifiers allow us to link any downloaded images with the same coordinates (via the unique source name) whilst allowing us to uniquely query the linked images for a specific survey image at a desired FOV. Using this logic, we separated the retrieved images into 3 data cube categories:

- Original image data cubes for a given FOV and survey: where each downloaded image is grouped into a data cube of continuum radio sources at a certain FOV for a specified survey frequency
- Multi-scale data cubes for a given survey: where a collection of downloaded images at multiple FOVs (10′, 20′, 30′, 45′, 84′, and 120′) are grouped into a data cube of continuum radio sources at multiple FOVs and a specified survey frequency
- Multi-scale and multi-frequency data cubes: where a collection of multi-scale data cubes from a combination of specified surveys (TGSS, WENSS, and TGSS, using the example surveys depicted in Fig. 6.1) are grouped into a data cube of continuum radio sources at multiple FOVs and survey frequencies

The motivation behind creating the multi-scale, and multi-scale and multi-frequency[5] data cubes is that different features appear and are dominant at different scales and frequencies. Thus, a functional framework has to take that into account and abstract features across multiple scales and frequencies. Domain experts also follow a similar prescription, in that preliminary, they will look at a given source for a particular survey and field of view. Based on the source's morphological extent and interaction with the surrounding environment, they then have to choose the right combination of surveys and FOV to characterise the features observed at different frequencies and scales. We exploit the fact that computers perceive images as arrays, granting them the ability to simultaneously "comprehend" the different frequencies and scales.

To an extent, we are investigating which combination of abstracted perspectives (surveys) and regions of interest (FOVs) allow us to group extended

[5]Shortened to multi-scale-frequency from hereon.

radio sources in a manner that is contextually correct and intuitive to radio astronomers. We, therefore, can intuitively think of the multi-scale data cubes as giving degrees of freedom to the pre-trained networks to abstract zoomed in and out features of radio sources. Extending this idea to the multi-scale-frequency data cubes, further adds the ability to switch between the different features (or perspectives) observed by surveys at different frequencies.

Image arrays are computationally intensive and use a sizeable amount of RAM memory. This problem is notable apparent with the creation of the stacked multi-scale and multi-scale-frequency data cubes. To ensure that we can map the 3 data cube categories to the same vector space, and compute the data cubes for analyses; we used four different pre-trained CNN architectures (namely: Inception, Inception-Resnet, PNASNET and Resnet) to reduce the dimensionality of the image arrays. The pre-trained networks take as input, a single specified source array (300x300x1 for a single survey image, 300x300x6 for a stacked multi-scale survey image, and 300x300x6xN for multi-scale and N number of [multi-frequency] observed surveys, image), returning a feature engineered and dimensionally reduced (2048x1) column vector.

This feature engineered and dimensionally reduced column vector, is what we refer to as the physically abstracted feature vector space. We make this claim, since it is a product of the latent space from the pre-trained networks, which took as inputs observed continuum radio images. If the noise and different radio source components contained within an image, have physically consistent structures; then in principle, the pre-trained networks should be able to systematically feature engineer a collection of patterns (as descriptors of the features contained within an image) to be stored within the produced abstracted column vector. Should this be true, an architecture with a collection of pre-trained networks that have been fine-tuned to pick up different features of radio sources, is feasible. We note that the search for the appropriate collection of pre-trained networks to focus on the right features, is out of scope for this work. The collection of abstracted column vectors are then used as inputs to the SOM, producing binned classes of similar abstracted column vectors and presented in Tab. 1, for a data cube of 2137 radio sources observed by the WENSS survey at a FOV of 20′.

SOMs (self-organising maps) developed by Kohonen (1997, 2013) were used to systematically cluster column vectors with similar original or model abstracted image features (see Kohonen (1997, 2013) for details of the generalised metric and parameters, and Mostert et al. (2021) provide specifics on the adaptations made for the image domain). We trained the SOM run on an initialised 6x6 SOM map, with a learning rate of 0.5, neighbourhood radius of 0.5, and 2000 number of iterations.

To create controls/baselines of how the SOM would bin different sources from radio continuum surveys, we carry out the procedure described in Fig. 6.1, without using pre-trained models. In this step, each feature column vector (with dimensionality 90000x1) is passed into the SOM without abstraction. We then repeated the process for the different model abstracted column vectors (with dimensionality 2048x1). The results from these steps were used to construct the results presented in Tab. 1.

Although the concept of using latent features from the pre-trained networks, which are designed to have some degree of rotational and translational equivariance, is reasonable, we note that the idea of passing the flattened images directly lacks a solid theoretical basis. This is essentially trying to pass the pixels directly to a very simple neural network. Due to its extreme reliance on the source's position and rotation, it cannot reliably learn physically motivated features. As mentioned, its inclusion is for the purpose of establishing controls.

Since SOMs are an unsupervised technique, we do not have control over how they bin or name the outputted classes. However, they bin given inputs based on the similarities of the input column vectors. Using this fact, we can ignore the difference in the names of the produced classes, focusing on the bin counts and the intersection of sources that appear in the same classes across the binned original and model abstracted feature column vectors.

This is done by using a combination of class names as a barcode. That is, if source x is binned in classes (5, 3), (5, 5), (2, 5), (2, 5), and (4, 2) for the original image, Inception, Inception-Resnet, PNASNET, and Resnet model abstracted column vectors, respectively. Then, source x is given 53:55:25:25:42 as a barcode. Thus, sources with similar binned features to source x will have the same barcode, allowing us to construct a SOM class agreement table as presented in Tab. 1, and gauge which sources are consistently binned in the same classes.

The first part of Tab. 1, shows results from each conducted SOM experiment, considering data cubes from the WENSS, NVSS and TGSS surveys. The second part of Tab. 1 displays categorised binned counts based on sources that consistently occur in the same classes throughout the various SOM experiments, with groupings formed from the abstractors only results G(i:ir:p:r), and abstractors including the control group results G(o:i:ir:p:r).

Due to computational resources and time constraints, for the case of running images without using pre-trained models; we only ran the SOM on images from the original image data cubes, which have a specified FOV and survey. The un-abstracted column vectors could not be ran in reasonable time to produce results for both the multi-scale and multi-scale-frequency analyses. We also faced the same resource constraints with repeating the analyses for all of the abstractors using stacked arrays. Because of this, we opted to use PNASNET as the pre-trained network of choice, to abstract features from the multi-scale and multi-scale-frequency data cubes. We now present the finding from the followed approach.

6.3 Results and Discussions

To bin features/components of radio sources based on pixel intensities from observed radio continuum images, we used the original image, multi-scale, and multi-scale-frequency data cubes, described in Chap. 6.2. These were used as computational descriptors of different contexts based on data with a specified scale-frequency, multi-scales, and multi-scales-frequencies.

6.3.1 Single Scale and Frequency Features

Data cubes with a single specified FOV and survey form the simplest dateset to work with, and were treated differently to the other two categories. Using these data cubes, we carried out five experiments where we ran these single scale and frequency features on the SOM without treating the data to create control groups for each downloaded survey FOV. This process was then repeated using the data cubes to train each of the networks. This created feature-engineered column vectors, which made up the abstracted features of each downloaded survey FOV. The results are presented in Tab. 1 for the the smaller sample size of WENSS and (comparatively large) NVSS and TGSS surveys with (frequency, FOV) pairs of (325 MHz, 0.333°), (1400 MHz, 0.333°) and (150 MHz, 0.176°), respectively.

Comparing the sorted source counts from the different SOM outputted class bins, a consistent trend is observed. The different networks abstract similar features for each given survey, forming a prototypical grouping of extended radio sources with different orientations and extents. This forms an interesting contrast with the sorted source counts of the control group from each survey. The control group bins tend to have a comparatively larger second grouping, with the bin counts falling off quicker for the grouped sources because of the larger second grouping. We see a high level of agreement between the abstracted features G(i:ir:p:r), meaning even though the pre-trained use different architectures, they have similar descriptions of what is contained within the presented views (since they were trained on ImageNet). An expected and reproducible[6] finding from the TGSS survey, is that surveys with designs that are optimised for point sources produce single SOM bins, when using unabstracted features.

[6]Using data cubes from the FIRST and VLSSr surveys, with runs from each of the 6 FOVs.

Counts of SOM outputted class bins trained on 2137 images with 0.333° FOV from the WENSS survey				
original	inception	inc.-res.	pnasnet	resnet
1771	1839	1842	1840	1833
141	30	30	29	39
64	28	30	28	30
29	27	25	23	25
29	27	21	22	25
27	24	20	18	24
12	19	18	16	17
7	14	12	16	14
6	13	11	12	13
5	10	10	12	11
4	9	9	12	9
4	8	9	9	8
4	7	9	9	7
3	7	8	7	7
3	7	8	7	6
2	6	7	6	6
2	6	6	6	6
2	6	6	6	5
2	5	6	6	5
2	5	5	5	5
2	4	5	5	5
2	4	5	5	4
1	4	4	4	4
1	4	4	4	4
1	3	4	4	4
1	3	3	4	4
1	3	3	3	3
1	3	3	3	2
1	3	2	3	2
1	2	2	3	2
1	2	2	2	2
1	1	2	2	2
1	1	2	2	1
1	1	2	2	1
1	1	1	1	1
1	1	1	1	1

Counts of SOM outputted class bins trained on 10477 images with 0.333° FOV from the NVSS survey				
original	inception	inc.-res.	pnasnet	resnet
9753	10183	10187	10183	10188
385	68	68	68	69
176	34	34	34	34
72	27	27	27	27
22	25	25	25	26
13	22	22	22	18
7	11	11	11	11
5	11	11	11	11
5	8	8	11	10
3	8	8	8	8
3	7	7	8	7
3	6	6	7	6
3	5	6	5	5
3	5	5	5	5
2	5	5	5	5
2	5	5	4	4
2	4	5	4	4
2	4	4	3	4
1	4	4	3	3
1	3	3	3	3
1	3	3	3	3
1	3	3	3	3
1	3	3	3	3
1	3	2	3	3
1	3	2	3	2
1	2	2	2	2
1	2	2	2	2
1	2	2	2	2
1	2	2	2	2
1	2	1	2	2
1	2	1	1	1
1	1	1	1	1
1	1	1	1	1
1	1	1	1	1
-	1	-	1	1
-	1	-	-	-

Counts of SOM outputted class bins trained on 12304 images with 0.176° FOV from the TGSS survey				
original	inception	inc.-res.	pnasnet	resnet
12304	11594	11603	11594	11594
-	115	132	102	116
-	72	93	69	72
-	56	55	59	56
-	52	48	55	52
-	50	48	52	51
-	49	46	51	51
-	43	39	48	46
-	35	33	47	39
-	33	29	26	33
-	30	23	23	27
-	19	15	22	22
-	16	14	19	14
-	13	11	13	13
-	11	9	13	10
-	11	8	10	10
-	9	8	10	9
-	7	8	10	9
-	7	8	10	8
-	7	7	9	8
-	7	7	7	8
-	7	6	7	6
-	7	6	6	6
-	7	6	6	6
-	6	5	5	5
-	6	5	5	5
-	5	5	4	4
-	5	4	4	4
-	5	4	4	4
-	4	4	3	3
-	4	3	2	3
-	4	3	2	3
-	3	3	2	2
-	2	3	2	2
-	2	2	2	2
-	1	1	1	1

	WENSS correlated SOM barcoded class bins trained on 2137 images with 0.333° FOV				NVSS correlated SOM barcoded class bins trained on 10477 images with 0.333° FOV				TGSS correlated SOM barcoded class bins trained on 12304 images with 0.176° FOV			
	G(o:i:ir:p:r)	Counts	G(i:ir:p:r)	Counts	G(o:i:ir:p:r)	Counts	G(i:ir:p:r)	Counts	G(o:i:ir:p:r)	Counts	G(i:ir:p:r)	Counts
0	32:55:25:25:42	1517	55:25:25:42	1832	21:45:00:53:00	9477	45:00:53:00	10183	00:33:10:40:05	11594	33:10:40:05	11594
1	43:55:25:25:42	122	00:42:23:51	29	32:45:00:53:00	377	53:13:35:02	68	00:32:20:50:14	100	32:20:50:14	100
2	42:55:25:25:42	61	41:05:22:02	20	42:45:00:53:00	172	31:12:34:55	34	00:13:12:20:21	54	13:12:20:21	54
3	53:55:25:25:42	25	40:11:42:05	19	52:45:00:53:00	68	54:22:24:40	26	00:43:00:30:04	52	43:00:30:04	52
4	52:55:25:25:42	24	50:10:43:13	17	21:53:13:35:02	63	42:03:25:03	24	00:22:35:00:33	42	22:35:00:33	42
5	32:00:42:23:51	22	33:43:13:21	17	21:31:12:34:55	32	55:21:45:30	18	00:02:15:03:31	41	02:15:03:31	41
6	54:55:25:25:42	19	30:21:44:03	14	21:54:22:24:40	25	34:20:44:20	11	00:34:00:41:15	40	34:00:41:15	40
7	32:41:05:22:02	18	11:02:10:14	9	21:42:03:25:03	22	13:41:33:23	11	00:31:13:12:34	37	31:13:12:34	37
8	32:33:43:13:21	16	43:33:35:22	9	22:45:00:53:00	21	25:01:43:11	8	00:52:14:01:42	35	52:14:01:42	35
9	32:50:10:43:13	13	04:05:32:30	8	21:55:21:45:30	18	51:15:10:05	7	00:15:24:15:24	24	15:24:15:24	24
10	32:30:21:44:03	12	30:21:45:13	8	35:45:00:53:00	13	33:33:42:34	5	00:21:32:51:22	19	21:32:51:22	19
11	32:40:11:42:05	12	35:15:14:44	7	21:34:20:44:20	11	30:34:03:22	5	00:42:25:12:12	18	42:25:12:12	18
12	33:55:25:25:42	10	22:30:34:03	7	21:25:01:43:11	8	41:14:02:04	5	00:43:20:30:04	17	43:20:30:04	17
13	32:43:33:35:22	8	44:35:15:32	7	21:13:41:33:23	8	23:53:30:33	5	00:32:20:21:14	14	32:20:21:14	14
14	32:30:21:45:13	7	54:45:04:53	6	11:45:00:53:00	7	52:04:04:04	5	00:42:25:01:12	12	42:25:01:12	12
15	32:04:05:32:30	7	42:55:12:54	6	21:51:15:10:05	6	55:21:45:41	4	00:53:45:04:44	10	53:45:04:44	10
16	32:22:30:34:03	7	12:54:31:00	5	21:30:34:03:22	5	24:43:42:24	4	00:20:31:24:43	10	20:31:24:43	10
17	32:35:15:14:44	6	31:31:01:45	5	21:52:04:04:04	5	50:24:20:14	4	00:24:21:31:23	9	24:21:31:23	9
18	32:44:35:15:32	6	41:01:22:02	5	21:33:33:42:34	5	05:30:05:31	4	00:34:10:41:15	9	34:10:41:15	9
19	45:55:25:25:42	6	43:44:03:22	5	31:45:00:53:00	5	35:00:52:00	3	00:12:40:52:11	7	12:40:52:11	7
20	32:42:55:12:54	6	20:20:00:24	4	03:45:00:53:00	5	52:04:04:03	3	00:11:42:22:45	7	11:42:22:45	7
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

This is because the SOM is unable to learn correlations between pixels (see Fig. 6.4b) in the same way that CNNs can (see Fig. 6.5b), and so has no real distinguishing power. These factors are the reasons why the abstractor's approach is recommended for obtaining tangible outcomes from the data cubes.

We further see a small variation in the highest source count classes, across all abstractors, and sources that are perceived to have unique features tend to be binned in their own classes. The sources between these extremes are interesting. They exhibit minor shifts between the binned classes, for the different pre-trained architectures. The cause of these variations is of particular interest for this study since, we speculate that, some of the variation might be due to non-thermal diffuse emission. If this is true, there should be groupings that have a similar statistical structure due to the contribution of this emission.

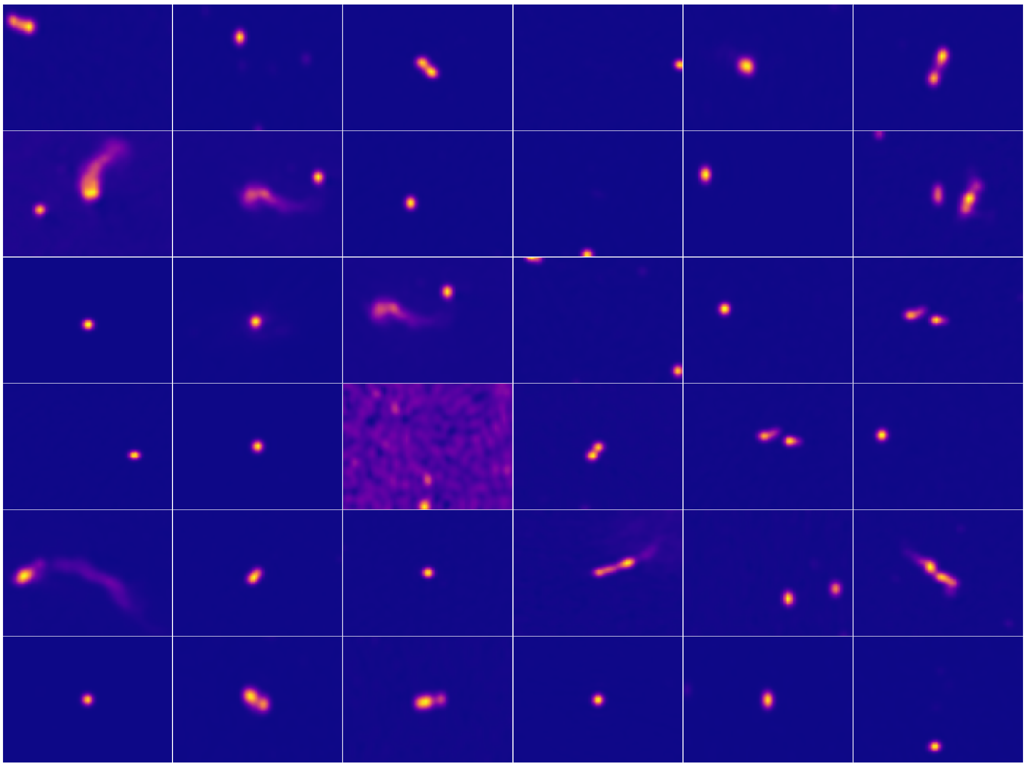
This form of tabular-based reasoning showcases the possible level of granularity we can achieve by using a combination of abstractors to search for and track changes between abstracted features and the original data features (in the case of rotation and translation equivariant SOMs). It grants us a representation of the data that effectively links the insights gained from the different sets of analyses. The data from the organised tables may be easily connected to the visual representations of the radio sources using the SOM count matrix and class representative images (as presented in Figs. 6.2b [the control bins] and 6.3b [the PNASNET abstracted features bins] for the WENSS).

We can visually compare these two representations of the bins, computed using the original and abstracted feature vectors, and link it to the analyses of the partitioned classes. From Fig. 6.2b we observe that the classes (3, 2), (4, 3), (4, 2), (5, 3), (5, 2), (5, 4) and (4, 3) are clustered around the same region, forming a local grouping of objects with similar features. They differentiate the largest count bin of Fig. 6.3b into sub-groupings with varying extents of double, extended, elongated, and point-like sources. This illustrates that most of the largest SOM bins are a collection of a generalised view of the source populations within the data sets, and the proceeding bins (with lesser counts are) formed from populations which show variance from the main grouping.

Even though the control group features show the capacity to differentiate sub-groups within the main grouping, from the abstracted features grouping; its failure to distinguish that the sources contained within the pairs of classes (1, 1) and (2, 2) should be in the same grouping, and the source in class (3, 4) should be binned into class (2, 5), highlights shortcomings of SOM bins formed from untreated data. That is, it is not optimal to use four bins when the sources should be binned into two bins. PNASNET abstracted features

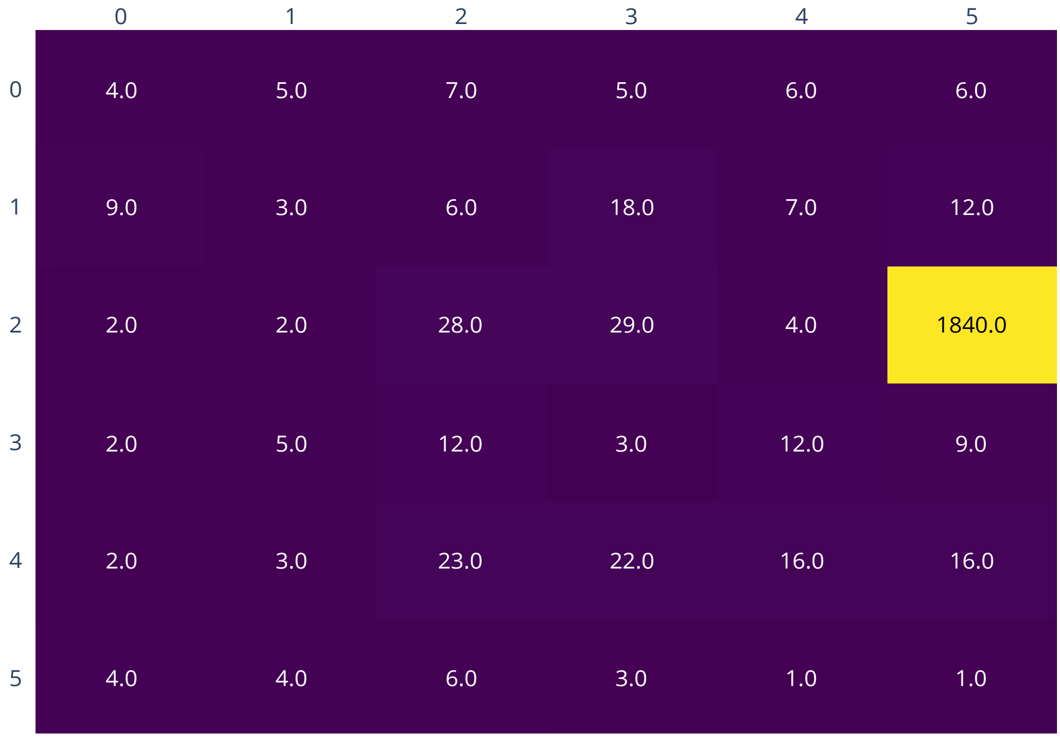
	0	1	2	3	4	5
0	1.0	1.0	4.0	2.0	4.0	4.0
1	2.0	1.0	2.0	1.0	1.0	3.0
2	1.0	6.0	1.0	1.0	1.0	2.0
3	1.0	5.0	1771.0	12.0	1.0	2.0
4	1.0	2.0	64.0	141.0	2.0	7.0
5	3.0	1.0	29.0	27.0	29.0	1.0

(A)

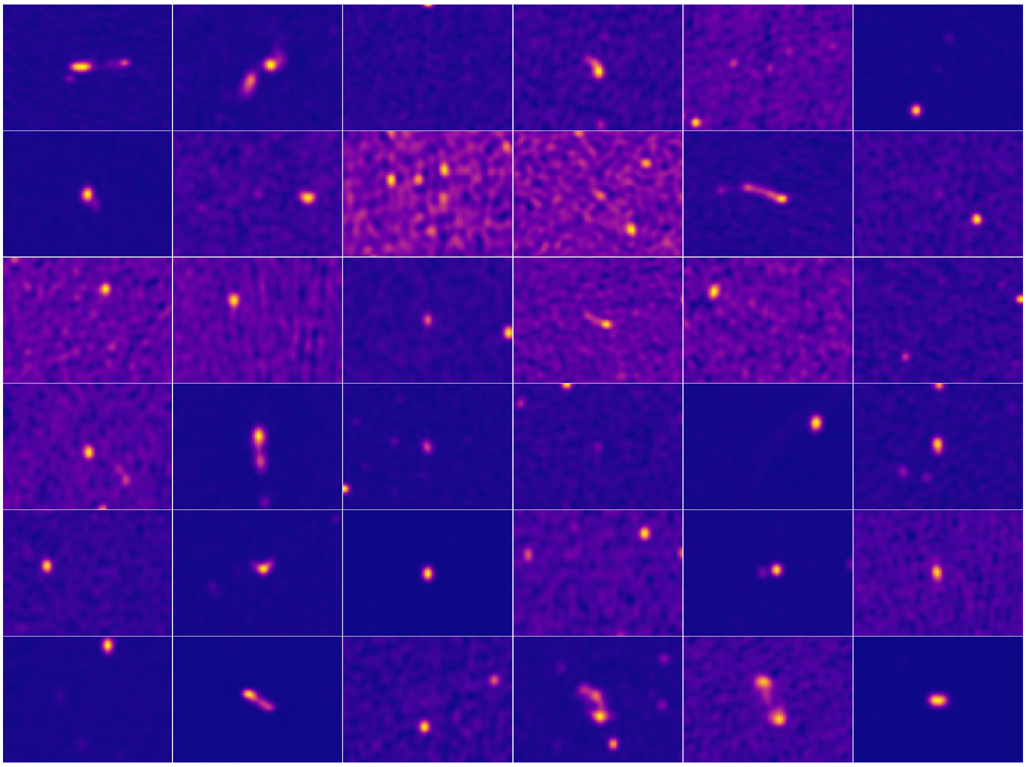


(B)

FIGURE 6.2: (A): Count matrix for the WENSS survey at 0.333° using 2137 untreated images and its (B) SOM class representative images for the above mentioned data.



(A)

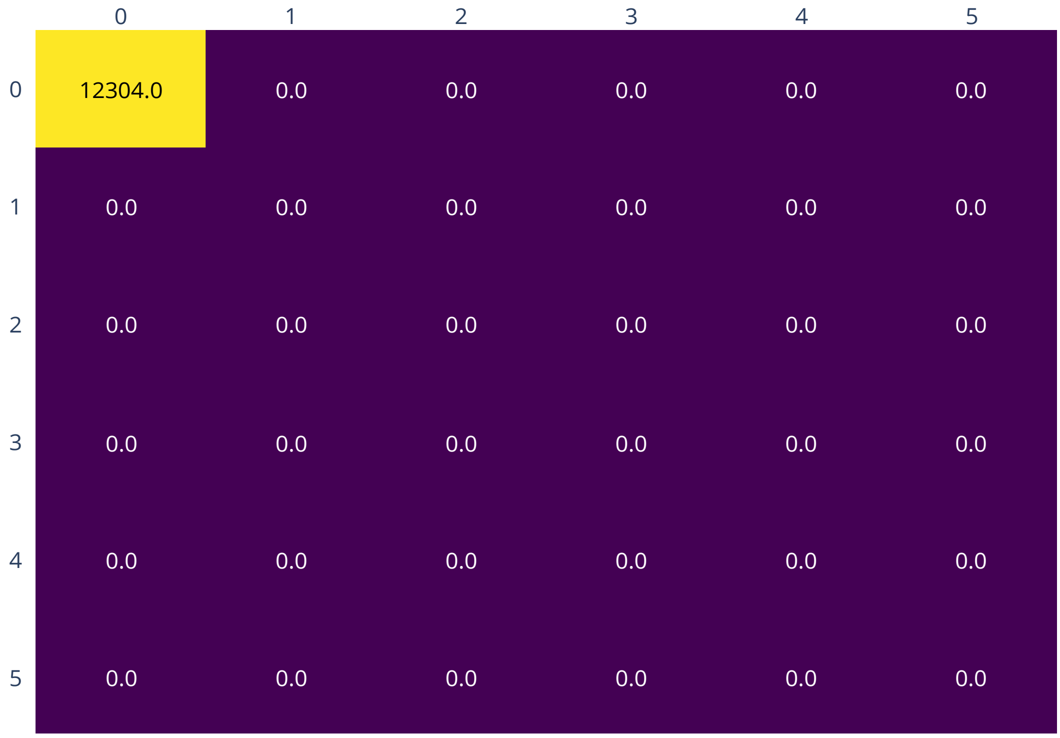


(B)

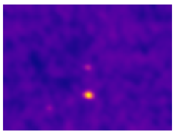
FIGURE 6.3: (A): Count matrix for the WENSS survey at 0.333° using 2137 PNASNET abstracted features and its (B) SOM class representative images.

are marginally favoured, thus far, based on their ability to distinguish features from telescopes optimised for point source detection (see Fig. 6.5b), and forming bins which are not susceptible to small (image scene) changes.

It is also interesting to observe that SOMs can perceive images with masked regions, Fig. 6.7b, a useful feature for sieving out data with blank pixels. Having shown that SOMs can bin source components (based on whether an image scene has noisy backgrounds, single, elongated, extended, multiple sources, or a combination of these features, with varying extents) using pixel intensities, from features that describe radio source dynamics at a single scale and frequency; we now extend our analyses to data cubes that have multiple scales.

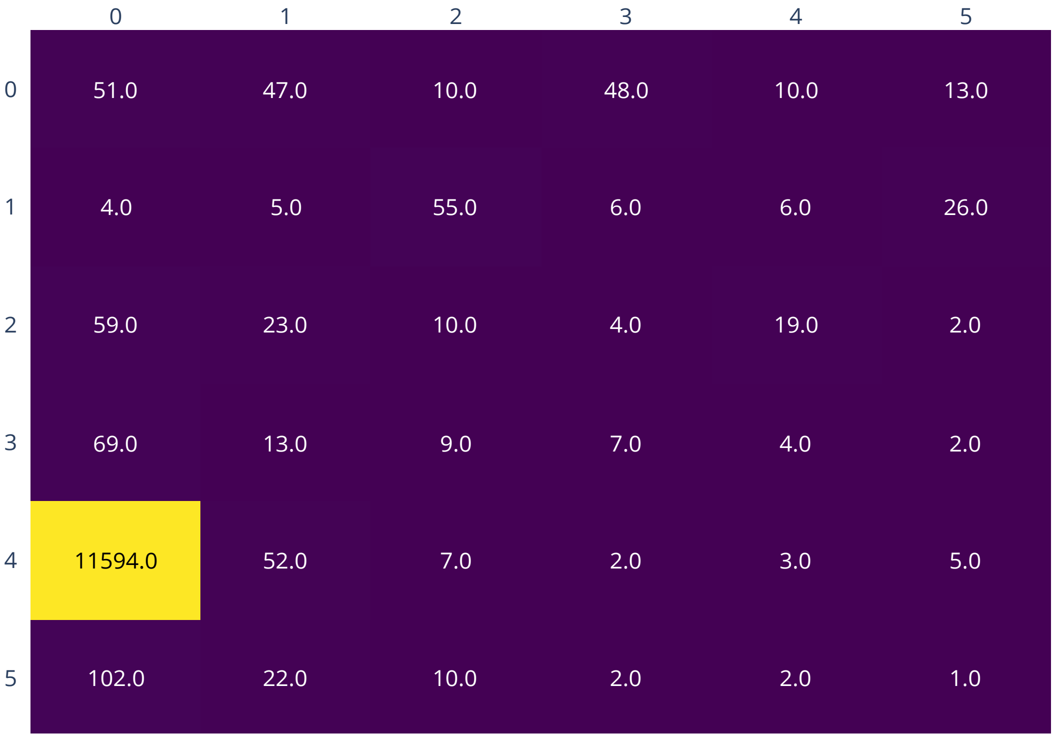


(A)

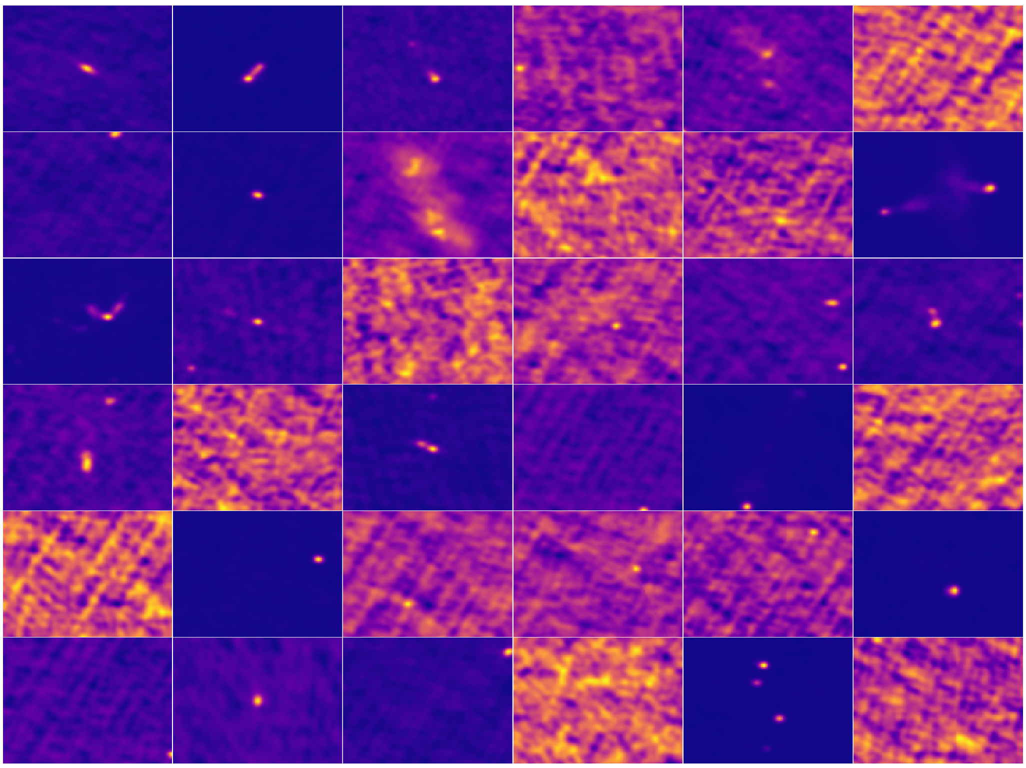


(B)

FIGURE 6.4: (A): Count matrix for the TGSS survey at 0.176° using 12304 untreated images and its (B) SOM class representative images for the above mentioned data.



(A)



(B)

FIGURE 6.5: (A): Count matrix for the TGSS survey at 0.176° using 12304 PNASNET abstracted features and its (B) SOM class representative images.

6.3.2 Multi-scale and Single Frequency Abstracted Features

Proceeding to work on data cubes that have a collection of FOVs and a specified survey frequency, we abstracted the multi-scale features and used them to produce multi-scale binned SOM classes (as presented in Figs. 6.6a, 6.9a, 6.7a, 6.8a, and 6.10a for the WENSS, FIRST, NVSS, TGSS, and VLSSr surveys, respectively). These data cubes are smaller in size as a consequence of missing images from some of the downloaded survey FOVs.

Multi-scale data cubes are hampered by the ability to abstract features from radio galaxy to galaxy cluster and cosmological scales. The computed bins, from multi-scale abstracted features, show a notable improvement in binning by geometry and how pixel intensities fall off, within images with differing dominant source components. Comparing the features abstracted at a single scale and frequency (Fig. 6.3b) with the multi-scale abstracted features (Fig. 6.6b), we note that the trained SOM only made use of 32 bins out of the available 36 initialised bins. This is encouraging to observe, since, it shows that the SOMs are focusing more on binning the similarities of the abstracted features, rather than optimising how to spread them into as many bins as is possible.

The images presented in Fig. 6.6b, have a more "expert" like curated view. Conveniently binning elongated, centred, and bright point-like sources (with high pixel intensities compared to the background noise) in bins close to the SOM class (0, 0). Within this neighbourhood, classes (1, 1) and (2, 0) show groupings of extended radio sources that have point-like sources that are similar to the surrounding sources. More point-like sources appear in classes (1, 3), in the neighbourhood of multi-point-like sources, doubles and elongated doubles. Looking at trends closer to the bottom right quadrant, we see groupings of extended sources with varying spatial extent, noise dominated images [in class (3, 5)], and triples being the other extreme [class (5, 5)].

Finally, looking at the dominant class (5, 2) showcases the earlier expressed thought on SOMs tending to form a "generalised" major grouping of heterogeneous source features that have statistical similarities, and smaller surrounding bins with more homogeneous source features. The surrounding neighbourhood of the dominant class is a collection of small groupings exhibiting extended emission features [class (4, 3)], multi-component sources that exhibit similar geometric features to triples [class (5, 3)], doubles [class (5, 1)], and a combination of extend-elongated features [class (4, 2)]. Looking at the trend within the first column and starting from the bottom [class (5, 0)], we see the centred bins transition from doubles to the earlier described single elongated sources. Similarly binned, physical regions are observed in Figs. 6.7b and 6.8b, for the NVSS and TGSS survey, respectively. The differences in how the bin classes are curated is due to the SOM map having autonomy in

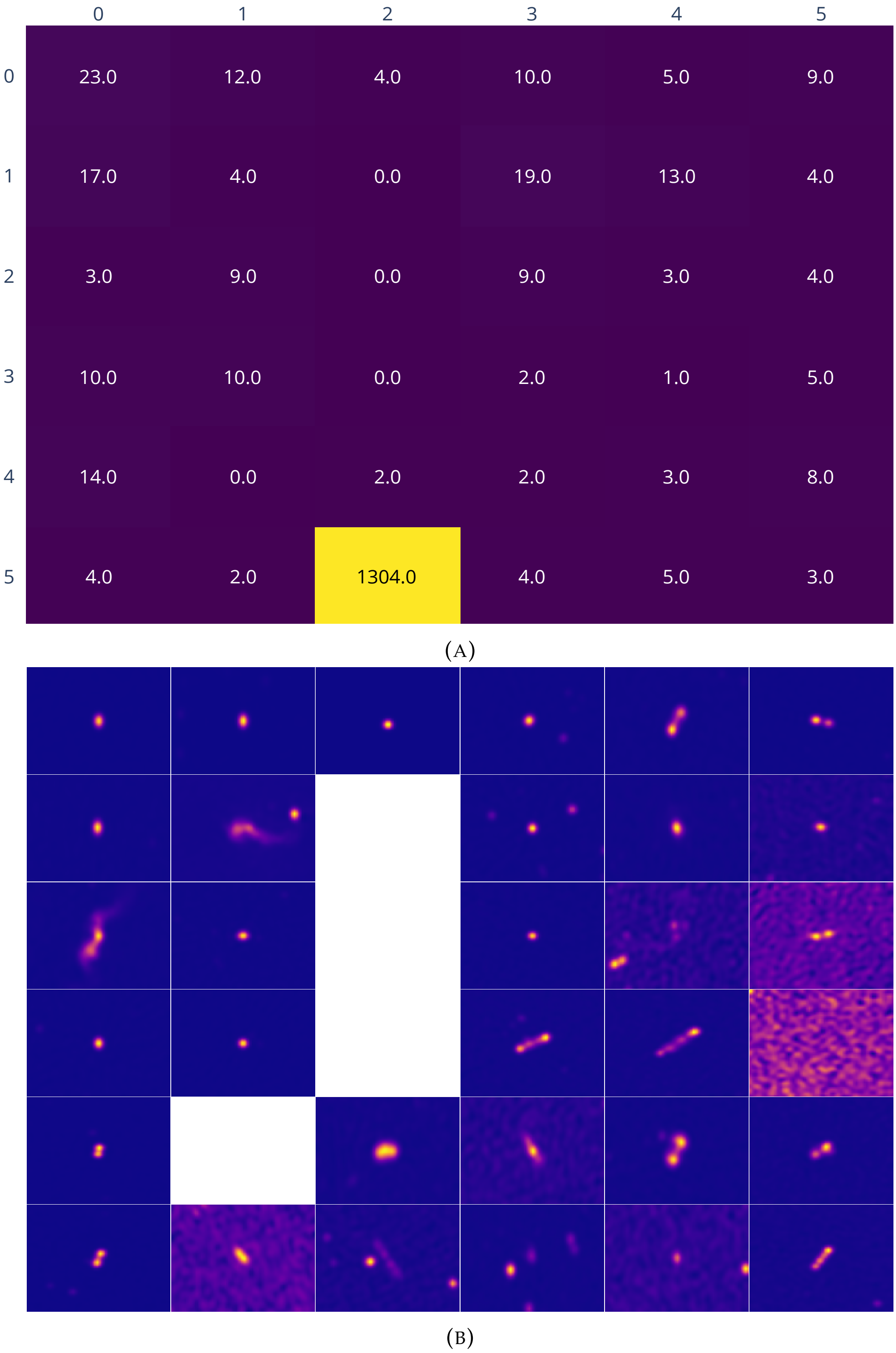


FIGURE 6.6: (A): Count matrix for the WENSS survey at 0.333° using 1527 multi-scale WENSS PNASNET abstracted features and its (B) SOM class representative images.

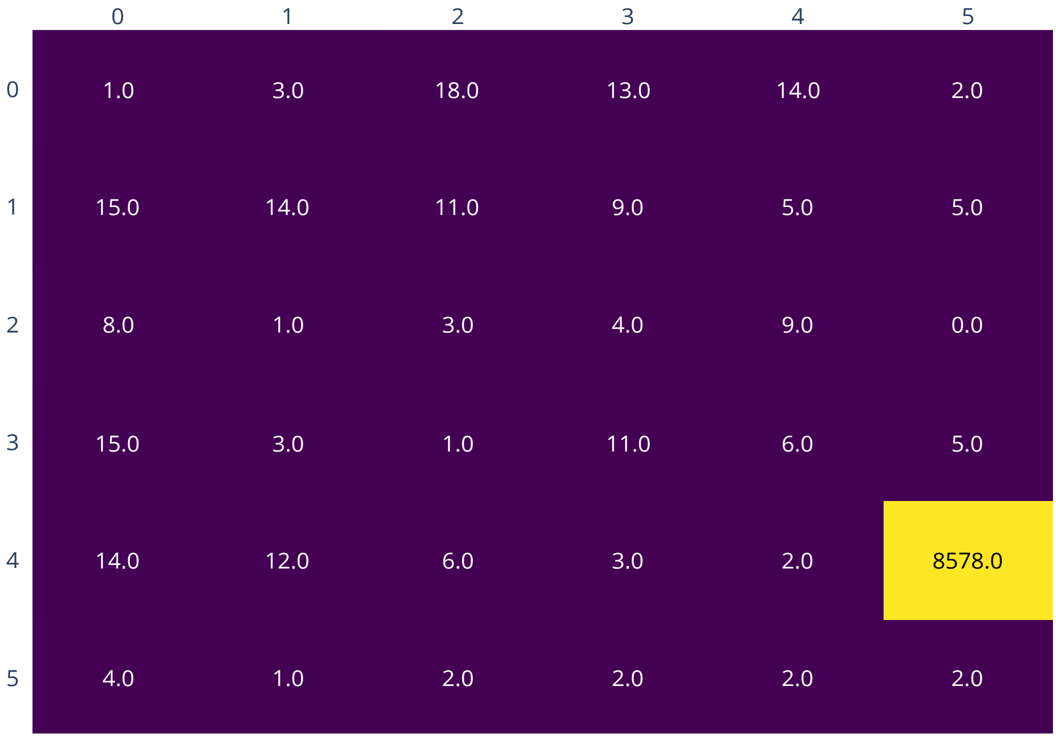
choosing where to place the perceived differences.

The produced TGSS (as well as FIRST) SOM multi-scale feature bins, are the best motivating examples for favouring the usage of multi-scale survey abstracted features over unabstracted features. By comparing the produced SOM class representative images, of the bins produced from: unabstracted features (Fig. 6.4b), single scale and survey abstracted features (Fig. 6.5b), and multi-scale and single survey abstracted features (Fig. 6.8b). We have effectively progressed from the perspective of SOM bins which cannot distinguish features contained within continuum radio images (Fig. 6.4b), to physical bins which show atypical morphologies as seen in class (5, 0) of Fig. 6.8b.

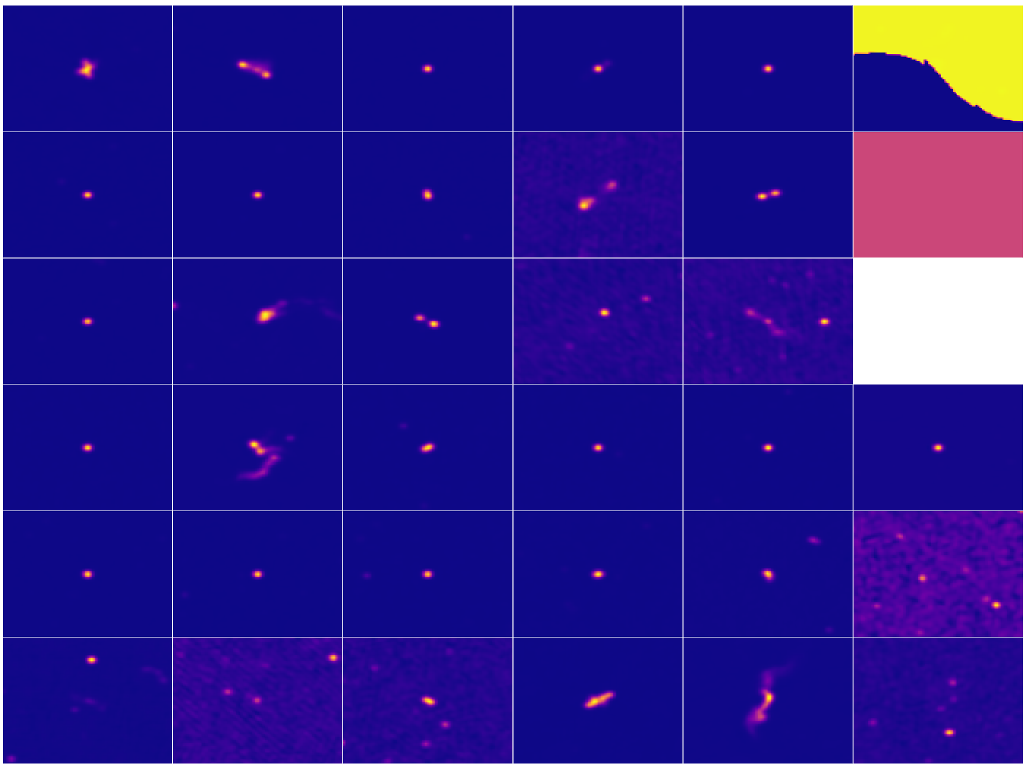
Class (5, 0) also highlights how we can link insights from different SOM runs, to find similar source populations in one multi-scale survey abstracted features bin based on another SOM run. Observing that classes (5, 0) and (5, 4) share the same abstracted features source in the TGSS and NVSS surveys (Figs. 6.8b and 6.7b), respectively, grants us the ability of tracking similar source populations across different surveys.

Fig. 6.9b shows the rarity of other source populations, especially ones with extended emission features, in the FIRST survey. This is not a shocking discovery, since the long survey baselines were optimised for picking up point sources, at a high angular resolution. The NVSS survey, is the complimentary VLA survey for picking up extended sources. Its short survey baseline and low angular resolution, are optimised for extended radio emission features, Figs. 6.7b.

Gaining physical insights from low-frequency radio surveys is typically a daunting task, since most legacy low-frequency surveys have low angular resolutions (see the synthesized beam sizes in Tab. 6.1 for comparison) and the images tend to be noise dominated. To observe differentiated extended feature bins [as seen in classes (2, 0), (0, 2), (0, 3), (2, 3), (4, 2), and (5, 2) from Fig. 6.10b] in a low-frequency legacy radio survey optimised for point source detection, is a notable result, as seen 6.10a. The abstracted multi-scale features approach, therefore, provides an effective, autonomous framework for mining insights from low-frequency legacy radio survey. We now look at the final set of data cubes.

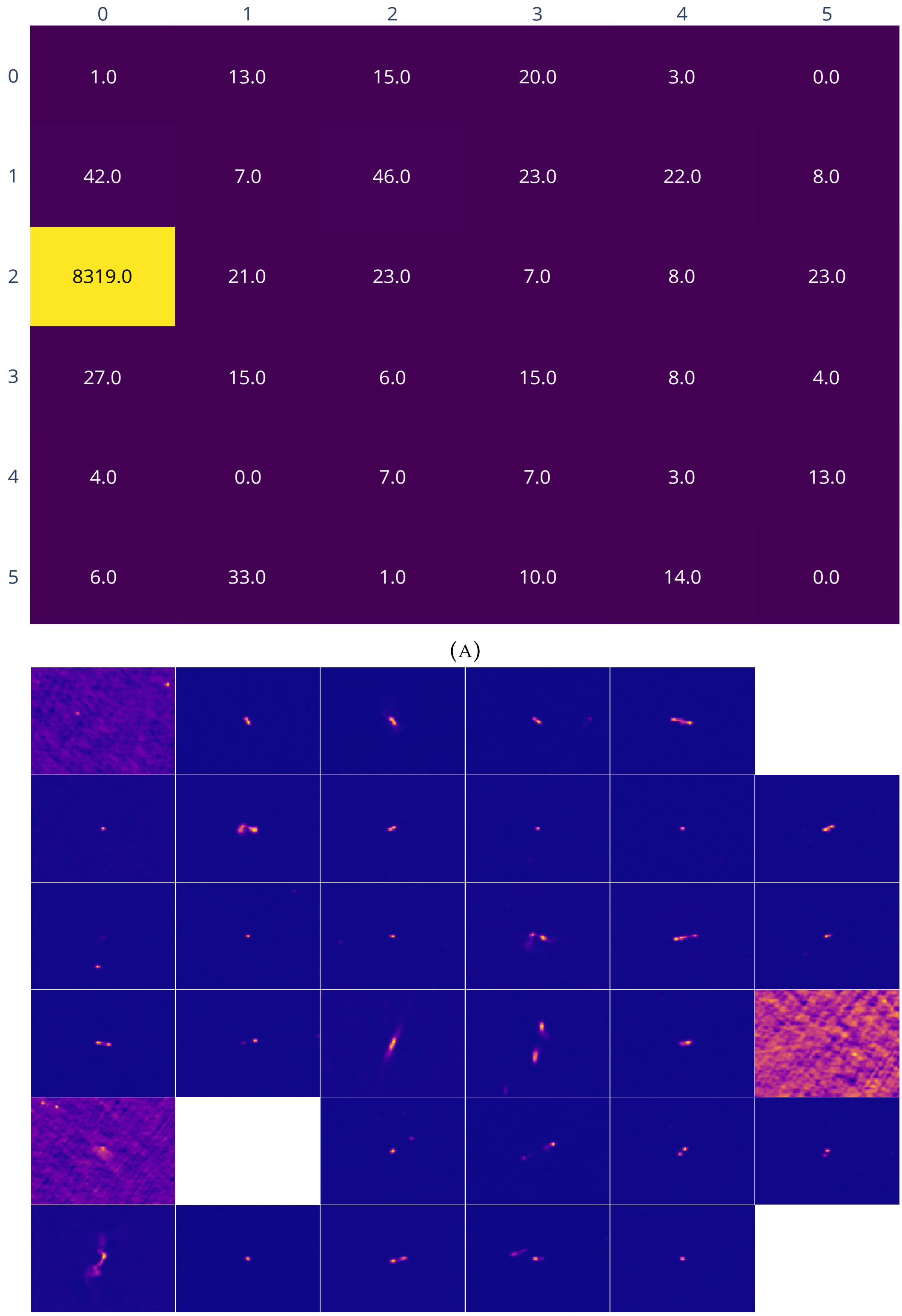


(A)



(B)

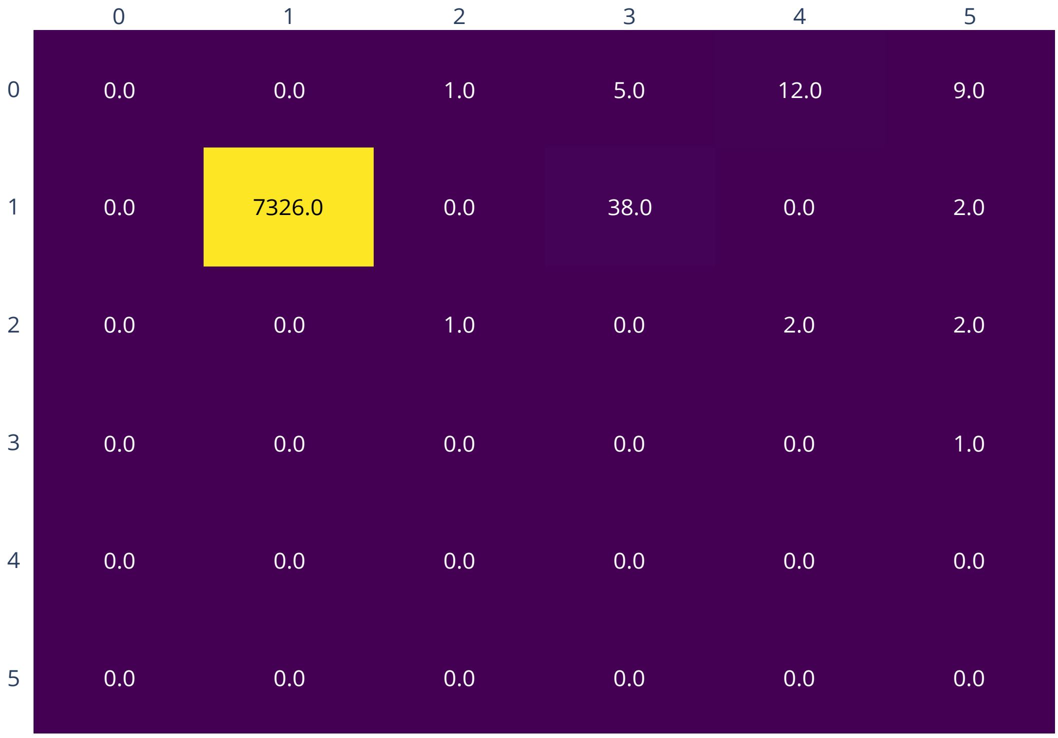
FIGURE 6.7: (A): Count matrix for the NVSS survey at 0.333° using 8804 multi-scale NVSS PNASNET abstracted features and its (B) SOM class representative images.



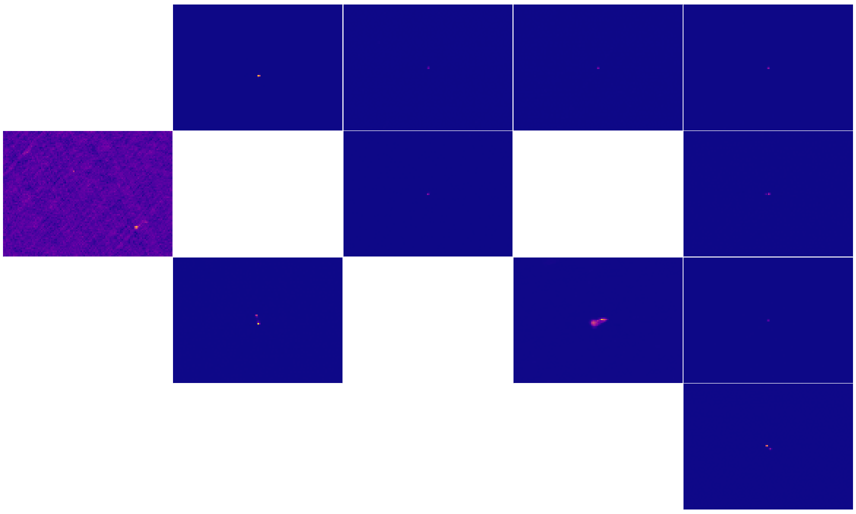
(A)

(B)

FIGURE 6.8: (A): Count matrix for the TGSS survey at 0.333° using 8774 multi-scale TGSS PNASNET abstracted features and its (B) SOM class representative images.



(A)

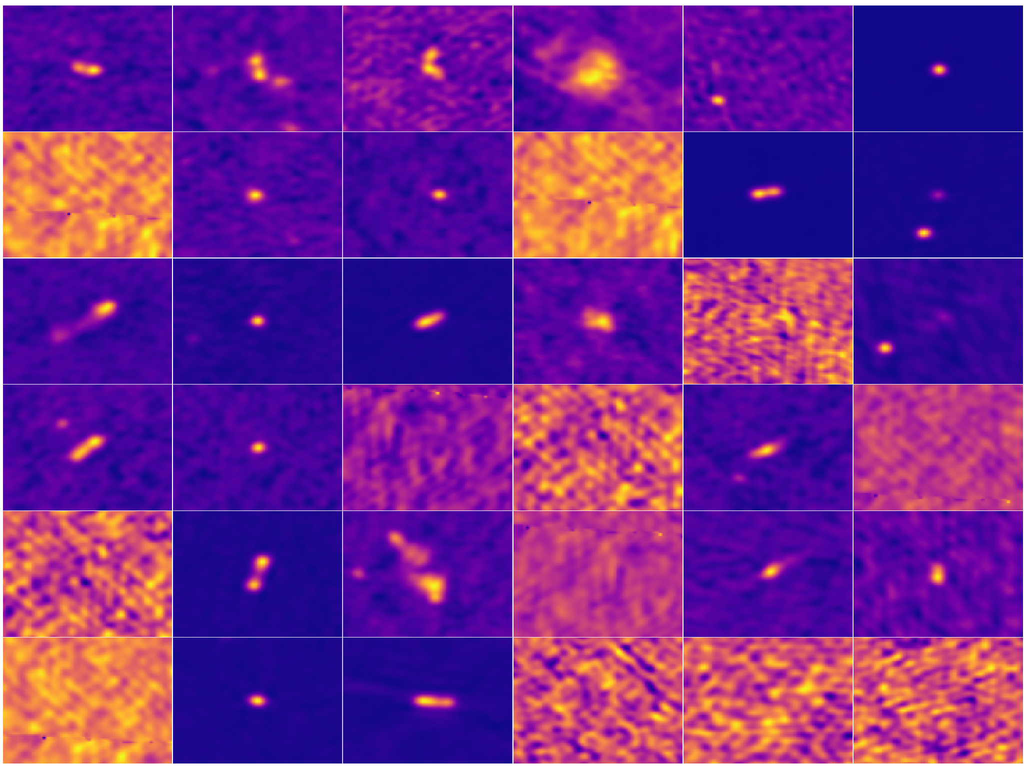


(B)

FIGURE 6.9: (A): Count matrix for the FIRST survey at 0.176° using 7377 multi-scale FIRST PNASNET abstracted features and its (B) SOM class representative images.

	0	1	2	3	4	5
0	25.0	67.0	59.0	2.0	5.0	12.0
1	6.0	62.0	21.0	1.0	4.0	96.0
2	36.0	72.0	20.0	41.0	23.0	38.0
3	33.0	49.0	10.0	58.0	54.0	6.0
4	13.0	21.0	16.0	202.0	44.0	69.0
5	11.0	96.0	37.0	13.0	111.0	7234.0

(A)



(B)

FIGURE 6.10: (A): Count matrix for the VLSSr survey at 0.333° using 8667 multi-scale VLSSr PNASNET abstracted features and its (B) SOM class representative images.

6.3.3 Multi-scale-frequency Abstracted Features

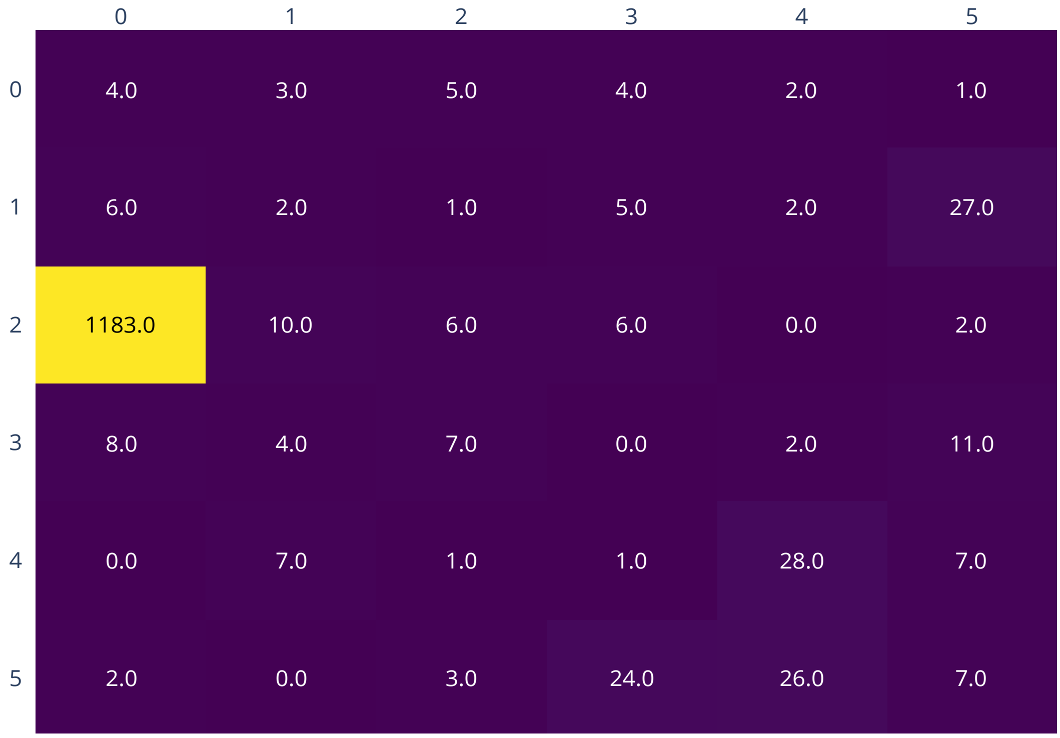
Having shown that multi-scale abstracted features form improved SOM bins, clustering radio source components based on physically explainable features. We investigate if training on a combination of multi-scale features from a number of radio survey frequencies, improves the observed SOM binning capabilities. Working on data cubes with multiple FOVs and observed frequencies, we abstracted multi-scale-frequency features and used them to produce multi-scale-frequency binned SOMs classes (as presented in Figs. 6.11b, 6.13b, and 6.12b for the combination of (TGSS:WENSS:NVSS), (TGSS:NVSS), and (VLSSr:NVSS:FIRST) survey features, respectively.

With this approach, instead of trying to link insights from different source populations by tracking similar source populations across different surveys; we can directly abstract multi-scale-frequency features for survey combinations of interest. For example, using the combination of TGSS (as a point source detector at low-frequencies, with enough sensitivity to detect extended features at a 150 MHz), WENSS (as an extended features detector at a 325 MHz) and NVSS (as another extended features detector at a 1400 MHz) we can produce Fig. 6.11b.

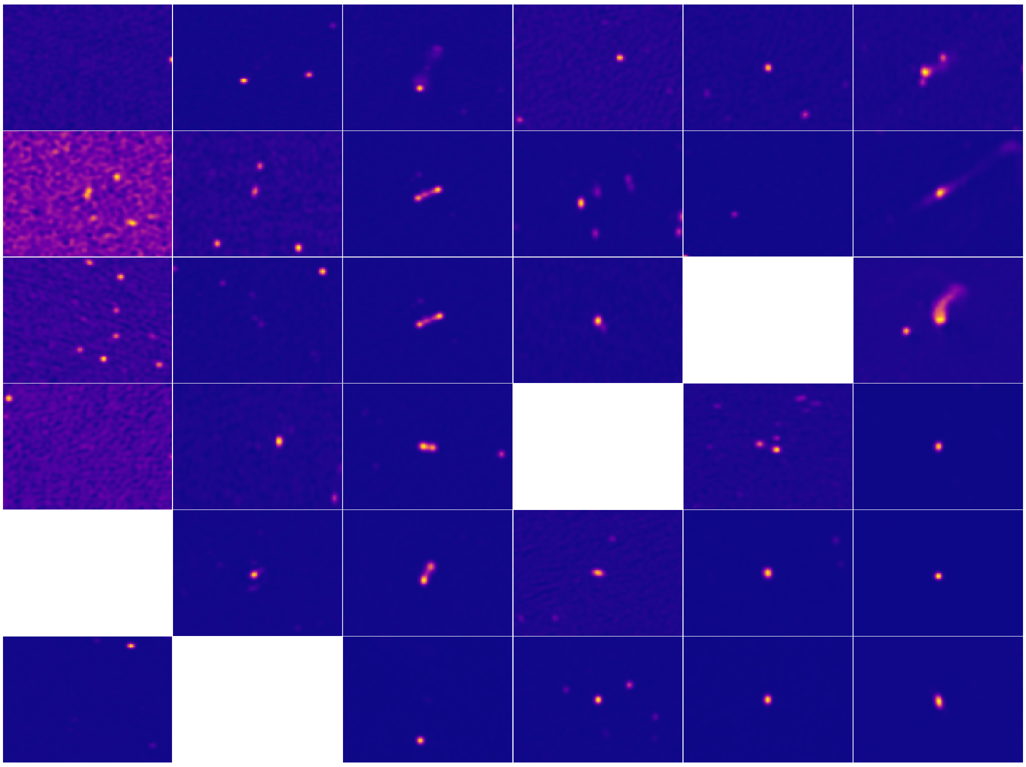
Comparing the results of the (TGSS:WENSS:NVSS) multi-scale-frequency binned SOMs classes, Fig. 6.11b, with the multi-scale abstracted features of the WENSS survey, Fig. 6.6b. We observe that the multi-scale abstracted features bins focused on differentiating the type of extension (elongated, doubles, triples, extended and so on) and its physical extent. Which makes sense since the WENSS survey was optimised for extended sources. Thus, the abstractors then abstract features based on multi-scale representations (of inputs) that are optimised for extended features.

The multi-scale-frequency binned SOMs classes in Fig. 6.11b seem to add a level of abstraction that further differentiates the bins, based on the number of sources in a given image scene. They also seem to form groupings based on: plumes class (2, 5), long extensions class (1, 5), disconnected "patchy" extensions class (1, 3), symmetric (bright) jets [most of the classes in column 2], off-centred sources classes (5, 0) and (5, 2), and multi-sources bins [classes in the neighbourhood of class (2, 0)]. This is an opportune time to reiterate that, the presented images are randomly sampled images of each class, and there was no human intervention in curating the observed classes.

Another combination of surveys we considered for curating a larger data set, is the VLA combination of: VLSSr (as a point source detector at 74 MHz), NVSS (as a extended features detector at a 1400 MHz) and FIRST (as another point source detector at 1400 MHz) surveys, used to produce Fig. 6.12b. An interesting observation from using this combination is, since we used two point source optimised surveys, the observed bins show diverse groupings



(A)

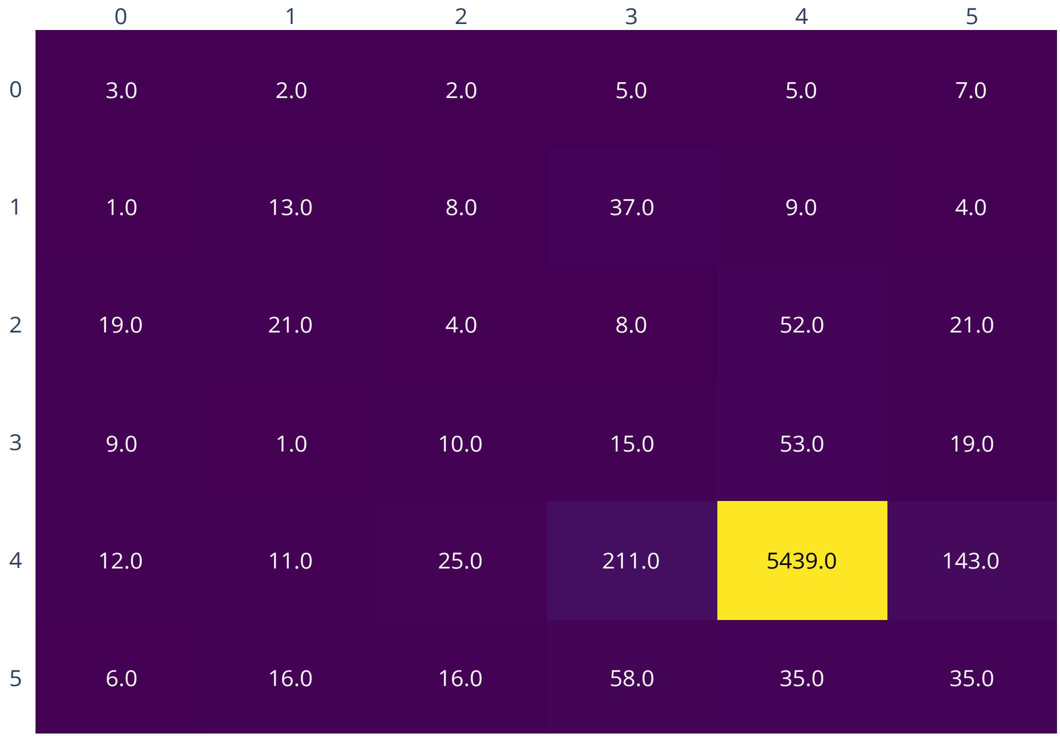


(B)

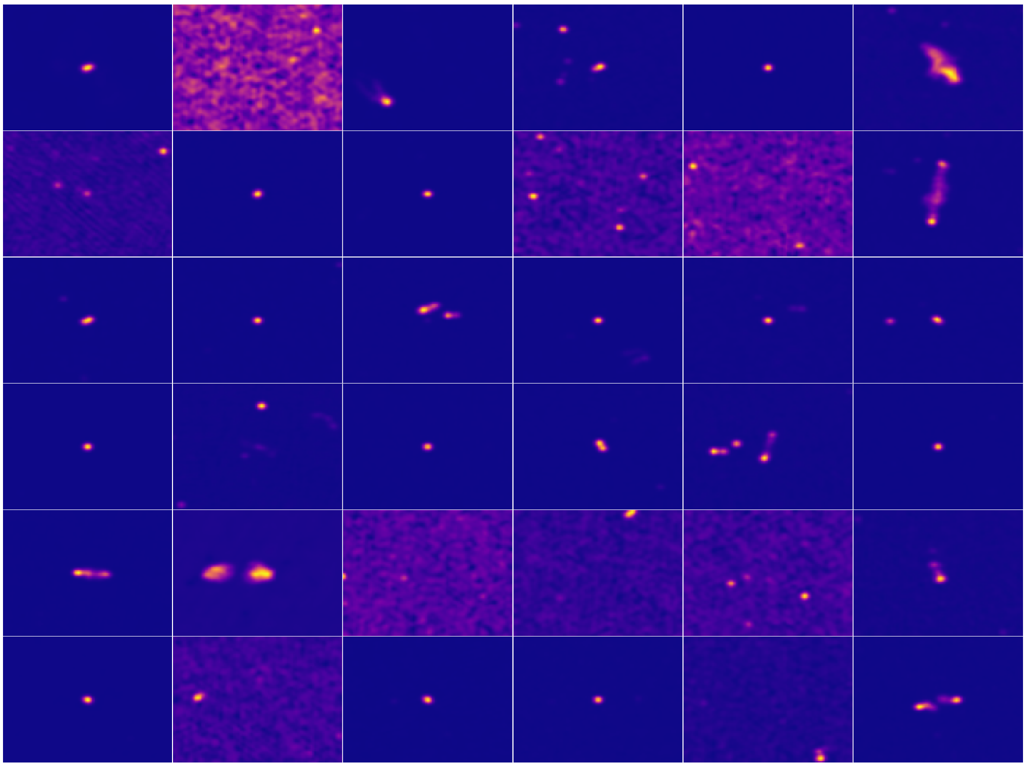
FIGURE 6.11: (A): Count matrix for the multi-scale-freq (TGSS:WENSS:NVSS) features, imaged using WENSS survey at 0.5° with 1407 PNASNET abstracted features and its (B) SOM class representative images.

of point source configurations, as well as extended features. This highlights the importance of (choice when) curating combinations of surveys to abstract. These should be based on physical features (and dynamics) of interest, in an investigation.

Considering the last combination of surveys we investigated in this exploration, we chose to produce binned classes from features that were abstracted using TGSS (as a point source/sufficient extended features detector at low-frequencies) and NVSS (as an extended features detector at mid-frequencies) surveys, producing Fig. 6.13b. This exploration uncovered a wide range of extended sources with bent-tails classes (0, 2) and (5, 4), different configurations of extended features, and generally interesting groupings such as those in classes (3, 4), (1, 5), (2, 5) and (4, 2). We therefore, have demonstrated how to use dimensional reduction techniques as abstractors of physical features in radio continuum images, using the previously described data cubes as inputs with different contexts.



(A)

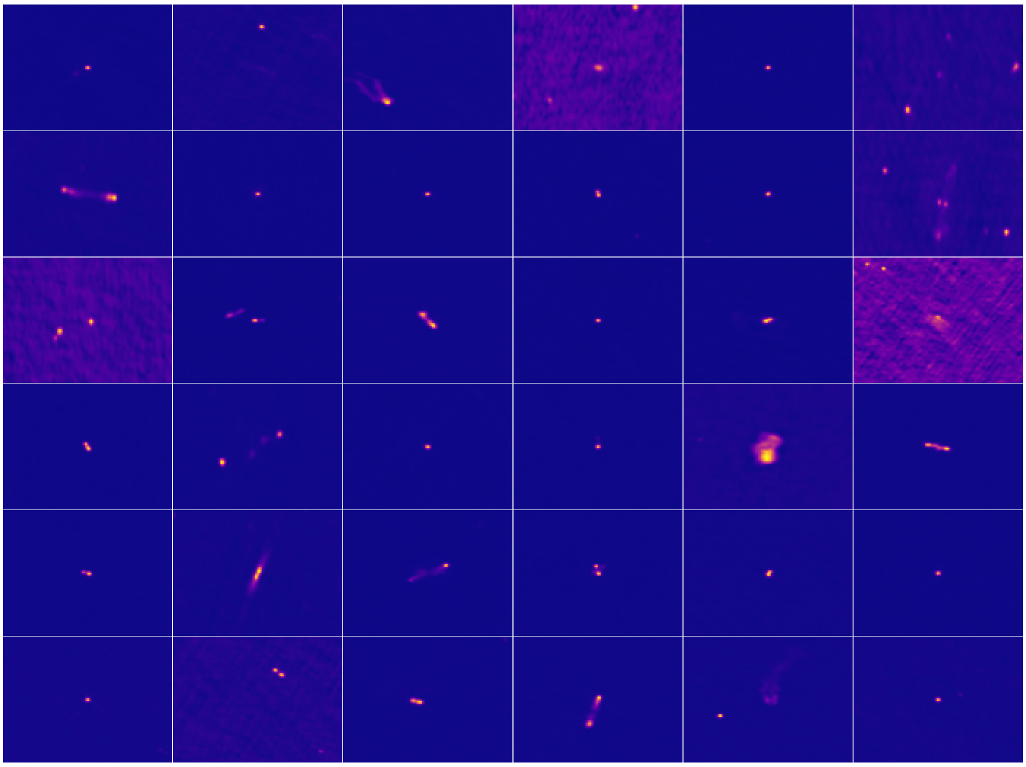


(B)

FIGURE 6.12: (A): Count matrix for the multi-scale-freq (VLSSr:NVSS:FIRST) features, imaged using NVSS survey at 0.333° with 6335 PNASNET abstracted features and its (B) SOM class representative images.

	0	1	2	3	4	5
0	11.0	1.0	8.0	2.0	8.0	5.0
1	18.0	10.0	7.0	12.0	4.0	3.0
2	4.0	12.0	8.0	9.0	6.0	4.0
3	10.0	20.0	10.0	17.0	3.0	8.0
4	13.0	18.0	9.0	15.0	5.0	9.0
5	78.0	7777.0	31.0	8.0	2.0	22.0

(A)



(B)

FIGURE 6.13: (A): Count matrix for the multi-scale-freq (TGSS:NVSS) features, imaged using TGSS survey at 0.333° with 8187 PNASNET abstracted features and its (B) SOM class representative images.

6.4 Conclusions

A range of publicly available legacy radio continuum surveys, with varying FOVs and frequencies, were utilised to design a functional architecture, for the task of abstracting features contained within extended radio sources. Using pre-trained networks (as abstractors), an unsupervised model, and applying transfer learning principles to images from continuum radio surveys; we used data cubes from three categories (namely: original image, multi-scale, and multi-scale-frequency data cubes) to test the binning capabilities of the functional architecture for extended radio sources.

We found that the functional architecture can differentiate between images which have: extended versus point source features, single versus multiple sources, centred versus sources with off-centred components, and a combination of these attributes at varying extents. Noting that the data cubes were created using untreated radio continuum images, and passed through the functional architecture. The produced physical bins were curated from specified orders of pixel intensities, at varying scales and frequencies.

From the original image data cubes, we formed SOM control group bins, by using untreated image inputs with a single FOV and specified survey. We found that surveys optimised for point source detection (VLSSr, TGSS and FIRST) produce single SOM bins. Meaning, the SOM could not distinguish differences within a scene of an image feature column vector. We clarified that attempting to pass pixels directly to a very basic neural network was the cause of this, and its inclusion was for the purpose of establishing control groups. Another caveat was the SOMs inclination towards binning the same source, with minor image changes (say the retrieved image had coordinates that were off by 5 arcseconds) into different classes. Using feature column vectors abstracted (from the original image data cubes) by the pre-trained networks, remedied both of these caveats and produced differentiated bins.

Comparing the abstracted original image data cubes results with the multi-scale data cubes results, highlighted that, the multi-scale abstracted features show some physical binning of the sources. We found that SOMs tend to form a "generalised" major grouping of heterogeneous source features which have statistical similarities, and smaller surrounding bins with more homogeneous source features. These include extended sources with varying spatial extent, noise dominated images, multi-component sources exhibiting similar geometric features to triples, triples, doubles, elongated doubles, and combination of extend-elongated features. From this analyses, we are able to produce physically explainable groupings, using data sets produced by surveys optimised for point source detection. One of the major results came from the effective use of the functional architecture to mine and gain physical insights from VLSSr, a low-frequency legacy radio survey.

The final multi-scale-frequency data cubes added the functionality of abstracting multi-scales features from a collection of specified survey combinations. We found that the multi-scale-frequency binned SOMs classes seem to add a level of abstraction that further differentiates the bins, based on the number of sources in a given image scene. This differentiated the bins into classes with: plumes, long extensions, disconnected "patchy" extensions, symmetric (bright) jets, off-centred sources, and multi-sources. Noting that the presented images are randomly sampled images of each class, and there was no human intervention in curating the observed classes. The importance of (choice when) curating combinations of surveys to abstract, was highlighted. These should be based on physical features (and dynamics) of interest, in an investigation.

We round off this discussion with a proposed approach to finding non-thermal diffuse emission. Using the observed changes between class bins produced by SOMs and tracking known extended radio source populations with non-thermal diffuse emission (such as the papers detailed in the Treatment of Data section); we can find unknown groupings of sources with non-thermal diffuse emission from the three types of used data cubes. This ultimately requires the assistance of computer vision techniques to capture, and focus on, the physical extent(s) of radio source components. The above analyses would produce a plethora of SOM bins with varying radio source populations. Computer vision techniques would assist in forming up generalised prescriptions of the classes, and our claim is that in principle; if non-thermal diffuse emission has generalised observable features and/or is contained within some SOM bins, then it is possible to capture its structure from the (computer vision) generalised prescriptions of those SOM bins.

We have shown the potential of using dimensional reduction techniques as physical abstractors. Grouping these abstractions, we systematically presented a functional architecture for curating continuum radio survey data for radio source component identification. The multi-layered analyses approach shows how we can use insights gained from different data cubes to reach a level of physical explainability using a deep learning technique on radio continuum images. These preliminary results show the potential abstractors have in providing a scalable solutions to curating large, fundamental datasets autonomously, especially noting their effectiveness in working with low-frequency data.

Chapter 7

Discussion and Conclusions

In pursuit of finding radio sources with non-thermal diffuse emission from a combination of publicly available radio continuum surveys, at low-mid radio frequencies. We carried out a heuristic based search for extended radio sources in the LoTSS survey, as described in Chap. 4. This investigation led to the identification of 11 potential extended radio sources, with PGC2285791 being chosen for further studies; due to the observable low-brightness extended features and bubble-like, previously undetected, diffuse radio lobe structures with an angular extent of giant radio galaxies.

PGC2285791, previously poorly understood, was characterised to be a star-forming galaxy at redshift $z = 0.14126 \pm 0.00001$. This radio galaxy has two diffuse bright lobes forming a double structure spanning a projected linear size of ~ 0.86 Mpc. The source is seen to have a similar morphology at 141 MHz and 1400 MHz. However, we notice that in the LoTSS image, the core of the source is brighter, and (observed for the first time) two additional compact sources or knots appear in the northern lobe.

The integrated flux emitted by the core is $S_{\mathrm{core(141MHz)}} = 8.74 \pm 1.11$ mJy and $S_{\mathrm{core(1400MHz)}} = 3.40 \pm 1.07$ mJy. Its radio luminosity is, therefore, $\mathrm{L}_{\mathrm{141MHz}} \sim 4.35 \times 10^{23}$ W Hz^{-1} and $\mathrm{L}_{\mathrm{1400MHz}} \sim 1.69 \times 10^{23}$ W Hz^{-1}. The two knots have an integrated flux of $S_{\mathrm{nk(141MHz)}} = 11.26 \pm 1.51$ mJy for the northern knot, and $S_{\mathrm{sk(141MHz)}} = 20.53 \pm 1.42$ mJy for the southern knot.

Modelling the stellar population of PGC2285791 yielded that it is approximated by an 8.7 Gyr old stellar population with lower metallicity. The total and black hole masses are $\sim 5 \times 10^{10}$ $\mathrm{M}_{\odot}$ and 2.7×10^{8} $\mathrm{M}_{\odot}$, respectively, with a SFR of ~ 2 $\mathrm{M}_{\odot}$ yr^{-1}. If correctly represented, the low metallicity could be reflective of gas outflows and the interplay between the galaxy's ISM and the IGM. On the other hand, star formation may have been triggered by galaxy mergers or interaction, though we did not see any sign of interaction or disturbance.

Spectral index distribution maps were produced by combining the LoTSS image at 141 MHz and NVSS image at 1400 MHz. The spectral index of the extended radio source PGC2285791 varies from $\alpha = -1.1$ to $\alpha = -0.5$. An island of a flat spectral region ($\alpha = -0.6$) and steep spectral regions, associated with the hot spots ($\alpha = -1.0$), are observed. Possible interpretations of the spectral-index variation could be density and magnetic field enhancement due to shock compression by external plasma.

Thus, Chap. 4 looked at how we can use publicly available radio continuum surveys for searches (and characterisation) of non-thermal diffuse emission at low-mid radio frequencies, as guided by observations from the LoTSS survey.

We characterised A1689, a massive optically rich galaxy cluster at redshift $z = 0.187$, to investigate the extension of this method at galaxy cluster scales, in Chap. 5. This characterisation revealed an environment with diffuse emission (of LLS $\sim$720 kpc at 147.5 MHz) and multiple radio galaxies that outshine the diffuse emission. The cluster's radio halo was compared to other well known radio halos and was found to be small to medium sized, when compared to radio halos of the same redshift range.

A spectral analysis of the cluster, using the integrated flux densities of the diffuse emission at 5 frequencies, shows that the diffuse emission in the cluster follows a power law relation. We estimated the spectral index value of the diffuse emission to be α_{diffuse} = -1.34. $\pm$ 0.14. This spectral index value and the flux density value at 1400 MHz corresponds to the radio power of the diffuse emission to be $P_{1400\,\text{MHz}}$= 9.81 $\pm$ 0.77 $\times$ 10^{23} W Hz^{-1}.

Spectral index distribution maps were produced between 147.5 and 1400 MHz, to look at the surrounding environment, showed no enhanced spectral features for the main central source. This is inline with the characterisation of A1689 as a cool-core cluster. The most interesting features (from the computed spectral index maps) were the inverted spectral features associated with regions D and E. We note that this could be due to the difference in sensitivities of the surveys used, and a targeted search (with telescope time) would be ideal to verify the inverted spectral features. Spectral index maps between 1400 and 3000 MHz were also produced, to map out the sources from the different regions, using the higher angular resolution maps. They showed a few regions enclosed by steep spectral indices. We note the same caveats pertaining to the difference in sensitivities of the considered surveys.

With both sources (PGC2285791 and A1689) being identified by eye, and showing results that favour the feasibility of using publicly available radio continuum surveys for searches of non-thermal diffuse emission; we looked at the feasibility of using unsupervised learning to scale the search, in Chap. 6. Using pre-trained networks (as abstractors), an unsupervised model, and applying transfer learning principles to images from continuum radio surveys. We showed how self-organising maps (SOMs; in combination with pretrained networks) can be trained on multi-frequency and multi-scale pixel intensity inputs for searches of diffuse emission.

We used data cubes from three categories (namely: original image, multi-scale, and multi-scale-frequency data cubes) to test the binning capabilities of the functional architecture for extended radio sources, in different contexts. Through an iterative process, using the developed functional architecture, we produced SOM bins for each data cube. The analyses iteratively showed that physical bins can be formed from multi-scale-frequency abstracted features. These SOM bins presented groupings which differentiated between the "generalised" major grouping, of heterogeneous source features that have statistical similarities, and smaller surrounding bins with more homogeneous source features. These include extended sources with varying spatial extent, noise dominated images, multi-component sources exhibiting similar geometric features to triples, triples, doubles, elongated doubles, and combination of extend-elongated features.

We rounded off the Chapter discussion with a proposed approach to finding non-thermal diffuse emission. That is, using the observed changes between class bins produced by SOMs, and tracking known extended radio source populations with non-thermal diffuse emission; we can find unknown groupings of sources with non-thermal diffuse emission from the three types of used data cubes. This ultimately requires the assistance of computer vision techniques to capture, and focus on, the physical extent(s) of radio source components. The Chapter's analyses produces a plethora of SOM bins with varying radio source populations. Computer vision techniques would assist in forming up generalised prescriptions of the classes, and our claim is that in principle; if non-thermal diffuse emission has generalised observable features and/or is contained within some SOM bins, then it is possible to capture its structure from the (computer vision) generalised prescriptions of those SOM bins.

One of the major results came from, the effective use of the functional architecture to mine and gain physical insights from VLSSr, a low-frequency legacy radio survey.

We, therefore, have showcased the use of publicly available data (from deep legacy continuum radio surveys) as feasible repositories for searches and characterisations of radio sources with diffuse emission at multiple frequencies and scales; from a single source radio galaxy (presented in Chap. 4), to a galaxy cluster (presented in Chap. 5), and using an automated approach (presented in Chap. 6).

This contribution is significant since the study of astrophysical objects that produce radio emission is expected to become one of the greatest generators of scientific data with the advent of the Square Kilometre Array (SKA). The presented multi-field-of-view and multi-frequency approach provides an effective, autonomous framework for mining insights from legacy continuum radio surveys. Assisting the community with ways of interfacing humans with machine learning for accelerated scientific discovery.

Bibliography

Abell G.O., 1958, ApJS, 3, 211

Abolfathi B. et al., 2018, ApJS, 235, 42

Ackermann M. et al., 2010, ApJ, 717, L71

Ackermann M. et al., 2014, ApJ, 787, 18

Ahn C.P., Alexandroff R., et al., 2012, 203, 21

Allen J. T. et al., 2011, MNRAS, 410, 860

Ammons S.M., Wong K.C., Zabludoff A.I., et al., 2014, ApJ, 781, 2

Andersson K.E., Madejski G.M., 2004, ApJ, 607, 190

Andrei A. H. et al., 2009, A&A, 505, 385-404

Aniyan A. K., Thorat K., 2017, ApJS, 230, 20

ASTRON Team, online page titled “LOFAR STATIONS: DESCRIPTION AND LAYOUT". (accessed 06 January, 2018) `https://www.astron.nl/radio-observatory/astronomers/users/technical-information/lofar-stations/lofar-stations-description-`

Astropy Collaboration et al., 2018, AJ, 156, 123

Bacchi M., Feretti L., Giovannini G., Govoni F., 2003, A&A , 400, 465

Baldi R. D., Capetti A., Massaro F., 2018, A&A, 609, A1

Baltz E.A., 2004, arXiv preprint astro-ph/0412170

Barber T., Meiksin A., Murphy T., 2007, MNRAS, 377, 787

Baron D., 2019, arXiv e-prints, arXiv:1904.07248, arXiv:1904.07248

Bautz L.P., Morgan W.W., 1970, AJ, 162, L149

Beardsley A. P., Johnston-Hollitt M., Trott C. M., et al., 2019, PASA, 36, e050

Becker R. H., White R. L., Edwards A. L., 1991, ApJS, 75, 1

Becker R. H., White R. L., Helfand D. J., 1995, ApJ, 450, 559

Beckmann V., 2006, lecture notes, Astrophysical Spring School, delivered 14 Apr 2006 https://eud.gsfc.nasa.gov/Volker.Beckmann/school/download/Longair_Radiation3.pdf

Begelman M. C., Blandford R. D., Rees M. J., 1984, Reviews of Modern Physics, 56, 255

Bertone G.,Hooper D., Silk J., 2005, Physics Reports, 405, 279

Best P. N., Kauffmann G., Heckman T. M., Ivezić Ž., 2005, MNRAS, 362, 9

Best P. N., Heckman T. M., 2012, MNRAS, 421, 1569

Blandford R., Meier D., Readhead A., 2019, ARA&A, 57, 467

Blanton M. R. et al., 2017, AJ, 154, 28

Blasi P., Colafrancesco S., 1999, Astropart. Phys., 12, 169

Bonafede A., Brunetti G., Vazza F., et al., 2021, ApJ, 907, 32

Bonnarel F. et al., 2000, A&AS, 143, 33

Bornancini C. G., O'Mill A. L., Gurovich S., Lambas D. G., 2010, MNRAS, 406, 197

Bradt H., 2010, "Astrophysics Processes", Cambridge University Press, First edition, 290

Brau J., online image titled "Galaxies", 3 April 2010. (accessed December 22, 2017)
http://pages.uoregon.edu/jimbrau/BrauImNew/Chap24/6th/24_35Figurea-F.jpg

De Breuck C., Klamer I., Johnston H. et al., 2006, MNRAS, 366, 58

Brienza M. et al., 2016, A&A, 585, A29

Broadhurst T., Benítez, N., et al., 2005a, ApJ, 621, 53

Broadhurst T., Takada M., 2005b, ApJ, 619, L143

Brunetti G., Setti G., Feretti L., Giovannini G., 2001, MNRAS , 320, 365

Brunetti G., Blasi P., Cassano R., Gabici S., 2004, MNRAS , 350, 1174

Brunetti G., Cassano R., Dolag K., Setti G., 2003, A&A , 507, 661

Bullock J. S. et al., 2001, MNRAS, 321, 559

Buote D. A., 1998, ApJ , 553, L15

Capetti A., Massaro F., Baldi R. D., 2017, A&A, 598, A49

Capetti A., Massaro F., Baldi R. D., 2017, A&A, 601, A81

Cantwell, T. M., Scaife, A. M. M., Oozeer N., et al., 2016, MNRAS , 458, 1803

Cavagnolo K. W., Donahue M., Voit G. M., Sun M., 2009, ApJS, 182, 12

Chabrier G., 2003, PASP, 115, 763

Chen Y., Reiprich T. H., Böhringer H., et al. 2007, A&A, 466, 805

Chen Z. F. et al., 2014, ApJS, 210, 7

Chen Z. F., Gu Q. S., Chen Y. M., 2015, ApJS, 221, 32

Clowe D., Bradač M., Gonzalez A., et al., 2006, ApJ , 648, L109

Cohen A. S., Lane W. M., Cotton W. D., et al. 2007, ApJ , 134, 1245

Colafrancesco S., 2010, Mem. S.A.It., 81, 104

Colafrancesco S., 2010, AIP Conference Proceedings, 1206, 5

Colafrancesco S., Melé B., 2001, ApJ, 562, 24

Colafrancesco S., Profumo S., Ullio P., 2006, A&A, 455, 21

Colafrancesco S., 2011, Lieu R., Marchegiani P., Pato M., Pieri L., BuonannoR., A&A,527, A80

Colafrancesco S., Marchegiani P., Beck G., 2015, JCAP, 02, 032

Colafrancesco S. et al., 2016, MNRAS, 456, 512

Colafrancesco S., Marchegiani P., Paulo C. M., 2017, MNRAS, 471, 4747

Comparat J. et al., 2017, ArXiv e-prints, 1711.06575

Condon J. J., 1992, ARA&A , 30, 575

Condon J. J., Cotton W. D., et al., 1998, ApJ , 115, 1693

Cooksey K. L. et al., 2013, ApJ, 763, 37

Croston J. H. et al., 2008, ApJ, 688, 190

Dabhade P., Röttgering H. J. A., Bagchi J., et al., 2020, A&A, 139, 313

Dahle H., Kaiser N., Irgens R. J., et al., 2002, ApJS, 139, 313

DeBoer D. R., Parsons A. R., Aguirre J. E., et al., 2017, PASP, 129, 045001

Dewdney P. E., Hall P. J., Schilizzi R. T., Lazio T. J. L. W., 2009, Proceedings of the IEEE, 97, 1482

Duchesne S. W., Johnston-Hollitt M., Offringa A. R., 2021, PASA, 38, e010

Duric N., Seaquist E. R., 1988, ApJ, 326, 574

Egger J. M., 2006, ArXiv Mathematics e-prints

Enßlin T. A., Gopal-Krishna, 2001, A&A, 366, 26

Fanaroff B. L., Riley J. M., 1974, MNRAS, 167, 31

Fanti C., Fanti R., 1994, Astronomical Society of the Pacific Conference Series, 54, 341

Farnes J. S., Gaensler B. M., Carretti E., 2014, ApJS, 212, 15

Feretti L., Dallacasa D., Govoni F., et al., 1999, A&A , 344, 472

Feretti L., Fusco-Femiano R., Giovannini G., Govoni F., 2001, A&A , 373, 106

Feretti L., 2005, Advances in Space Research, 36, 729

Feretti L., Giovannini G., Govoni F., Murgia M., 2012, A&A Rev., 20, 54

Ferrari C., Govoni F., Schindler S., Bykov A.M., and Rephael, Y., 2008, Clusters of Galaxies, 93

Ficarra A., Grueff G., Tomassetti G., 1985, A&AS, 59, 255

Flesch E., and Hardcastle M. J., 2004, A&A , 427, 387

Funk S., Hinton J.A., 2013, Astroparticle Physics, 43, 348

Gaia Collaboration et al., 2016, A&A, 595, A2

Galvin T. J., Huynh M. T., Norris R. P. et al., 2020, MNRAS, 497, 2730–2758

de Gasperin F., Intema H. T., Shimwell T. W., et al., 2017, Science Advances, 3, 10

Gendre M. A., Best P. N., Wall J. V., 2010, MNRAS, 404, 1719

Gendron-Marsolais M., Hlavacek-Larrondo J., van Weeren R. J., et al., 2020, MNRAS, 499, 5791

Giacintucci S., Markevitch M., Venturi T., et al., 2014, ApJ, 781, 9

Ginzburg V. L., Syrovatskii S. I., 1964, “The Origin of Cosmic Rays", New York: Macmillan, 66

Giovannini G., Feretti L., 2000, A&A ,5, 335

Giovannini G., Feretti L., Bacchi M., Govoni F., 2003, A&A , 301, 483

Giovannini G., Bonafede A., Feretti L., et al., 2009, A&A , 507, 1257

Gitti M., Brunetti G., Setti G., 2002, A&A , 386, 456

Gitti M., Brunetti G., Feretti L., Setti G., 2004, A&A , 417, 1

Gitti M., Ferrari C., Domainko W., Feretti L., Schindler S., 2007, A&A , 470, L25

Gitti M., 2013, MNRAS, 436, L84

Govoni F., Enßlin T. A., Feretti L., Giovannini G., 2001a, A&A , 369, 441

Govoni F., Feretti L., Giovannini G., et al., 2001b, A&A , 376,803

Govoni F., Markevitch M., Vikhlinin A., et al., 2004, ApJ , 605, 695

Govoni F., Feretti L., 2004, International Journal of Modern Physics , 13, 1549

Govoni F., Murgia M., Markevitch M., et al., 2009, A&A , 499, 2

Green D. A., Riley J. M., 1995, MNRAS, 274, 324

Gregory P. C., Condon J. J., 1991, ApJS, 75, 1011

Gregory P. C., Scott W. K., Douglas K., Condon J. J., 1996, ApJS, 103, 427

Griem H. R. (1997) "Continuous spectra," in Principles of Plasma Spectroscopy, Cambridge: Cambridge University Press (Cambridge Monographs on Plasma Physics), 132–155.

Goddard D. et al., 2017a, MNRAS, 465, 688

Goddard D. et al., 2017b, MNRAS, 466, 4731

van Haarlem M. P., Wise M. W., Gunst A. W., 2013, A&A, 556, A2

Hales S. E. G., Baldwin J. E., Warner P. J., 1988, MNRAS, 234, 919

Hao J. et al., 2010, ApJS, 191, 254

Helou G., Madore B. F., Schmitz M., 1995, springer , 96

He K.,. Zhang X., Ren S., Sun J., 2016, ECCV 2016, 630

Hoekstra H., Yee H. K. C., Gladders M. D., 2004, ApJ , 606, 67

Holt J. et al., 2006, MNRAS, 370, 1633

Holt J., Tadhunter C. N., Morganti R., 2008, MNRAS, 387, 639

Hotan A. W., Bunton J. D., Tuthill J., Voronkov M., Wu X., 2020, Publications of the Astronomical Society of Australia, xx, xx

Hu W., 2021, lecture notes, Astro 305, delivered 27 Apr 2021 `http://background.uchicago.edu/~whu/Courses/Ast305_10/ast305_08.pdf`

Hunter J. D., 2007, Computing In Science & Engineering, 9, 90

Hurley-Walker N., Callingham J. R., et al., 2017, MNRAS , 464, 1146

Ignesti A., Brunetti G., Gitti M., Giacintucci S., 2020, A&A , 640, A37

Inductiveload, NASA, online image titled "A diagram of the electromagnetic spectrum", 26 October 2007. (accessed December 22, 2017) `https://en.wikipedia.org/wiki/Electromagnetic_spectrum#/media/File:EM_Spectrum_Properties_edit.svg`

Intema H. T., Van der Tol S., et al., 2009, A&A , 501, 1185

Intema H. T., 2014, Astrophysics Source Code Library, record ascl:1408.006

Intema H. T., Jagannathan P., Mooley K. P., Frail D. A., 2017, A&A , 598, A78

Jaeger S., 2008, Astronomical Data Analysis Software and Systems XVII , 394, 623

Jaffe W. J., 1977, ApJ , 212, 1

Jayalakshmi D. S., Srinivasan R., and Srinivasa K. G., 2016, IGI Global, ch029, 639

Jester S., Röser H. J., Meisenheimer K., et al., 2001, A&A , 373, 447

Johnston S., et al., 2007, Publications of the Astronomical Society of Australia, 24, 174

Jonas J., MeerKAT Team, 2016, in MeerKAT Science: On the Pathway to the SKA, 1

Jin Jonghoon, Dundar A., Culurciello E., 2015, ICLR, abs/1412.5474

Kewley L. J., Groves B., Kauffmann G., Heckman T., 2006, MNRAS, 372, 961

Khatri R., 2016, VizieR Online Data Catalog, 359

Kohonen T., 1997, Proceedings of International Conference on Neural Networks (ICNN'97), 1, PL1-PL6 vol.1

Kohonen T., 2013, Neural Networks, 37, 52

Krawczyk C. M. et al., 2013, ApJS, 206, 4

Kuźmicz A., Jamrozy M., Bronarska, K., et al., 2018, ApJS, 238, 9

Lacy M., Baum S. A., Chandler C. J., et al., 2020, IOP Publishing, 132, 035001

Lane W.M., Cotton W.D., van Velzen S., et al., 2014, ApJ, 440, 327

Leonard A., King L. J., and Goldberg D. M., 2011, MNRAS, 413, 789

LibreTexts, Physics, online image titled "The Electromagnetic Spectrum", 17 December 2016. (accessed December 22, 2017) `https://phys.libretexts.org/TextMaps/Astronomy_and_Cosmology_TextMaps/Map%3A_Astronomy_(OpenStax)/5%3A_Radiation_and_Spectra/5.2%3A_The_Electromagnetic_Spectrum`

Lin Y. T. et al., 2010, ApJ, 723, 1119

Lister M. L., Aller M. F., Aller H. D., et al., 2018, ApJS, 234, 12

Liu C., Zoph B., Neumann M., et al., 2018, ECCV 2018, 19

Macario G., Venturi T., Intema H. T., et al., 2013, A&A, 551, A141

Mahatma V. H., Hardcastle M. J., Williams W. L., et al., 2019, A&A, 622, A13

Malarecki J. M. et al., 2013, MNRAS, 432, 200

Mandal S., Intema H. T., van Weeren R. J., et al., 2020, A&A, 634, A4

Mann A. W., Ebeling H., 2012, MNRAS, 420, 2120

Maraston C., Strömbäck G., 2011, MNRAS, 418, 2785

Markevitch M., Vikhlinin A., 2001, ApJ , 563, 95

Markevitch M., Gonzalez A.H., David L., et al., 2002, ApJ ,567, L27

Markevitch M., 2005, arXiv preprint astro-ph/0511345

Markevitch M., Govoni F., 2005, ApJ, 627, 733

Martino R., Mazzotta, P., Bourdin H., et al., 2014, MNRAS, 443, 2342

Maslowski J., 1972, ACTAA, 22, 227

Maslowski J., Pauliny-Toth I. I. K., Witzel A., Kuehr H., 1984, A&A, 139, 85

Massaro F. et al., 2014, ApJS, 213, 3

McMahon R. G., White R. L., 2002, ApJS, 143, 1

McMullin J. P. et al., 2007, Astronomical Data Analysis Software and Systems XVI, 376, 127

Mhardcastle, online image titled "Radio galaxy 3C31", 5 February 2006. (accessed January 19, 2018)
`https://commons.wikimedia.org/wiki/File:Radio_galaxy_3C31Fr.png#/media/File:Radio_galaxy_3C31.png`

Mhardcastle, online image titled "Radio galaxy 3C98", 5 February 2006. (accessed January 19, 2018)
`https://upload.wikimedia.org/wikipedia/commons/a/aa/Radio_galaxy_3C98.png`

Mingo B. et al., 2011, ApJ, 731, 21

Minniti D., Ramos R., Alonso-García J., 2015, ApJ , 810, L20

Mohan N., Rafferty, D., 2015, Astrophysics Source Code Library, record ascl:1502.007

Mostert R. I. J., Duncan K. J., Röttgering H. J. A. et al., 2021, A&A 645, A89 (2021)

Mulcahy D. D. et al., 2014, A&A, 568, A74

Nandi S., Saikia D. J., 2012, BASI, 40, 121

O'Dea C. P. et al., 2002, AJ, 123, 2333

Oguri M. & Blandford R.D., 2009, MNRAS , 392, 930

Ogrean G. A., Brüggen M., van Weeren R. J., Röttgering H., Croston J. H., Hoeft M., 2013, MNRAS, 433, 812

Padovani P., 2016, A&A Rev., 24, 13

Padovani P., Alexander D.M., Assef R.J. et al., 2017, A&A Rev., 25, 2

Pan S. J., Kwok J. T., Yang Q., 2008, AAAI 08, 677

Panadzhyan V. G., 1998, Astrophysics, 41, 246

Pâris I. et al., 2014, A&A, 563, A54

Paturel G. et al., 2003, A&A, 412, 45

Pedersen K., Dahle H., 2007, AJ , 667, 26

Pieri L. et al., 2011, Phys. Rev. D., 83, 023518

Pollack J. B., Fazio G. G., 1965, ApJ, 141, 730

Polsterer K. L., Gieseke F., Igel C., 2015, Astronomical Data Analysis Software an Systems XXIV (ADASS XXIV), 495, 81

Pooley G. G., Kenderdine S., 1968, MNRAS, 139, 529

Price-Whelan, A. M., Sipőcz B. M., Günther H. M., et al., 2018, AJ, 156, 123

Prochaska J. X., Hennawi J. F., Herbert-Fort S., 2008, ApJ, 675, 1002

Proctor D. D., 2016, ApJS, 224, 18

Quillen A. C., Bland-Hawthorn J., Brookes M. H., et al., 2006, ApJL, 641, L29

Ransom S., online page titled “Brightness and Flux Density", 06 December 2010. (accessed January 25, 2018) `http://www.cv.nrao.edu/course/astr534/Brightness.html`

Ransom S., online page titled “Radiation Fundamentals", 23 September 2016. (accessed December 22, 2017) `http://www.cv.nrao.edu/~sransom/web/Ch2.html`

Ransom S., online page titled “Free–Free Radiation", 23 September 2016. (accessed December 22, 2017) `http://www.cv.nrao.edu/~sransom/web/Ch4.html`

Ransom S., online page titled “Synchrotron Radiation", 25 October 2016. (accessed December 22, 2017) `http://www.cv.nrao.edu/~sransom/web/Ch5.html`

Ransom S., online page titled “Spectral Lines", 29 November 2016. (accessed December 22, 2017) `https://www.cv.nrao.edu/~sransom/web/Ch7.html`

Reichert A., Böhringer H., Fassbender R., Mühlegger M., 2011, A&A, 535, A4

Rengelink R. B., Tang Y., de Bruyn A. G., et al., 1997, A&AS, 124, 259

Rephaeli Y., 1979, ApJ, 227, 364

Rines K., Geller M. J., Diaferio A., Kurtz M. J., 2013, ApJ , 767, 15

Sánchez-Blázquez P. et al., 2006, MNRAS, 371, 703

Sand D. J., Treu T., Ellis R. S., Smith G. P., 2005, ApJ , 627, 32

Sault R. J., Teuben P. J., Wright M. C. H., 1995, ADASS IV, ASP Conf. Series, 77, 433

Schneider D. P. et al., 2007, AJ, 134, 102

Schneider D. P. et al., 2010, AJ, 139, 2360

Schoenmakers Arno P. et al., 2000, MNRAS, 315, 371

Sereno M. & Umetsu K., 2011, MNRAS , 416, 3187

Sereno M., Ettori S., Umetsu K., & Baldi A., 2013, MNRAS , 428, 2241

Seyffert E. N. et al., 2013, ApJ, 779, 161

Shafi N. et al., 2015, MNRAS, 454, 1404

Shan H. Y., Qin B., Fort B., et al., 2010, MNRAS , 406, 1134

Shen Y. et al., 2008, ApJ, 680, 169

Shen Y. et al., 2011, ApJS, 194, 45

Shimwell T. W., Markevitch M., et al. 2015, MNRAS , 409, 1486

Shimwell T. W. et al., 2017, A&A, 598, A104

Shulevski A. et al., 2017, A&A, 600, A65

Struble M. F., Rood, H. J., 1999, ApJS, 125, 35

Subrahmanyan R., Saripalli L., Hunstead R. W., 1996, MNRAS, 279, 257

Suchkov A. A., Hanisch R. J., Voges W., Heckman T. M., 2006, AJ, 132, 1475

Swinburne University of Technology, online pdf titled "Mechanisms of Radio Emission: Thermal and Non-Thermal Radio Emission", 2010. (accessed December 22, 2017)
`http://astronomy.swinburne.edu.au/sao/downloads/HET608-M03A02.pdf`

Szabo T. et al., 2011, ApJ, 736, 21

Szegedy C., Vanhoucke V., Ioffe S., Shlens J., Wojna Z., 2016, CVPR 2016, 2818

Szegedy C., Ioffe S., Vanhoucke V., Alemi A. A., 2017, AAAI 17, 4278

Taylor A. N., Dye S., 1998, ApJ , 501, 539

Thompson A. R., Moran, J. M., Swenson G. W. , 2017, Interferometry and Synthesis in Radio Astronomy 3rd edition, Springer International Publishing

Thompson T. A., Quataert E., Murray N., 2009, MNRAS , 397, 1410

Thwala S. A., Shafi N., Colafrancesco S., Govoni F., Murgia M., 2019, MNRAS, 485, 1938

Tuccillo D., Huertas-Company M., Decencière E. et al., 2018, MNRAS, 475, 894

Tyson J. A., Valdes F., Wenk R. A., 1990, ApJ , 349, L1

Tyson J. A., Fischer P., 1995, arXiv preprint astro-ph/9503119

Vacca V., Govoni F., et al., 2011, A&A , 535, A82

Vallee J. P., Roger R. S., 1989, A&AS, 77, 31

van Haarlem M. P. et al., 2013, A&A, 556, 2

van Weeren R. J., Röttgering H. J. A., Brüggen M., Hoeft M., 2010, Science, 330, 347

van Weeren R. J., Brüggen M., Röttgering H. J. A., Hoeft M., 2011, MNRAS, 418, 230

van Weeren R. J., de Gasperin F., Akamatsu H. et al., 2019, SSRv, 215, 16

Vardoulaki E., Jiménez Andrade E. F., Karim A. et al., 2019, A&A, 627, A142

Veilleux S., Cecil G., Bland-Hawthorn J., 2005, ARA&A , 43, 769

Véron-Cetty M.-P., Véron P., 2006, A&A, 455, 773

Véron-Cetty M.-P., Véron P., 2010, A&A, 518, A10

Vollmer B. et al., 2010, A&A, 511, A53

Wardzinski G., Zdziarski A. A., 2000, MNRAS, 314, 183

Wayth R. B., Lenc E., et al., 2015, PASA, 32

Weltman A., Bull P., Camera S., et al., 2020, PASA, 37, e002

West A. A. et al., 2008, AJ, 135, 785

West A. A. et al., 2011, AJ, 141, 97

Wilkinson D. M. et al., 2015, MNRAS, 449, 328

Wilkinson D. M. et al., 2017, MNRAS, 472, 4297

Wright E. L. et al., 2010, AJ, 140, 1868

Zhou W., 2002, Chinese Journal of Astronomy and Astrophysics, 2, 106