Վ. Համբարձումյանի անվան Բյուրականի աստղադիտարան

Գոռ Արեգի Միքայելյան

ԵՆԹԱԿԱՐՄԻՐ ԳԱԼԱԿՏԻԿԱՆԵՐԻ ՈՐՈՆՈՒՄ ԵՎ ՈՒՍՈՒՄՆԱՍԻՐՈՒԹՅՈՒՆՆԵՐ

Ատենախոսություն

Ա.03.02 – «Աստղաֆիզիկա, ռադիոաստղագիտություն» մասնագիտությամբ

Ֆիզիկամաթեմատիկական գիտությունների թեկնածուի գիտական աստիճանի համար

Գիտական ղեկավար՝

Ֆ.մ.գ.թ. Արեգ Մարտինի Միքայելյան

Բյուրական - 2026

V. AMBARTSUMIAN BYURAKAN ASTROPHYSICAL OBSERVATORY

Gor A. Mikayelyan

SEARCH AND STUDIES OF INFRARED GALAXIES

DOCTORAL THESIS

For the degree of Doctor of Philosophy in the

Specialization A.03.02 – Astrophysics, Radio Astronomy

Supervisor:

Dr. Areg M. Mickaelian

Byurakan – 2026

TABLE OF CONTENTS

INTRODUCTION 7

Aim of the Thesis 10

Objectives of the Thesis 11

Scientific Novelty 11

Practical and Scientific Significance 12

Structure of the Thesis 12

The Main Points to Be Presented for Defense 13

Personal Contribution to Co-authored Publications 14

Publications Forming the Basis of the Thesis 15

Scientific Presentations of the Thesis Results 16

1. INFRARED GALAXIES AND THEIR ASTROPHYSICAL CONTEXT

1.1. Infrared Galaxies 17

1.2. The IRAS Survey 23

1.3. The IRAS PSC/FSC Combined Catalogue 25

1.4. Infrared Luminosity Classification 25

1.5. Star Formation and Starburst Activity 26

1.6. Active Galactic Nuclei 26

1.7. Starburst-AGN Connection 27

1.8. Interacting, Binary, and Multiple Infrared Galaxies 28

1.9. Optical Spectroscopy and Activity Classification 28

1.10. Multiwavelength Data and Astronomical Archives 29

1.11. Scientific Motivation for the Present Study 30

2. BYURAKAN-IRAS GALAXIES: OPTICAL IDENTIFICATION AND SPECTROSCOPIC STUDIES

2.1. Introduction 31

2.2. The Byurakan-IRAS Galaxy Sample 32

2.3. Optical Identification and the Role of Low-Dispersion Spectroscopy 33

2.4. Spectroscopic Observations and Derived Parameters 34

2.5. Spectral Activity Classification 42

2.6. Activity Distribution of the BIG Sample 42

2.7. Star-Forming and HII Galaxies 44

2.8. AGN, Seyfert Galaxies, and LINERs 44

2.9. Composite and Uncertain Spectral Classes 45
2.10. Redshift, Distance, and Absolute-Magnitude Distributions 46
2.11. Pairs and Multiple Systems 48
2.12. Methodological Limitations and Interpretation 49
2.13. Discussion 50
2.14. Connection with the Following Chapters 51
2.15. Summary 51
3. THE EXTRAGALACTIC SAMPLE OF THE IRAS PSC/FSC COMBINED CATALOGUE
3.1. Introduction 53
3.2. The IRAS PSC/FSC Combined Catalogue 53
3.3. Methodological Design of the Extragalactic-Sample Construction 54
3.4. Choice of Search Radius and Positional Uncertainty 56
3.5. Identification of Previously Known Galaxies and Active Objects 56
3.6. Galaxy-Candidate Selection from Optical Morphology 58
3.7. Radio Information as an Independent Extragal. Criterion and Gaia-based Stellar Rejection 59
3.8. Preliminary Classification of the IRAS Population 60
3.9. Connection to the ULIRG and HLIRG Sample 61
3.10. Limitations and Sources of Uncertainty 62
3.11. Methodological Perspective 62
3.12. Discussion 63
3.13. Summary 64
4. HIGH-LUMINOSITY INFRARED GALAXIES: ULIRGs AND HLIRGs
4.1. Introduction 65
4.2. Physical and Scientific Motivation 65
4.3. Parent Catalogue and Positional Refinement 66
4.4. Optical Counterpart and Spectroscopic Selection 67
4.5. Four-Band IRAS Photometry Requirement 67
4.6. Infrared and Far-Infrared Luminosities 68
4.7. Construction of the High-Luminosity Sample 68
4.8. Composition of the Sample and Newly Identified Objects 71
4.9. Luminosity-Redshift Distribution and Selection Effects 71
4.10. Spectroscopic Classification: General Method 72
4.11. Classification Criteria and Sample Spectra 73

4.12. Detailed Spectral Activity Types 77
4.13. Mean Properties by Activity Type 78
4.14. Starburst and AGN Activity 80
4.15. Comparison with Existing Activity Information 80
4.16. Comparison with the BIG Population 81
4.17. Relation to the IRAS Extragalactic Sample 81
4.18. Relation to the ArVO and BAO Archive Framework 82
4.19. Summary 82
5. ARMENIAN VIRTUAL OBSERVATORY AND BAO PLATE ARCHIVE: DATA INFRASTRUCTURE AND METHODOLOGICAL APPLICATIONS
5.1. Introduction: ArVO, the BAO Plate Archive and DFBS 84
5.2. The DFBS as a Digital Astronomical Resource 85
5.3. Automated Classification of Low-Dispersion Spectra 87
5.4. Preprocessing and Spectral-Image Extraction 88
5.5. Supervised and Unsupervised Learning 89
5.6. Experimental Results and Limitations 91
5.7. Cloud-Based Machine Learning Service for Astronomical Sub-Object Classification 93
5.8. Revised Extraction, CNN Architecture and Transfer Learning 94
5.9. Cloud-Based Service and Workflow 96
5.10. Evaluation, Large-Scale Inference and Limitations 97
5.11. Comparative Discussion and Relevance to Infrared-Galaxy Research 100
5.12. Summary 102
6. OVERALL DISCUSSION AND SCIENTIFIC SYNTHESIS
6.1. Scope of the Research Programme 103
6.2. Infrared Selection as the Starting Point 103
6.3. BIG as the Detailed Source-Level Component 104
6.4. Spectroscopic Classification as a Physical Diagnostic 105
6.5. From the BIG Population to ULIRGs and HLIRGs 106
6.6. Infrared Luminosity and Its Observational Interpretation 107
6.7. Star Formation, AGN, and the Activity Sequence 108
6.8. Interactions and the Nature of the Infrared Source 109
6.9. The Large Extragalactic Sample and the Problem of Completeness 109
6.10. Role of ArVO and the BAO Archive in the Overall Methodology 110

6.11. Unified Methodological Framework 111
6.12. Methodological Strengths and Limitations 112
6.13. Scientific Implications and Future Research 114
6.14. Chapter Summary 115
CONCLUSION 116
BIBLIOGRAPHY 120
LIST OF ACRONYMS 139
ACKNOWLEDGEMENTS 141

INTRODUCTION

Infrared astronomy provides an important observational window for the study of galaxies and their evolution. A significant fraction of the energy emitted by some galaxies is released in the infrared, making infrared observations particularly effective for revealing systems with intense star-formation activity, large amounts of interstellar dust, and active galactic nuclei. Infrared-selected galaxies therefore provide valuable information on the processes that regulate star formation, nuclear activity, galaxy interactions, and the evolution of galaxies across cosmic time. The study of such objects is also important for understanding the relationship between starburst activity and active galactic nuclei and the role of interactions and mergers in the evolution of galaxies [1, 2].

The InfraRed Astronomical Satellite (IRAS) was the first space-based observatory to carry out an infrared survey of almost the entire sky. IRAS mapped approximately 96% of the sky in four bands centred at 12, 25, 60, and 100 μm and detected approximately 350,000 infrared sources. The IRAS Point Source Catalogue (PSC) [3] and Faint Source Catalogue (FSC) [4] subsequently became major resources for infrared and extragalactic studies. In particular, the 60 and 100 μm measurements provide far-infrared information that is highly useful for investigations of infrared-selected galaxies.

Among infrared-selected galaxies, the most luminous systems are of particular interest. Following the commonly used infrared-luminosity classification [5], galaxies with $L_{IR} > 10^{11}\ L_{\odot}$ are classified as Luminous Infrared Galaxies (LIRGs), galaxies with $L_{IR} > 10^{12}\ L_{\odot}$ as Ultra-Luminous Infrared Galaxies (ULIRGs), and galaxies with $L_{IR} > 10^{13}\ L_{\odot}$ as Hyper-Luminous Infrared Galaxies (HLIRGs). Their high luminosities allow them to be detected at relatively large distances, making them useful objects for studying energetic processes in galaxies and the history of star formation.

The high infrared luminosity of a galaxy may be produced by intense star formation, nuclear activity, or a combination of both. Starburst activity is associated with enhanced star-formation rates over relatively short evolutionary timescales, while active galactic nuclei are powered by energetic processes in compact nuclear regions and are generally associated with accretion onto supermassive black holes. A major unresolved question concerns the physical relationship between these two forms of activity and the circumstances under which one may accompany or

influence the other. The question becomes especially relevant for luminous infrared galaxies because many such systems also exhibit interaction, multiple structure, or merging features.

Reliable identification of infrared sources is therefore a fundamental step in their scientific study. The positional uncertainties of the original IRAS catalogues, together with the possibility of uncertain detections, make direct identification of their optical counterparts difficult. To improve the available data set, the IRAS PSC and FSC were combined into the IRAS PSC/FSC Combined Catalogue [6]. This catalogue contains 345,163 infrared sources. Its positional information was improved by cross-correlation with more recent catalogues, including AKARI-IRC [7], AKARI-FIS [8], and WISE [9], providing a substantially more accurate basis for identifying optical counterparts and carrying out multiwavelength studies.

Optical identification and spectroscopy of IRAS sources have been a major research direction at the Byurakan Astrophysical Observatory since 1995 [10]. These studies resulted in the discovery and investigation of the Byurakan-IRAS Galaxies (BIG) [11], including galaxies with strong star-formation activity, active nuclei, enhanced infrared luminosity, and interacting or multiple structures. Spectroscopic observations have been obtained with the BAO 2.6-m telescope [12, 13], the 6-m telescope of the Special Astrophysical Observatory (SAO, Russia) [14, 15], the 1.93-m telescope of the Observatoire de Haute-Provence (OHP, France) [16], and the Sloan Digital Sky Survey (SDSS) [17, 18]. A total of 257 BIG objects have been spectroscopically studied and classified [19].

The spectroscopic study of BIG objects provides redshifts, radial velocities, distances, absolute magnitudes, and information about the physical activity of the galaxies. Among the 257 studied objects, 149 were classified as HII/star-forming galaxies, while the remaining objects were distributed among AGN, composite, emission-line, normal, LINER, Seyfert, absorption, and unclassified categories. The sample also contains pairs and multiple systems, providing an opportunity to study the relationship between infrared emission, star formation, nuclear activity, and galaxy interactions [19].

The next stage of the research involves extending the study from a relatively limited set of identified BIG objects to the much larger population represented by the IRAS PSC/FSC Combined Catalogue [6]. Because the catalogue contains hundreds of thousands of sources, reliable

determination of their nature requires the combination of several complementary data sets. Cross-correlation with catalogues of known galaxies, quasars, and blazars can provide secure identifications, while optical morphological information can be used to identify galaxy candidates. Radio catalogues provide an additional extragalactic criterion, and Gaia proper-motion information can be used to identify and remove stellar sources [20].

The resulting large-scale investigation has produced a preliminary classification of the IRAS source population into stars, galaxies, quasars, galaxy candidates, and unidentified sources. From the 345,163 sources of the Combined Catalogue, 183,168 were identified as star candidates, 68,940 as galaxy candidates, and 12,057 as confirmed or genuine galaxies or quasars, while 80,998 sources remained unidentified at the stage reported in the study [20].

A further step is the identification of the most luminous infrared galaxies within this large data set. Using improved WISE positions, SDSS spectroscopic information, and the four IRAS flux measurements, 6,346 sources with reliable spectroscopic optical counterparts were selected. Of these, 6,270 had measurements available in all four IRAS bands required for the infrared-luminosity calculation [21]. Selecting objects with $L_{IR} > 10^{12} L_{\odot}$ resulted in a sample of 114 high-luminosity infrared galaxies belonging to the ULIRG and HLIRG classes [22].

The 114-object sample was subsequently studied spectroscopically in order to determine the types of nuclear and star-forming activity present in these extremely luminous systems [23]. The classification identified quasars and BL Lac objects, Seyfert galaxies with broad and narrow spectral components, LINERs, composite objects, HII galaxies, emission-line objects, and absorption-line systems. The analysis showed that the high-luminosity infrared sample contains a substantial population of broad-line active galaxies in addition to the HII and lower-ionization systems commonly associated with less luminous infrared galaxies [24].

The study of infrared galaxies can be extended further through the use of the Armenian Virtual Observatory (ArVO) and archival observations preserved at the Byurakan Astrophysical Observatory. The BAO Plate Archive contains approximately 37,000 photographic plates and films obtained during decades of observations. The digitization programme was developed to preserve this observational heritage, provide access to the data, and enable new scientific investigations [25]. The Digitized First Byurakan Survey (DFBS), based on 1,874 plates, contains

approximately 40 million extracted low-dispersion spectra and represents an additional valuable source of information for the investigation and classification of 20 million astronomical objects [26, 27, 28, 29].

The present Thesis brings these lines together into a systematic study of infrared galaxies selected from IRAS data. The work combines optical identification, spectroscopy, catalogue cross-correlation, infrared and far-infrared luminosity determination, spectral activity classification, and the study of interacting, pairs, and multiple systems. Particular emphasis is placed on the Byurakan-IRAS Galaxy population and on the high-luminosity ULIRG and HLIRG population, while the larger IRAS extragalactic sample provides the basis for extending the investigation to a much broader population of infrared galaxies.

Aim of the Thesis

The main aim of this Thesis is to investigate the optical, spectroscopic, infrared, and multiwavelength properties of IRAS-selected galaxies, to improve the identification and classification of their extragalactic counterparts, to characterise their infrared luminosities and types of activity, and to investigate the role of interactions and multiple systems in the observed properties of infrared galaxies.

Objectives of the Thesis

- To investigate the optical counterparts and spectroscopic properties of Byurakan-IRAS Galaxies.
- To determine redshifts, distances, absolute magnitudes, and spectral activity types for the studied BIG objects.
- To develop and extend the extragalactic sample of the IRAS PSC/FSC Combined Catalogue using cross-correlation with optical, infrared, radio, and Gaia catalogues.
- To identify high-luminosity infrared galaxies within the IRAS sample and determine their infrared luminosities.
- To construct and investigate the sample of ULIRGs and HLIRGs and identify newly revealed objects.

- To classify the high-luminosity infrared galaxies according to their spectral activity and investigate the distribution of HII, Seyfert, LINER, composite, quasar, and other activity types.
- To investigate pairs, multiple, interacting, and merging infrared systems and determine the relation between their structure and infrared emission.
- To explore the use of the Armenian Virtual Observatory and the BAO Plate Archive as resources for the further investigation of infrared and active galaxies.

Scientific Novelty

The scientific results presented in this Thesis include the systematic spectroscopic analysis of a large sample of Byurakan-IRAS Galaxies, the construction and analysis of an extensive extragalactic sample from the IRAS PSC/FSC Combined Catalogue, and the identification and classification of 114 high-luminosity infrared galaxies belonging to the ULIRG and HLIRG classes. Among these, 48 ULIRGs and 7 HLIRGs were reported as newly identified objects [19, 20, 22, 23].

The Thesis also brings together the classification of the high-luminosity objects with their infrared luminosities, redshifts, optical colours, and spectral properties, while extending the analysis toward interacting and multiple systems and the broader use of astronomical archival resources.

Practical and Scientific Significance

The results provide observational samples that can be used for statistical studies of infrared galaxies, active galaxies, star-forming systems, and high-luminosity infrared objects. The identification and classification of the sources also provide a basis for selecting individual objects and subsamples for more detailed optical, infrared, radio, ultraviolet, X-ray, and morphological investigations. The development of large and homogeneous samples is particularly important for studying the statistical properties of infrared galaxies and their evolution.

Structure of the Thesis

Chapter 1 presents the scientific background required for the study, including infrared galaxies, the IRAS survey and catalogues, LIRGs, ULIRGs and HLIRGs, star-formation and nuclear activity, galaxy interactions, spectroscopic classification, and the main astronomical data resources used in the research.

Chapter 2 is devoted to the Byurakan-IRAS Galaxies. It presents the optical and spectroscopic observations, the determination of redshifts and distances, the classification of the galaxies according to their activity, and the statistical properties of the BIG sample, including pairs and multiple systems.

Chapter 3 presents the construction of the extragalactic sample of the IRAS PSC/FSC Combined Catalogue. The cross-correlation methodology and the use of optical, radio, infrared, and Gaia catalogues are described, followed by the preliminary classification of the IRAS population.

Chapter 4 focuses on high-luminosity infrared galaxies. The selection of ULIRGs and HLIRGs, the calculation of their infrared luminosities, their spectral classification, and the statistical relations between activity type, luminosity, redshift, and optical colours are investigated.

Chapter 5 introduces the Armenian Virtual Observatory (ArVO), the BAO Plate Archive and the Digitized First Byurakan Survey (DFBS), then focuses on two studies of automated DFBS classification. The first investigates supervised and unsupervised classification of low-dispersion spectral images; the second extends the approach to selected spectral subtypes and presents a cloud-based machine-learning service.

Chapter 6 synthesizes the studies of Chapters 2–5, considers their scientific implications and methodological limitations, and identifies directions for future research.

The Conclusion summarizes the principal results of the Thesis.

The Main Points to Be Presented for Defense

1. **Results of the optical identification and spectroscopic study of Byurakan-IRAS Galaxies.** A spectroscopic sample comprising 257 objects was compiled and investigated using observations obtained with the 2.6-m telescope of the Byurakan Astrophysical Observatory

(BAO), the 6-m telescope of the Special Astrophysical Observatory (SAO), the 1.93-m telescope of the Observatoire de Haute-Provence (OHP), and data from the Sloan Digital Sky Survey (SDSS). Redshifts, distances, absolute magnitudes, and spectral activity classes were determined for the studied objects.

2. **Methodology and results of constructing an extragalactic sample from the IRAS PSC/FSC Combined Catalogue.** An approach for the preliminary identification of the extragalactic nature of infrared sources was developed and applied. It is based on cross-matching with catalogues of known galaxies and active objects, optical morphological information, radio-source data, and astrometric information from Gaia.

3. **Selection of high-luminosity infrared galaxies and identification of new objects.** A sample of 114 galaxies belonging to the ultraluminous infrared galaxy (ULIRG) and hyperluminous infrared galaxy (HLIRG) classes was selected from the IRAS PSC/FSC Combined Catalogue. Among these objects, 48 ULIRGs and 7 HLIRGs were reported as newly classified.

4. **Results of the spectroscopic classification of high-luminosity infrared galaxies.** The spectral properties and activity types of the selected objects were investigated, distinguishing Seyfert galaxies, quasars, LINERs, HII galaxies, composite objects, and other spectral classes. The roles of active galactic nuclei and star-formation activity in the infrared emission of high-luminosity systems were examined.

5. **Results of the investigation of pairs and multiple systems of infrared galaxies.** The optical identification of infrared sources and the characteristics of interacting galaxies and multiple systems were investigated. Possible relationships between their structure, star-formation activity, and nuclear activity were considered.

6. **Applications of astronomical archival data and digital methods.** The potential of the Armenian Virtual Observatory (ArVO), the BAO Plate Archive, and the Digitized First Byurakan Survey (DFBS) was investigated for the identification of astronomical objects and the classification of low-dispersion spectra.

Personal Contribution to Co-authored Publications

I made a substantial personal contribution to the research presented in this Thesis.

In **Chapter 2**, I compiled and analysed the previously obtained observational and spectroscopic results for the Byurakan-IRAS Galaxies (BIG). Although the original observations were carried out by other researchers, I collected and systematised the available results, reviewed and corrected the parameters, prepared the summary of the spectroscopic dataset, and conducted a dedicated investigation of pairs and multiple systems. I also analysed the spectroscopic classifications and derived parameters to provide a consistent overview of the studied sample.

In **Chapter 3**, I carried out the principal research activities related to the construction and analysis of the extragalactic sample based on the IRAS PSC/FSC Combined Catalogue, including the processing and interpretation of catalogue data and the investigation of source-identification criteria. The cross-correlations with all catalogues have been done by me.

In **Chapter 4**, I conducted the principal analysis of the high-luminosity infrared galaxy sample, including the selection and investigation of ULIRGs and HLIRGs, the analysis of infrared luminosities, and the study of their spectroscopic classifications and activity types.

For the research presented in **Chapter 5**, my contribution consisted of data acquisition, data analysis and/or interpretation.

I also contributed to the overall synthesis and interpretation of the research results presented in this Thesis and to the preparation of the scientific material.

Publications Forming the Basis of the Thesis

1. G.A. Mikayelyan, A.M. Mickaelian, H.V. Abrahamyan, G.M. Paronyan, “Discovery of New Bright ULIRGs from the IRAS PSC/FSC Combined Catalogue”, Communications of the Byurakan Astrophysical Observatory, vol. 65, pp. 13–20, 2018.
2. G.A. Mikayelyan, A.M. Mickaelian, H.V. Abrahamyan, G.M. Paronyan, M.V. Gyulzadyan, “Spectral Classification of Ultrabright IR Galaxies from the IRAS PSC/FSC Combined Catalog”, Astrophysics, vol. 62, pp. 452–466, 2019.
3. G.A. Mikayelyan, A.M. Mickaelian, H.V. Abrahamyan, G.M. Paronyan, M.V. Gyulzadyan, “Observations and Study of Byurakan-IRAS Galaxies: Summary”, Communications of the Byurakan Astrophysical Observatory, vol. 66, pp. 31–40, 2019.

4. H. Astsatryan, G. Gevorgyan, A. Knyazyan, A. Mickaelian, G. Mikayelyan, “Astronomical Objects Classification Based on the Digitized First Byurakan Survey Low-Dispersion Spectra”, Astronomy and Computing, vol. 34, id 100442, 2021.
5. G.A. Mikayelyan, A.M. Mickaelian, H.V. Abrahamyan, G.M. Paronyan, “Building the Largest Sample of IR Galaxies: Preliminary Results”, Communications of the Byurakan Astrophysical Observatory, vol. 70, pp. 321–327, 2023.
6. H. Astsatryan, S. Babayan, A. Mickaelian, G. Mikayelyan, M. Astsatryan, “Cloud-Based Machine Learning Service for Astronomical Sub-Object Classification: Case Study on the First Byurakan Survey Spectra”, Data Science Journal, vol. 23 (1), id 6, 14 pp., 2024.
7. G.A. Mikayelyan, “Recent Results of Studies of IR Galaxies”, Communications of the Byurakan Astrophysical Observatory, vol. 72, pp. 409–415, 2025.

Scientific Presentations of the Thesis Results

Seminars

- Leiden University, Netherlands, 18.06.2025 - *Multiwavelength Studies of Infrared Galaxies*
- Byurakan Astrophysical Observatory, Armenia, 15.10.2026 – *Search and Studies of Infrared Galaxies*

Oral Talks

- European Astronomical Society Annual Meeting 2020, Virtual, 01-03.06.2020 - *Digitized First Byurakan Survey low-dispersion spectroscopic archive for future studies*
- International Conference “Non-Stable Phenomena in the Universe”, Byurakan, Armenia, 18-21.09.2023 - *The extragalactic sample of the Combined IRAS PSC/FSC catalogue*
- Symposium “Relation of Astronomy to Other Sciences, Culture and Society”, Byurakan, Armenia, 6-8.05.2024 – *Pairs and Multiples in Infrared Galaxies*
- European Astronomical Society Annual Meeting 2025, Cork, Ireland, 23-27.06.2025 - *The largest IR galaxy sample for Starburst and AGN study*
- International Symposium “Astronomical Surveys and Big Data 3”, Byurakan, Armenia, 15-19.09.2025 - *Recent Results of Studies of IR Galaxies*

1. INFRARED GALAXIES AND THEIR ASTROPHYSICAL CONTEXT

1.1. Infrared Galaxies

Infrared galaxies are an important class of extragalactic objects whose energy output is strongly concentrated in the infrared part of the electromagnetic spectrum. Their infrared emission is commonly associated with dust-reprocessed radiation from regions of intense star formation, although an active galactic nucleus may also make a significant contribution to the observed luminosity. Infrared observations therefore provide an efficient way to identify galaxies in which large amounts of energy are being released by star formation and nuclear activity.

The observational foundation for identifying such a population was established with the development of space-based infrared astronomy and, in particular, with the launch of the Infrared Astronomical Satellite (IRAS). IRAS demonstrated the importance of infrared observations for extragalactic astronomy and revealed a large population of galaxies that emit a major part of their energy at infrared wavelengths. Many of these galaxies are rich in gas and dust, and their infrared emission may be produced by vigorous star formation, AGN activity, or both.

These properties make infrared-selected galaxies relevant to several connected scientific problems: the evolution of galaxies, the history of star formation, the growth of supermassive black holes, the interaction between starburst and nuclear activity, and the role of galaxy interactions and mergers in triggering energetic phenomena.

1.1.1. Physical Basis of Infrared Astronomy

To understand why infrared observations are so effective for detecting luminous infrared galaxies, particularly dusty systems with intense star formation and active galactic nuclei, it is necessary to consider how infrared radiation is produced and emitted. Infrared astronomy occupies a distinctive position within modern multiwavelength astrophysics because thermal and dust-reprocessed emission carries information that is often strongly attenuated or completely hidden at optical wavelengths. The infrared regime is not limited to a particular class of astronomical objects: all bodies with temperatures above absolute zero emit thermal radiation, while the

relative importance of that emission shifts toward longer wavelengths as the characteristic temperature decreases. Consequently, infrared observations provide access to populations and physical environments that may be inconspicuous in the optical domain.

For astronomical applications, the infrared range is commonly divided into near-infrared (NIR), mid-infrared (MIR), and far-infrared (FIR) domains. The boundaries are not unique and may depend on detector technology, atmospheric transmission, and the observational context. Table 1.1 below emphasises this practical definition and associates the different wavelength domains with characteristic classes of objects: relatively cool stars and stellar envelopes in the NIR; warm dust, planets, comets, asteroids, and circumstellar or protoplanetary material in the MIR; and cold gas, cool dust, molecular clouds, and the central regions of galaxies in the FIR [1].

Table 1.1. Practical infrared spectral domains and representative astronomical applications.

**The temperature ranges are presented as characteristic scales associated with the wavelength ranges in the source material; the boundaries between NIR, MIR, and FIR are not universal and depend on the observational system. [1]*

Infrared domain	Approximate wavelength	Characteristic temperature*	Representative objects
Near-infrared (NIR)	(0.7-1)–5 μm	~740–(3000-5200) K	Cool/red stars, stellar envelopes, planetary nebulae
Mid-infrared (MIR)	5–(25-40) μm	~(92.5-140)–740 K	Planets, comets, asteroids, warm dust, circumstellar and protoplanetary material
Far-infrared (FIR)	(25-40)–(200-350) μm	~(10.6-18.5)–(92.5-140) K	Cold gas and dust, molecular clouds, galactic central regions

The physical reason for this wavelength dependence can be expressed through Wien's displacement relation: the wavelength at which a thermal spectrum reaches its maximum increases as the temperature decreases. Thus, cold astrophysical material emits preferentially at longer wavelengths. In the context of galaxy research, this makes the infrared particularly sensitive to interstellar dust and to energy absorbed from ultraviolet and optical radiation and

subsequently re-emitted thermally. Regions of vigorous star formation can therefore become prominent infrared sources even when they are heavily obscured in visible light.

This temperature dependence also explains why the infrared view of a galaxy can differ markedly from its optical appearance. In the near-infrared, the effects of dust obscuration are reduced relative to the visible domain, allowing structures hidden by extinction to be investigated. At longer wavelengths, emission from cooler material becomes increasingly important. Figures 1.1 and 1.2 illustrate this change using images of dusty Galactic regions and the Galactic centre, emphasizing that objects may appear, disappear, or change their relative prominence as the observing wavelength is increased.



Figure 1.1. Appearance of the Horsehead Nebula in the visible, near-IR (2MASS), and mid-IR (ISO) ranges [1].

These physical effects are directly relevant to infrared galaxies. Strong far-infrared emission may trace large reservoirs of cool dust heated by young stars, while mid-infrared emission can contain contributions from warmer dust and compact nuclear regions. In active galaxies, radiation associated with an energetic nucleus can also heat surrounding dust and contribute substantially to the infrared output. The observed infrared luminosity is therefore an integrated indicator of several possible energy-production mechanisms, and its interpretation is most reliable when combined with optical spectroscopy and other wavelengths.

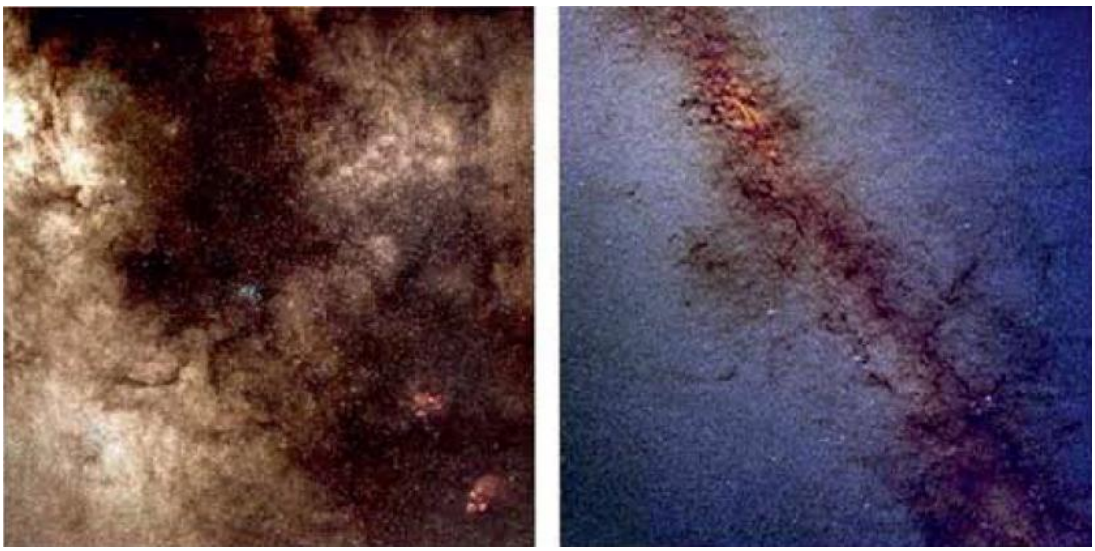


Figure 1.2. View of the Galactic Centre in the visible and near-infrared (2MASS) ranges [1].

1.1.2. Atmospheric Constraints and the Need for Space-Based Infrared Observations

The usefulness of particular infrared wavelength ranges on Earth is strongly constrained by the terrestrial atmosphere. Water vapour absorbs over substantial portions of the infrared spectrum, while the atmosphere itself produces thermal emission. The combination of atmospheric absorption and background radiation limits sensitivity, especially at longer wavelengths. Ground-based infrared observatories are therefore preferentially located at high, dry sites, where the water-vapour column is reduced. Even at such sites, only selected atmospheric transmission windows are available for observations. Figure 1.3 shows the atmospheric transmission across the electromagnetic spectrum.

The detector environment introduces an additional constraint. Infrared instruments can themselves emit thermal radiation and therefore introduce background signals. Cooling the telescope and its detectors to very low temperatures reduce this contribution and improve sensitivity.

Together, the atmospheric and instrumental limitations explain the central role of space missions in the development of infrared astronomy. Space-based telescopes avoid the dominant effects of atmospheric absorption and can observe wavelength ranges that are inaccessible, or extremely difficult, from the ground. They also provide more stable backgrounds and enable systematic surveys over large portions of the sky. These advantages were essential for the transition from specialized ground-based infrared observations to large homogeneous infrared catalogues,

beginning with IRAS and continuing with subsequent missions such as ISO [30], Spitzer [31], AKARI [7, 8], and WISE [9].

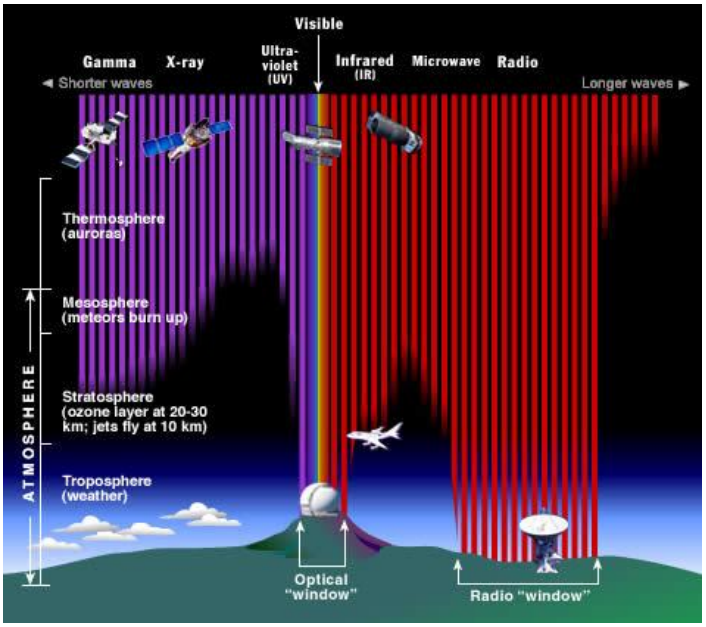


Figure 1.3. Atmospheric transmission across the electromagnetic spectrum, illustrating the wavelength regions accessible to ground-based astronomy (Credit: STScI/JHU/NASA).

1.1.3. Development of Infrared Astronomy and the IRAS Breakthrough

Following the physical and observational foundations described above, infrared astronomy developed from early ground-based measurements toward increasingly sensitive surveys and dedicated space observatories. The evolution of infrared astronomy included the identification of infrared emission from both Galactic and extragalactic objects, and the subsequent extension of observations to progressively longer wavelengths. These developments established infrared astronomy as an important branch of observational astrophysics capable of detecting objects and structures that can be weak or obscured at optical wavelengths. Table 1.2 describes the major infrared telescopes and surveys.

The increasing importance of space-based observations followed directly from the atmospheric and thermal limitations discussed in Section 1.1.2. Space missions made it possible to obtain systematic measurements over wavelength ranges that are difficult or inaccessible from the ground and to build large, homogeneous infrared surveys. This progression culminated in the

Infrared Astronomical Satellite (IRAS), whose mission, survey characteristics, principal catalogues, and importance for extragalactic astronomy are discussed in detail in Section 1.2.

Table 1.2. Major infrared telescopes and large-scale infrared surveys, with their operating periods, wavelength coverage, and principal results [1].

Telescope / Survey	Country	Period	Wavelength bands, μm	Main results
IRAS [3,4]	USA	1983	8–120	All-sky survey
ISO [30]	Europe	1996-1998	2.5–240	Infrared spectra
Spitzer [31]	USA	2003-2020	3–180	Deep infrared images and spectra
AKARI [7,8]	Japan	2006-2007	1.8–180	All-sky survey
Herschel [32]	Europe	2009-2013	55–672	Far-infrared observations
WISE [9]	USA	2010-2011	3–28	All-sky survey
JWST [33]	USA	2022-	0.6–28.5	Deep infrared images
DENIS [34]	Europe	1996-2001	0.82; 1.24; 2.16	Southern-sky survey (ground based)
2MASS [35]	USA	1997-2001	1.24; 1.66; 2.16	All-sky survey (ground based)

1.1.4. Infrared Galaxies, Star Formation, AGN, Interactions and Merging

With the observational importance of infrared galaxies established, the next question is what physical processes produce their strong infrared emission. Two major sources of energetic activity are intense star formation and active galactic nuclei. In star-forming systems, young massive stars emit strong ultraviolet radiation that is absorbed by surrounding dust and re-radiated thermally in the infrared. In AGN, compact nuclear activity can heat surrounding dusty material and produce an additional infrared component. The relative contribution of these mechanisms varies from object to object, making multiwavelength classification an essential part of infrared-galaxy research.

At the population level, this physical diversity is reflected in the luminosity classes LIRG, ULIRG, and HLIRG, which provide a useful phenomenological ordering by total infrared luminosity. The commonly used thresholds are 10^{11}–10^{12} $L_\odot$ for LIRGs, 10^{12}–10^{13} $L_\odot$ for ULIRGs, and above 10^{13}

$L_{\odot}$ for HLIRGs [5]. The increase in infrared luminosity is accompanied by a changing mixture of physical activity, with an increasing importance of AGN among the most luminous systems [24].

The physical interpretation of the most luminous systems is further complicated by the frequent presence of interacting and multiple structures. Highly luminous infrared galaxies often display interacting or multiple morphologies [36], and the coexistence of enhanced infrared emission, starburst activity, and nuclear activity raises the question of how these processes are physically related. This question is especially relevant to the ULIRG population and is carried into the later chapters of this Thesis, where both the high-luminosity sample and BIG pairs and multiples are considered.

1.2. The IRAS Survey

A decisive step in the development of infrared astronomy was the launch of the Infrared Astronomical Satellite (IRAS) on 25 January 1983. IRAS operated for approximately ten months and surveyed about 96% of the sky repeatedly in four bands centred at 12, 25, 60, and 100 μm. The mission detected roughly 350,000 infrared sources and produced the first major homogeneous all-sky infrared source catalogues [3,4].



Figure 1.4. The Infrared Astronomical Satellite (IRAS) (Credit: IRAS/IPAC).

The two principal catalogues produced from the IRAS survey are the IRAS Point Source Catalogue (PSC) [3] and the IRAS Faint Source Catalogue (FSC) [4]. Together they provide measurements at four infrared wavelengths and became fundamental historical data sets for infrared and extragalactic studies. The PSC provides an all-sky source list, while the FSC provides a deeper sample and excludes the Galactic region with $|b| < 10°$.

For extragalactic research, an important strength of IRAS is its homogeneous far-infrared coverage, especially at 60 and 100 μm. The survey revealed a large population of infrared-emitting galaxies, including systems whose optical counterparts were not always securely identified. Consequently, the scientific use of IRAS galaxy samples depends not only on the infrared measurements themselves, but also on accurate source identification and on combining the IRAS data with more precise modern catalogues. This naturally leads to the construction of the IRAS PSC/FSC Combined Catalogue [6] discussed in Section 1.3.

1.3. The IRAS PSC/FSC Combined Catalogue

The separate PSC and FSC catalogues contain overlapping sources, but their direct cross-identification is complicated by the positional uncertainties and the varying reliability of individual detections. To address this problem, a dedicated cross-correlation procedure was developed within the framework of the Armenian Virtual Observatory [37]. Instead of adopting one fixed search radius, the method used a radius corresponding to the positional uncertainty of each individual source. This resulted in 73,770 associations between PSC and FSC sources.

The resulting IRAS PSC/FSC Combined Catalogue contains 345,163 sources. Its positional information was further improved through cross-correlations with AKARI-IRC, AKARI-FIS, and 2MASS/WISE. The combined data set contains photometric information extending from approximately 1.25 to 160 μm and provides a homogeneous basis for large-scale studies of infrared sources [6].

The improved positional accuracy is particularly important for optical identification. The original IRAS positional uncertainties can make the identification of an optical counterpart ambiguous, whereas WISE and other modern catalogues provide substantially improved astrometric

information. This improvement forms one of the methodological foundations of the research presented in the subsequent chapters.

1.4. Infrared Luminosity Classification

Infrared galaxies cover a wide range of luminosities. A commonly used classification separates them into LIRGs, ULIRGs, and HLIRGs according to their total infrared luminosity [5]. LIRGs have $10^{11} < L_{IR}/L_{\odot} < 10^{12}$, ULIRGs have $10^{12} < L_{IR}/L_{\odot} < 10^{13}$, and HLIRGs have $L_{IR}/L_{\odot} > 10^{13}$.

The increase of infrared luminosity is accompanied by important changes in the observed galaxy population. The available studies indicate an increasing importance of AGN among the most luminous infrared systems, while starburst-dominated systems are common among less luminous infrared galaxies [24]. This makes luminosity an important parameter when investigating the activity mechanisms of infrared-selected galaxies.

ULIRGs are particularly important because their infrared luminosities can be comparable to those of quasars, while their optical appearance may remain strongly influenced by dust and star formation. Their high luminosities also make them accessible at larger distances and therefore useful for studies of the evolution of luminous galaxy populations.

1.5. Star Formation and Starburst Activity

Star formation is one of the principal mechanisms responsible for strong infrared emission in galaxies. Newly formed massive stars emit large amounts of ultraviolet and optical radiation, which can be absorbed by surrounding dust and subsequently re-emitted in the infrared. Thus, strong infrared emission can serve as an indicator of elevated star-formation activity.

The starburst phase corresponds to a period in which the star-formation rate is significantly enhanced over the normal level of a galaxy. The studies summarized in the present work note that typical star-formation rates may be of order a few solar masses per year, while in powerful starburst systems they can become orders of magnitude larger.

For the galaxies considered in this Thesis, star formation is mainly investigated through optical spectral classification and infrared luminosity measurements. HII-type spectra provide evidence

for ionized regions associated with star formation, while the infrared luminosity provides an integrated measure of the energetic output from the dust-emitting component.

1.6. Active Galactic Nuclei

Active Galactic Nuclei (AGN) represent another major source of energy in infrared galaxies. The central region is associated with accretion onto a supermassive black hole, producing radiation over a broad wavelength range. Depending on the physical conditions and orientation of the nuclear region, galaxies can exhibit different spectral signatures, including Seyfert, LINER, and quasar characteristics.

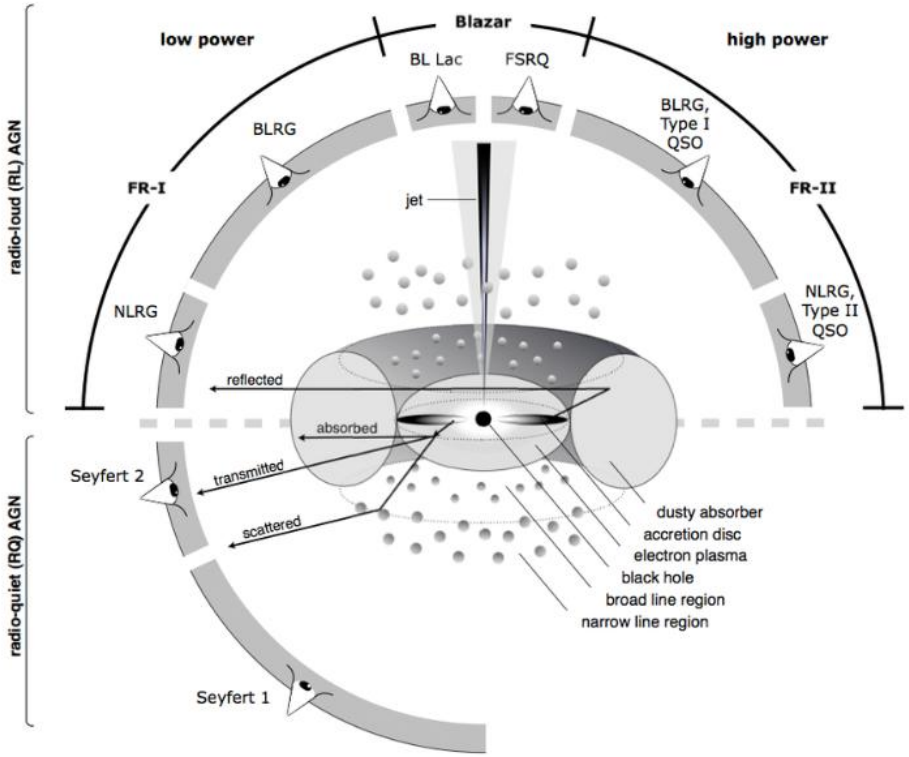


Figure 1.5. The unified model of Active Galactic Nuclei (AGN) [38].

Optical spectroscopy is particularly useful for distinguishing these activity types. Emission-line ratios and the presence or absence of broad components provide information about the excitation mechanism and the physical conditions in the nuclear region. In the high-luminosity infrared sample investigated in this Thesis, the presence of broad and narrow spectral components is particularly important for distinguishing the various Seyfert subclasses.

1.7. Starburst–AGN Connection

One of the central scientific questions associated with infrared galaxies is the relationship between star formation and AGN activity. Infrared emission can result from both mechanisms, while interactions and mergers may influence either or both processes. Consequently, the observation of a high infrared luminosity alone does not uniquely determine the physical source of the emission.

The samples investigated in this Thesis provide an observational framework for addressing this issue. The Byurakan-IRAS Galaxies (BIG) sample contains HII, AGN, LINER, Seyfert, and composite objects, while the ULIRG/HLIRG sample contains active galaxies together with HII and composite systems. Their comparison as a function of infrared luminosity, redshift, and optical colour provides an empirical basis for studying the diversity of energetic processes in infrared galaxies.

1.8. Interacting, Binary, and Multiple Infrared Galaxies

A notable property of infrared galaxies is the frequent occurrence of interacting, binary, and multiple systems. The interaction between galaxies can redistribute gas, compress the interstellar medium, and provide conditions favorable to enhanced star formation and nuclear activity. The identification of such systems is therefore an important part of the study of infrared galaxies.

The BIG observations contain both isolated objects and systems with two or more optical counterparts associated with an infrared source. Their investigation makes it possible to determine whether the infrared emission originates from one galaxy or from the interaction as a whole and to study the relationship between galaxy structure, star formation, and nuclear activity.

The preliminary investigation reported in the recent studies identified 361 groups containing 907 counterparts of binary and multiple systems [2]. For a subset of these groups, spectroscopic observations are available for two or more members, allowing a more direct investigation of their physical relationship.

1.9. Optical Spectroscopy and Activity Classification

Optical spectroscopy provides the principal method used in this Thesis for determining the physical nature and activity of the selected infrared galaxies. Spectroscopic observations provide redshifts and therefore distances, while the emission and absorption lines provide information about the ionization and excitation processes occurring in the galaxy.

The BIG studies employ spectra obtained from several observatories and from SDSS, allowing objects observed with different instruments and at different epochs to be combined into one statistical sample. For the 257 BIG objects, the available spectra were used to determine redshifts, radial velocities, distances, absolute magnitudes, and activity classes.

For the ULIRG/HLIRG sample, the classification required additional attention to the widths of the emission lines. Many of the spectra contain broad components that cannot be treated in the same way as narrow forbidden and permitted lines. Visual spectral inspection was therefore used as an important part of the classification process, particularly for Seyfert subclasses 1.0–1.9.

1.10. Multiwavelength Data and Astronomical Archives

The study of infrared galaxies benefits strongly from combining data obtained at different wavelengths. IRAS provides far-infrared measurements, while WISE and AKARI provide additional infrared information, SDSS provides optical imaging and spectroscopy, radio catalogues provide complementary information on extragalactic sources, and Gaia provides astrometric and proper-motion information useful for separating stars from extragalactic objects.

The Armenian Virtual Observatory provides an important framework for combining these data sets. It has been used for catalogue cross-correlation and has also been associated with the development of tools for the study of astronomical databases. The BAO Plate Archive further extends the available data into the historical domain. Its digitization provides a searchable digital resource containing tens of thousands of archival observations and approximately 40 million low-dispersion spectra from the Digitized First Byurakan Survey.

1.11. Scientific Motivation for the Present Study

The combination of large infrared catalogues, accurate optical positions, spectroscopic surveys, and archival astronomical data provides an opportunity to move beyond the study of individual infrared galaxies toward statistically meaningful samples. The research presented in this Thesis follows this progression from the identification and spectroscopy of individual BIG objects to the construction of a much larger extragalactic IRAS sample and finally to the selection and detailed classification of the most luminous infrared systems.

This approach allows several related questions to be addressed within one observational framework: What are the optical counterparts of IRAS sources? What types of activity are present in infrared galaxies? How are activity type and infrared luminosity related? How common are interacting and multiple systems? And how can improved catalogue cross-correlation and archival resources increase the number and quality of infrared galaxies available for scientific study?

These questions provide the scientific basis for the original investigations presented in Chapters 2–5.

2. BYURAKAN-IRAS GALAXIES: OPTICAL IDENTIFICATION AND SPECTROSCOPIC STUDIES

2.1. Introduction

The optical identification of infrared-selected sources is a necessary step in converting an infrared catalogue into a scientifically interpretable galaxy sample. Infrared observations are efficient at finding systems with strong dust-reprocessed emission, but the infrared measurement alone does not uniquely establish whether the dominant physical process is star formation, an active galactic nucleus (AGN), or a combination of processes. Optical imaging and spectroscopy provide complementary information that allows the nature of a source to be examined in greater detail, its redshift to be measured, and its nuclear or star-forming activity to be classified [19].

There have been a number of studies on identifications of IRAS galaxies since the release of IRAS catalogues: IRAS Revised Bright Galaxy Sample [39]; Far-InfraRed (FIR) sources [40]; IRAS galaxies towards the Boötes void [41]; IRAS point sources in the area of Fornax, Hydra I and Coma clusters [42]; IRAS 1.2 Jy survey [43]; IRAS galaxies in Virgo cluster area [44]; and some others.

The Byurakan-IRAS Galaxy (BIG) programme was developed at the Byurakan Astrophysical Observatory (BAO) in this observational context. Since the mid-1990s, optical identifications of IRAS sources have been carried out with the aim of revealing galaxies with enhanced star formation, active nuclei, high infrared luminosity, and interacting or multiple structures [10]. The programme made use of the First Byurakan Survey (FBS) [45] and its digitized form, the Digitized First Byurakan Survey (DFBS) [26, 27], as well as subsequent optical observations and spectroscopy [10–17].

The BIG study is closely connected with the later parts of this Thesis. The objects provide a relatively detailed spectroscopic sample against which the much larger IRAS PSC/FSC extragalactic population can be compared. At the same time, the classification of BIG galaxies establishes an observational reference for the ULIRG and HLIRG population considered in Chapter 4. The BAO Plate Archive and Virtual Observatory methodology discussed in Chapter 5

also has a direct historical connection with this programme because low-dispersion FBS/DFBS spectra were used to improve the selection of optical counterparts of IRAS sources [19].

The present chapter concentrates on the observational and analytical methodology of the BIG sample. Particular attention is given to the construction of the spectroscopic dataset, the combination of observations obtained with different facilities, the derived distances and absolute magnitudes, the activity-classification scheme, and the interpretation of binary and multiple systems. The numerical results are retained according to the revised 257-object catalogue [19].

2.2. The Byurakan-IRAS Galaxy Sample

The BIG programme was designed to investigate the optical counterparts of infrared sources and, through spectroscopy, to determine the physical nature of the corresponding galaxies. Earlier identification studies had shown that the optical identification of IRAS sources could be uncertain because the angular positions of infrared detections are less precise than those of modern optical surveys. Low-dispersion spectral information provided an additional discriminator because the spectral appearance of an optical source can distinguish galaxies from stars and can reveal emission features associated with active or star-forming systems [19].

The revised BIG dataset used here contains 257 spectra. The spectroscopic material was obtained from four sources: observations with the BAO 2.6-m telescope [12,13], observations with the 6-m telescope of the Special Astrophysical Observatory (SAO) [14,15], observations with the 1.93-m telescope of the Observatoire de Haute-Provence (OHP) [16], and spectra obtained from the Sloan Digital Sky Survey (SDSS) [17]. The source counts are 56, 54, 64, and 83, respectively, giving 257 spectra in total [19]. In cases where multiple spectra of the same object were available from different telescopes or observations, the spectrum with the best overall quality was selected for the final dataset.

The combination of these four observational sources is one of the strengths of the sample, but it also means that the dataset is not a single-instrument survey. The spectra were obtained under different observational conditions and with different instruments. Consequently, the sample should be interpreted as a compiled spectroscopic dataset rather than as a strictly homogeneous

flux-limited spectroscopic survey. Its primary strength is the accumulated spectroscopic coverage of identified IRAS galaxies and the resulting diversity of objects available for comparative study.

Table 2.1. Spectroscopic sources contributing to the revised 257-object BIG sample.

Spectroscopic source	Number of spectra	Fraction of total
BAO 2.6-m	56	21.8%
SAO 6-m	54	21.0%
OHP 1.93-m	64	24.9%
SDSS	83	32.3%

2.3. Optical Identification and the Role of Low-Dispersion Spectroscopy

The identification problem begins with an infrared position and requires the determination of the most plausible optical counterpart. In the BIG programme, optical identification was supported by the information contained in the FBS and DFBS. The original FBS was an objective-prism survey, and its low-dispersion spectra contain information about colour, continuum shape, emission features, and absorption features. These characteristics can be used as a first-level discriminator when examining optical counterparts of infrared sources [45].

The DFBS work is particularly relevant because the digitization transformed the original photographic material into a coordinate-associated digital resource. The DFBS contains approximately 20 million astronomical objects and about 40 million extracted low-dispersion spectra. Its original survey area covers about 17,000 square degrees. The availability of extracted spectra made it possible to inspect candidate counterparts computationally as well as through direct visual examination [26].

For infrared-source identification, the value of low-dispersion spectroscopy is not that it replaces higher-resolution spectroscopy, but that it provides an efficient method for determining whether an optical source has the broad spectral characteristics expected for a galaxy, an active object, or a stellar contaminant. This was already part of the historical identification methodology and later became a natural application of the digitized archive.

In this sense, the BIG programme represents an intermediate methodological stage between individual historical identifications and the systematic catalogue cross-correlation approach used in Chapter 3. The BIG work relies on careful identification of particular sources, whereas the later IRAS extragalactic sample applies a hierarchy of automated or catalogue-based associations to hundreds of thousands of sources.

2.4. Spectroscopic Observations and Derived Parameters

Each entry in the revised BIG table contains an IRAS identifier, an optical apparent magnitude, the source of the spectrum, the emission-line redshift, the corresponding radial velocity, the adopted distance, the absolute magnitude, and the activity type. For SDSS objects the *r*-band magnitude is used, while magnitudes from the historical observations are described as being close to *V*.

The emission redshift is the primary spectroscopic distance indicator in the tabulated sample. The source study gives radial velocities derived from the emission redshift using the relativistic relation and distances calculated with the adopted Hubble parameter $H_0 = 71$ km s^{-1} Mpc^{-1} [19]. This parameter choice is part of the original reduction and is retained here rather than replaced by a different cosmological convention.

The corresponding absolute magnitude is obtained from the apparent optical magnitude and distance according to the conventions used in the source dataset. Because the sample combines several types of spectroscopy and optical measurements, the absolute magnitudes should be interpreted as catalogue-level optical luminosity indicators rather than as a uniform precision photometric measurement for every object.

Table 2.2. Principal parameters recorded for the 257 BIG objects.

Parameter	Description	Role in the analysis
IRAS name	Identifier of the infrared source	Links the optical spectrum to the infrared catalogue
$m_{V/r}$	Optical apparent magnitude; r used for SDSS objects	Describes optical brightness and supports absolute-magnitude estimates

z_{em}	Emission-line redshift	Determines recession velocity and distance
v_r	Radial velocity derived from redshift	Provides the velocity scale used in the source table
R	Distance in Mpc, adopted H_0 = 71 km s^{-1} Mpc^{-1}	Used to characterise the spatial scale of the sample
M	Absolute magnitude	Characterises optical luminosity
Activity type	Spectral classification	Separates star-forming and nuclear-activity populations

Table 2.3. The list of all observed BIG objects with their parameters.

IRAS name	mv	Spectra source	z_{em}	V	R	Mv	Activity Type
				km/s	*Mpc*		
03304+8456a	16.8	OHP 1.93m	0.0509	14868	209.40	-19.80	HII
03304+8456b	16.4	BAO 2.6m	0.0500	14625	205.99	-20.16	Abs
03333+7851	15.0	OHP 1.93m	0.0561	16347	230.24	-21.81	HII
03347+7748	15.5	OHP 1.93m	0.0346	10209	143.79	-20.29	Composite
03386+7909	15.9	OHP 1.93m	0.0559	16313	229.76	-20.91	Sy2
03424+8424	16.0	SAO 6m	0.0744	21492	302.70	-21.41	LINER:
03424+8713	15.4	BAO 2.6m	0.0249	7365	103.74	-19.67	HII
03485+7703	15.6	OHP 1.93m	0.0701	20281	285.64	-21.68	HII
04033+6942	15.9	OHP 1.93m	0.0161	4803	67.65	-18.25	LINER
04079+7033	14.5	OHP 1.93m	0.0132	3934	55.41	-19.22	HII
04140+7448	16.0	OHP 1.93m	0.0338	9972	140.45	-19.74	AGN
04183+7457	14.7	OHP 1.93m	0.0325	9589	135.05	-20.95	Composite
04574+7639	17.1	BAO 2.6m	0.1209	34081	480.02	-21.29	Em
05097+7954	15.2	OHP 1.93m	0.0570	16602	233.83	-21.64	HII
05126+6516	15.4	OHP 1.93m	0.0402	11824	166.53	-20.71	HII
05196+7257	17.4	OHP 1.93m	0.1030	29303	412.72	-20.68	Sy2
05214+7741	16.3	OHP 1.93m	0.0769	22194	312.59	-21.17	AGN
05229+6826	14.5	OHP 1.93m	0.0166	4950	69.72	-19.72	HII
05275+6600	16.0	OHP 1.93m	0.0306	9037	127.28	-19.52	AGN:
05395+7550	15.0	OHP 1.93m	0.0248	7342	103.41	-20.07	HII
05401+6435	15.2	SAO 6m	0.0539	15735	221.62	-21.53	HII
05475+7449	16.0	OHP 1.93m	0.0515	15061	212.13	-20.63	Composite
05577+6141	16.1	OHP 1.93m	0.0306	9034	127.24	-19.42	Em
06022+7559	16.8	OHP 1.93m	0.0817	23509	331.11	-20.80	HII
06028+6734a	15.9	BAO 2.6m	0.0169	5015	70.64	-18.33	HII
06028+6734b	15.9	BAO 2.6m	0.0168	4983	70.18	-18.31	HII
06038+6239	14.4	SAO 6m	0.0413	12134	170.91	-21.76	Norm.
06261+7818a	18.4	BAO 2.6m	0.0939	26855	378.25	-19.45	HII

06261+7818b	17.4	BAO 2.6m	0.0940	26883	378.63	-20.48	LINER:
06273+6858a	17.2	BAO 2.6m	0.0664	19252	271.15	-19.96	Em
06273+6858b	16.4	OHP 1.93m	0.0665	19299	271.82	-20.77	HII
06273+6858c	15.7	OHP 1.93m	0.0656	19022	267.91	-21.44	Sy2
06319+7536	18.1	OHP 1.93m	0.0915	26210	369.15	-19.74	Composite
06432+8551	19.0	BAO 2.6m	0.0898	25743	362.58	-18.80	HII
06432+8551b	16.5	SAO 6m	0.0891	25544	359.77	-21.28	HII
06545+6647	16.1	OHP 1.93m	0.0160	4765	67.11	-18.03	HII
06584+6716a	17.0	BAO 2.6m	0.0734	21214	298.79	-20.38	Composite
06584+6716b	17.9	BAO 2.6m	0.0710	20546	289.38	-19.41	Unknown
07007+8242	16.9	SAO 6m	0.0586	17066	240.36	-20.00	HII
07021+7349	14.0	OHP 1.93m	0.1031	29335	413.17	-24.08	LINER:
07158+7706	15.5	OHP 1.93m	0.0485	14198	199.97	-21.00	HII
07205+7842	16.3	OHP 1.93m	0.0866	24859	350.13	-21.42	Sy2
07225+7653	16.1	OHP 1.93m	0.0484	14160	199.44	-20.40	AGN
07479+7832	18.3	OHP 1.93m	0.1734	47577	670.10	-20.83	Em
08036+7211	15.4	BAO 2.6m	0.0387	11374	160.20	-20.63	Em
08054+6824	14.1	OHP 1.93m	0.0413	12123	170.74	-22.06	Composite
08095+6445	14.1	SAO 6m	0.0294	8690	122.40	-21.34	Norm.
08142+6821	16.0	OHP 1.93m	0.0388	11414	160.77	-20.03	HII
08247+7311	14.9	OHP 1.93m	0.0655	19013	267.79	-22.24	HII
08259+7427	17.3	OHP 1.93m	0.1232	34685	488.53	-21.14	HII
08303+6118	16.8	SDSS	0.0865	24831	349.73	-20.89	HII
08317+7602	17.0	OHP 1.93m	0.0931	26632	375.10	-20.87	HII
08339+6517	14.4	BAO 2.6m	0.0190	5634	79.35	-20.12	HII
08379+6753	16.0	BAO 2.6m	0.0363	10687	150.52	-19.89	HII
08410+6124	16.3	SDSS	0.0769	22189	312.52	-21.18	HII
08567+6325	15.7	SDSS	0.0386	11347	159.81	-20.31	HII/Sy2
09002+8106	16.6	SAO 6m	0.0494	14454	203.58	-19.94	HII
09020+6751a	16.7	BAO 2.6m	0.0505	14754	207.80	-19.89	LINER
09020+6751b	16.1	BAO 2.6m	0.0564	16435	231.48	-20.74	Em
09037+6937	16.8	OHP 1.93m	0.0125	3738	52.65	-16.81	HII
09056+6538	16.3	BAO 2.6m	0.0652	18910	266.33	-20.88	Composite
09056+6538	16.3	SDSS	0.0649	18829	265.20	-20.80	HII
09062+8134a	16.8	SAO 6m	0.0495	14483	203.98	-19.75	LINER:
09062+8134b	17.9	SAO 6m	0.0491	14369	202.38	-18.63	HII
09103+8326	17.1	SAO 6m	0.0497	14540	204.79	-19.46	HII
09162+6539	15.5	OHP 1.93m	0.0379	11160	157.19	-20.48	HII
09162+6539	14.6	SDSS	0.0379	11143	156.94	-21.35	HII
09173+6231	15.6	SDSS	0.0473	13859	195.20	-20.90	HII
09176+6544	15.2	SDSS	0.0190	5640	79.44	-19.26	HII
09180+6532	16.0	OHP 1.93m	0.0382	11230	158.17	-20.00	HII
09229+7731	15.0	BAO 2.6m	0.0092	2741	38.61	-17.93	HII
09233+7825	14.7	OHP 1.93m	0.0747	21586	304.03	-22.71	Em
09244+7405	17.2	OHP 1.93m	0.1586	43847	617.57	-21.75	Sy2
09273+6232	16.3	SDSS	0.0420	12350	173.94	-19.89	AGN:

09305+6813	15.7	OHP 1.93m	0.0710	20548	289.41	-21.61	Sy2:
09343+6450	15.8	SDSS	0.0710	20542	289.33	-21.53	Sy2:
09363+6655	15.9	SDSS	0.0596	17355	244.43	-21.04	HII
09386+6240	15.6	SDSS	0.0397	11661	164.24	-20.48	Em
09406+6840	16.3	OHP 1.93m	0.0119	3561	50.15	-17.20	HII
09406+6840	15.6	SDSS	0.0118	3506	49.38	-17.89	HII
09418+8124	16.0	SAO 6m	0.0398	11703	164.82	-20.09	Abs
09427+7528	17.6	BAO 2.6m	0.1099	31168	438.99	-20.61	HII
09477+7050a	18.5	SAO 6m	0.1268	35645	502.05	-20.00	Norm.
09477+7050b	18.0	SAO 6m	0.1270	35698	502.79	-20.51	HII
09477+7050c	17.5	SAO 6m	0.1277	35881	505.37	-21.02	HII
09571+8435	16.5	SAO 6m	0.0924	26444	372.46	-21.36	HII
10045+7502b	16.0	BAO 2.6m	0.0481	14083	198.36	-20.49	LINER
10172+7548	16.2	OHP 1.93m	0.0592	17235	242.75	-20.73	HII
10210+7528	15.5	OHP 1.93m	0.0279	8245	116.12	-19.82	HII
10228+6227	17.2	SDSS	0.1163	32871	462.97	-21.17	HII
10252+7013	16.0	OHP 1.93m	0.1180	33319	469.28	-22.36	HII
10270+7302	12.5	OHP 1.93m	0.0224	6654	93.71	-22.36	HII
10272+6953	14.5	OHP 1.93m	0.0387	11391	160.44	-21.53	Composite
10276+7443	17.3	OHP 1.93m	0.0576	16769	236.18	-19.57	AGN
10298+6119	15.2	SDSS	0.0293	8666	122.06	-20.22	HII
10331+6338	15.9	SDSS	0.0381	11202	157.77	-20.04	HII
10361+7952a	15.9	BAO 2.6m	0.0388	11414	160.77	-20.16	HII
10361+7952b	15.5	BAO 2.6m	0.0394	11587	163.20	-20.56	HII/Comp:
10383+7637	16.2	OHP 1.93m	0.0315	9293	130.88	-19.38	AGN
10486+6558	16.7	SDSS	0.0338	9960	140.28	-18.99	HII
10527+7136	16.4	BAO 2.6m	0.1011	28790	405.50	-21.68	Composite
10529+7144	14.0	OHP 1.93m	0.0619	17996	253.47	-23.02	HII
10541+6614	17.3	SDSS	0.1306	36641	516.08	-21.22	Em
10589+6515	16.1	SDSS	0.0775	22362	314.95	-21.40	Sy2
11008+7915a	16.7	OHP 1.93m	0.0620	18013	253.71	-20.32	HII
11008+7915b	16.5	OHP 1.93m	0.0620	18013	253.71	-20.52	HII
11053+7037	14.0	OHP 1.93m	0.0412	12097	170.38	-22.16	HII
11059+7117	17.9	OHP 1.93m	0.1388	38761	545.94	-20.79	Composite
11067+7024a	15.5	BAO 2.6m	0.0389	11443	161.17	-20.54	HII/Comp:
11067+7024b	16.0	BAO 2.6m	0.0389	11443	161.17	-20.04	HII
11069+7438	16.2	OHP 1.93m	0.0395	11602	163.40	-19.87	HII
11085+7712a	15.6	OHP 1.93m	0.1003	28574	402.46	-22.42	Composite
11085+7712b	18.0	OHP 1.93m	0.0996	28404	400.06	-20.01	AGN
11161+6629	15.5	SDSS	0.0431	12663	178.36	-20.78	HII
11201+6305	16.9	SDSS	0.1905	51799	729.56	-22.41	Em
11201+6305	15.7	SDSS	0.0111	3311	46.63	-17.66	HII
11371+8106	16.7	BAO 2.6m	0.0445	13053	183.85	-19.60	HII
11401+6642	15.4	SDSS	0.0554	16169	227.73	-21.36	HII
11436+7438	14.8	OHP 1.93m	0.0560	16333	230.04	-22.01	Composite
11462+6424	15.3	SDSS	0.0414	12169	171.39	-20.87	HII

11472+6647	15.3	SDSS	0.0402	11809	166.33	-20.78	Sy2
11497+6139	16.0	SDSS	0.0427	12530	176.48	-20.21	HII
11587+6124	17.4	SDSS	0.1638	45166	636.15	-21.66	HII
12008+6141	15.6	SDSS	0.0644	18689	263.23	-21.48	HII/LINER
12040+8158	17.7	BAO 2.6m	0.0903	25858	364.19	-20.14	HII
12069+6753	17.1	SDSS	0.0585	17026	239.81	-19.83	HII
12077+8131	16.5	BAO 2.6m	0.0397	11671	164.38	-19.56	HII
12119+6117	17.3	SDSS	0.1630	44948	633.07	-21.72	Em
12120+6838	15.5	SDSS	0.0602	17512	246.64	-21.48	HII
12120+6838	15.5	SDSS	0.0606	17643	248.50	-21.44	Sy1.9
12120+6838a	15.8	BAO 2.6m	0.0581	16924	238.37	-21.09	HII
12120+6838b	16.2	BAO 2.6m	0.0604	17574	247.52	-20.77	HII
12120+6838c	15.6	BAO 2.6m	0.0597	17376	244.74	-21.34	Composite
12138+7537a	16.9	BAO 2.6m	0.0497	14540	204.79	-19.67	HII
12138+7537b	15.5	BAO 2.6m	0.0468	13712	193.12	-20.93	Em
12138+7537c	16.5	BAO 2.6m	0.0525	15337	216.02	-20.17	LINER
12147+6306	13.9	SAO 6m	0.0505	14768	208.00	-22.69	Norm.
12147+6306	15.4	SDSS	0.0501	14656	206.42	-21.14	HII
12164+6437	15.0	SDSS	0.0308	9104	128.22	-20.57	HII
12207+6329	15.0	SDSS	0.0589	17154	241.61	-21.87	HII
12226+6630	17.4	SDSS	0.0865	24819	349.56	-20.37	Em
12226+6630	16.4	SDSS	0.0899	25762	362.84	-21.45	HII
12235+6253	16.6	SDSS	0.0672	19479	274.35	-20.60	AGN
12267+6540	16.4	SDSS	0.0504	14728	207.44	-20.15	HII
12312+6939	15.4	OHP 1.93m	0.0692	20035	282.18	-21.85	HII
12395+6238	14.2	SAO 6m	0.0342	10085	142.04	-21.56	Composite
12395+6238	15.7	SDSS	0.0338	9959	140.26	-20.08	HII
12470+6705	17.4	SDSS	0.1295	36345	511.91	-21.14	HII
12477+7936a	14.7	OHP 1.93m	0.0338	9972	140.45	-21.04	HII
12483+7332	14.9	OHP 1.93m	0.0314	9269	130.55	-20.68	HII
12502+7625	16.5	OHP 1.93m	0.0634	18424	259.49	-20.57	HII
13014+6146	14.0	SAO 6m	0.0272	8049	113.37	-21.27	Norm.
13014+6146	15.3	SDSS	0.0269	7968	112.23	-19.93	HII
13030+6102	15.4	SAO 6m	0.0698	20211	284.66	-21.87	HII
13045+7016a	16.2	OHP 1.93m	0.0646	18752	264.12	-20.91	Em
13045+7016b	16.4	OHP 1.93m	0.0645	18736	263.88	-20.71	HII
13121+6646	16.0	SDSS	0.0677	19627	276.43	-21.20	HII
13209+6353	17.5	SDSS	0.1997	54016	760.79	-21.91	Sy2
13234+6239	16.0	SDSS	0.0418	12272	172.84	-20.17	HII
13286+7258	14.0	OHP 1.93m	0.0305	8996	126.70	-21.51	AGN
13291+6524	17.3	SDSS	0.0365	10738	151.24	-18.62	LINER:
13300+6652	17.4	SDSS	0.1044	29694	418.23	-20.71	AGN
13300+7219	15.2	OHP 1.93m	0.0360	10594	149.21	-20.67	Composite
13367+6703	17.2	SDSS	0.0933	26680	375.78	-20.68	HII
13386+6557	16.2	SDSS	0.0480	14047	197.85	-20.26	HII
13410+7837	16.0	BAO 2.6m	0.0580	16905	238.09	-20.93	Composite

13503+6104	16.6	SDSS	0.0880	25230	355.36	-21.11	HII
13513+6616	15.9	SDSS	0.0702	20331	286.35	-21.38	HII
13524+6213	16.5	SDSS	0.0729	21076	296.84	-20.86	HII
13524+6213	16.8	SDSS	0.0730	21111	297.33	-20.53	HII
14004+7445a	15.9	BAO 2.6m	0.0670	19422	273.56	-21.28	HII
14004+7445b	16.5	BAO 2.6m	0.0673	19523	274.98	-20.67	Composite
14013+6520	16.0	SDSS	0.0345	10173	143.29	-19.74	HII
14129+6111	17.8	SDSS	0.1514	42024	591.89	-21.11	HII
14132+6552	15.4	SDSS	0.0325	9585	135.00	-20.22	HII
14190+6432	17.2	SDSS	0.1525	42288	595.61	-21.70	LINER
14196+7734	16.6	BAO 2.6m	0.0469	13732	193.41	-19.87	AGN
14230+6158	16.4	SDSS	0.1112	31523	443.98	-21.80	Sy2
14244+6435	17.8	SDSS	0.1130	31992	450.59	-20.49	HII/Sy1.9
14263+6116	15.7	SDSS	0.0618	17967	253.05	-21.28	HII
14271+6437	17.7	SDSS	0.1411	39375	554.58	-20.98	HII
14360+6129	14.5	SAO 6m	0.0521	15223	214.41	-22.16	HII
14370+6254	16.6	SDSS	0.1154	32622	459.47	-21.75	HII
14386+6113	16.0	SDSS	0.0934	26721	376.35	-21.88	HII
14418+6131A	15.0	SDSS	0.0480	14044	197.80	-21.53	HII/LINER
14418+6131A	15.7	SDSS	0.0492	14410	202.95	-20.88	HII/LINER
14418+6131A	17.5	SDSS	0.0477	13958	196.59	-18.93	Em
14418+6131A	17.1	SDSS	0.0494	14458	203.63	-19.49	HII
14436+6258	14.9	SDSS	0.0375	11036	155.44	-21.07	HII
14445+7003	17.0	BAO 2.6m	0.0518	15138	213.21	-19.65	Em
14458+7926	16.5	BAO 2.6m	0.0316	9330	131.41	-19.09	Sy2:/Comp
14464+6119	16.9	SDSS	0.1364	38151	537.34	-21.76	Sy1.2
14501+6212	14.9	SAO 6m	0.0429	12594	177.38	-21.34	Abs
14501+6212	15.0	SDSS	0.0431	12666	178.39	-21.23	HII
14561+6312	14.7	SAO 6m	0.0432	12680	178.60	-21.56	Norm.
14561+6312	15.2	SDSS	0.0426	12510	176.20	-21.08	Sy2
14570+6339	16.2	SAO 6m	0.0475	13912	195.94	-20.26	LINER
14574+7641a	16.5	BAO 2.6m	0.0446	13082	184.25	-19.83	HII
15356+6119a	16.4	SAO 6m	0.0886	25407	357.84	-21.37	HII
15356+6119b	14.4	SAO 6m	0.0882	25297	356.30	-23.36	HII
15374+6822	15.7	BAO 2.6m	0.0578	16851	237.34	-21.17	HII
15427+6141	17.6	SAO 6m	0.0573	16698	235.19	-19.26	Norm.
15449+6459	15.3	SAO 6m	0.0373	10981	154.67	-20.65	Norm.
16030+6312a	16.8	SAO 6m	0.0548	15990	225.21	-19.96	Norm.
16030+6312b	15.9	SAO 6m	0.0569	16585	233.59	-20.94	HII
16044+6727	16.1	BAO 2.6m	0.0233	6906	97.26	-18.85	HII
16049+8802	16.8	BAO 2.6m	0.0541	15803	222.58	-19.95	Sy2
16101+6345	16.5	SAO 6m	0.0563	16415	231.20	-20.32	HII
16118+6231	13.3	SAO 6m	0.0321	9476	133.46	-22.33	Composite
16119+8551	17.1	BAO 2.6m	0.0847	24335	342.74	-20.54	HII
16358+6709a	15.0	BAO 2.6m	0.0534	15593	219.62	-21.69	Em
16358+6709b	15.0	BAO 2.6m	0.0537	15687	220.94	-21.70	HII

16365+6403a	15.7	SAO 6m	0.0628	18250	257.04	-21.35	HII
16365+6403b	17.5	SAO 6m	0.0628	18250	257.04	-19.55	HII
16365+6403c	12.6	SAO 6m	0.0632	18362	258.62	-24.46	HII
16452+6418a	16.3	SAO 6m	0.0697	20183	284.27	-20.97	Sy2
16452+6418b	16.6	SAO 6m	0.0684	19820	279.15	-20.63	HII
16452+6418c	14.2	SAO 6m	0.0685	19848	279.55	-23.03	HII
16452+6418d	17.1	SAO 6m	0.0697	20183	284.27	-20.17	HII
16533+6216	16.1	SDSS	0.1059	30089	423.79	-22.03	HII
16588+6357	15.6	SDSS	0.0604	17568	247.43	-21.40	Em
17008+6444	15.0	SDSS	0.0273	8076	113.75	-20.26	HII
17008+6444	12.6	SAO 6m	0.0278	8224	115.83	-22.72	HII
17017+6416	15.2	SDSS	0.0843	24238	341.38	-22.43	HII
17017+6416	14.1	SAO 6m	0.0848	24365	343.17	-23.58	HII
17037+6207	17.2	SDSS	0.1558	43138	607.58	-21.73	Em
17037+6207	15.8	SDSS	0.0823	23669	333.37	-21.86	HII
17046+6255	20.6	SDSS	0.0874	25086	353.33	-17.17	HII
17046+6255	16.4	SDSS	0.0883	25318	356.58	-21.39	HII
17062+7544	16.1	BAO 2.6m	0.0654	18974	267.24	-21.02	Em
17089+6558	17.3	SDSS	0.0278	8238	116.03	-18.04	HII/LINER
17089+6558a	17.6	BAO 2.6m	0.0279	8256	116.28	-17.73	HII
17089+6558b	15.6	BAO 2.6m	0.0282	8335	117.39	-19.75	HII
17102+6442	16.1	SDSS	0.0789	22735	320.21	-21.42	HII
17102+6442	14.4	SAO 6m	0.0789	22739	320.27	-23.13	HII
17173+6119	15.7	SDSS	0.0717	20732	292.00	-21.67	HII
17173+6119	14.7	SAO 6m	0.0735	21242	299.18	-22.68	Composite
17190+6219	16.7	SDSS	0.0803	23117	325.59	-20.84	AGN
17190+6219	17.0	SAO 6m	0.0808	23264	327.66	-20.58	Norm.
17207+6307	15.4	SDSS	0.0338	9973	140.46	-20.30	HII
17207+6307	15.3	SAO 6m	0.0336	9911	139.59	-20.42	HII
17330+7619	17.2	BAO 2.6m	0.0769	22185	312.47	-20.26	HII
17349+6139a	16.2	SAO 6m	0.0858	24639	347.03	-21.50	HII
17349+6139b	18.8	SAO 6m	0.0862	24749	348.58	-18.91	Norm.
17442+6130	14.0	SAO 6m	0.0365	10750	151.41	-21.90	Norm.
17469+6416	14.6	SAO 6m	0.0355	10461	147.34	-21.24	HII
17552+6209	16.4	SAO 6m	0.0836	24035	338.52	-21.25	Sy2
17591+8628	15.2	BAO 2.6m	0.0233	6917	97.43	-19.72	HII
18116+6328	13.9	SAO 6m	0.0481	14083	198.36	-22.59	Norm.
18169+6433c	12.4	SAO 6m	0.0208	6175	86.97	-22.30	HII
18169+6433d	15.1	SAO 6m	0.0209	6204	87.39	-19.61	HII
18192+8650	16.3	BAO 2.6m	0.0655	19002	267.64	-20.88	HII
18247+6102	15.7	SAO 6m	0.0732	21158	298.00	-21.67	LINER
18252+6315	15.9	SAO 6m	0.0838	24090	339.30	-21.75	HII
18380+8640b	16.1	BAO 2.6m	0.0778	22435	315.98	-21.43	HII
18380+8640c	16.0	BAO 2.6m	0.0779	22451	316.21	-21.46	HII
20537+8737	14.8	BAO 2.6m	0.0195	5784	81.47	-19.77	Em

2.5. Spectral Activity Classification

The central scientific use of the BIG spectra is the classification of the activity present in the optical counterpart. The revised dataset retains the classification terminology of the original study: HII, AGN, Seyfert (Sy), LINER, Composite, Em, Norm, Abs, and Unknown, together with mixed classes such as HII/LINER and HII/Sy. A colon (:) appended to a classification in Table 2.3 denotes an uncertain classification.

HII objects are interpreted as galaxies whose spectra are characteristic of ionized regions associated with star formation. AGN and Seyfert classifications indicate spectroscopic evidence for nuclear activity, while LINER denotes a population characterised by relatively strong low-ionization emission. Composite and mixed classes are retained because a single infrared-selected object can show contributions from more than one excitation mechanism.

The classification therefore serves two complementary purposes. At the level of an individual source, it provides a physical interpretation of the optical spectrum. At the sample level, it enables comparison between the frequency of star-forming and nuclear-activity signatures. This distinction becomes particularly important in Chapter 4, where the activity mixture of the high-luminosity sample differs substantially from that of the broader BIG population.

2.6. Activity Distribution of the BIG Sample

The activity distribution of the revised 257-record BIG dataset is summarized in Table 2.4. HII is the dominant category, comprising 149 objects (58.0%). The categories AGN, Seyfert, and LINER contain 13, 17, and 12 records, respectively, giving 42 records across these three nuclear-activity labels. The Composite category contains 22 records; when the four HII/LINER and two HII/Sy records are also included, 28 records have composite or mixed classifications. The remaining categories comprise 21 Em, 13 Norm, 3 Abs, and 1 Unknown record.

The numerical distribution illustrates why the BIG sample is useful for comparative studies. Star-forming systems dominate, but no single physical interpretation describes the entire infrared-selected population. The sample includes a substantial number of objects with nuclear activity, together with systems whose spectra are composite or difficult to classify at the available quality. This diversity is an observational manifestation of the broader scientific problem addressed

throughout this Thesis: strong infrared emission can be associated with different energetic processes.

Table 2.4. Distribution of the 257 BIG objects by activity type in the revised summary.

Activity type	Number	Percentage
HII	149	58.0
AGN	13	5.1
Sy 2	17	6.6
LINER	12	4.6
Composite	22	8.6
HII / LINER	4	1.5
HII / Sy 2	2	0.8
Norm	13	5.1
Em	21	8.2
Abs	3	1.1
Unknown	1	0.4
Total	**257**	**100.0**

It is important not to interpret the relative percentages as intrinsic population fractions for all IRAS galaxies. The BIG sample is the result of an optical identification and spectroscopic follow-up programme, and the availability and quality of spectra vary among objects. Objects that are easy to identify optically and objects for which useful spectral features are visible are more likely to enter a classified spectroscopic sample. The percentages are therefore descriptive properties of the revised BIG dataset rather than a volume-complete luminosity function.

2.7. Star-Forming and HII Galaxies

The dominance of HII-type objects is a major result of the BIG study. Of the 257 objects, 149 are classified as HII, corresponding to 58.0% of the sample. In an infrared-selected population this large fraction is physically significant because the infrared emission of dusty galaxies is closely related to the reprocessing of stellar radiation by interstellar dust. Optical HII-region spectra

provide an independent line of evidence that vigorous stellar activity is present in a substantial part of the sample.

The BIG sample therefore supplies a useful empirical connection between infrared selection and optical spectroscopic evidence for star formation. The infrared measurement identifies an energetically interesting source, while the spectrum helps determine whether the associated optical galaxy has the characteristic emission-line properties of a star-forming system. This complementary interpretation is one of the main reasons the spectroscopic follow-up was undertaken.

The large HII population also provides the reference against which the AGN and composite populations can be compared. This becomes particularly important for the high-luminosity sample in Chapter 4, where the fraction of broad-line and other active systems is substantially more prominent than in the broader BIG sample.

2.8. AGN, Seyfert Galaxies, and LINERs

The BIG dataset contains several forms of nuclear activity. The detailed classification table lists 13 AGN, 17 Seyfert galaxies (Sy 2 subtype), and 12 LINERs, together with mixed categories that include HII/LINER and HII/Seyfert spectra. The presence of these objects shows that IRAS selection is not restricted to star-forming galaxies and that nuclear activity can occur in a substantial subset of the infrared-selected population.

The distinction among these categories is based on spectral properties rather than infrared luminosity alone. In particular, emission-line characteristics provide evidence that the energetic source may be associated with an active nucleus. The mixed classes are useful because they show that the excitation mechanism cannot always be represented by a single category. A galaxy can simultaneously contain star-forming regions and a nucleus that contributes a different spectral signature.

This point is relevant to the broader starburst–AGN problem. Infrared luminosity can indicate that a large amount of energy is being processed by dust, but it does not by itself determine how much of the luminosity should be attributed to young stars and how much to an active nucleus.

Optical spectroscopy provides one of the principal empirical tools for separating these possibilities.

2.9. Composite and Uncertain Spectral Classes

Twenty-two objects in the detailed distribution are assigned to the Composite category, while four are HII/LINER and two are HII/Sy systems. Such objects are especially important because they demonstrate that the spectral activity of infrared galaxies may occupy an intermediate or mixed regime. Their existence argues against interpreting the infrared galaxy population through a simple binary division into star-forming and AGN systems.

The Em category contains 21 objects. These objects show emission features, but the available spectrum does not allow a more precise activity class to be assigned. Three objects are classified as absorption galaxies and 13 as Norm, for which the star-formation rate does not exceed the normal level represented in the adopted classification scheme. One object remains Unknown.

These categories should be considered part of the information content of the dataset rather than discarded failures. An uncertain or incomplete classification identifies a limitation of the observational data and can also define a target list for future observations. The methodology is therefore naturally hierarchical: a source can first be confirmed as extragalactic, then assigned to a broad spectral group, and only later, with improved spectroscopy, moved into a more precise activity category.

2.10. Redshift, Distance, and Absolute-Magnitude Distributions

The revised BIG sample covers the redshift interval $0.0092 \leq z \leq 0.1997$. The corresponding absolute magnitudes span the range $-24.46 \leq M \leq -16.81$. Using the adopted $H_0 = 71$ km s^{-1} Mpc^{-1}, the tabulated distances extend from 38.61 Mpc to 760.79 Mpc. The range is therefore sufficiently broad to include both nearby systems and objects at several hundred megaparsecs.

The redshift distribution shown in Figure 2.1 should be interpreted in the context of the optical spectroscopic selection. The sample is not uniform in distance because the ability to obtain or identify a useful optical spectrum depends on apparent brightness, source structure, and the

availability of observations. The upper-redshift tail is therefore represented by objects that are sufficiently observable for inclusion in the spectroscopic compilation.

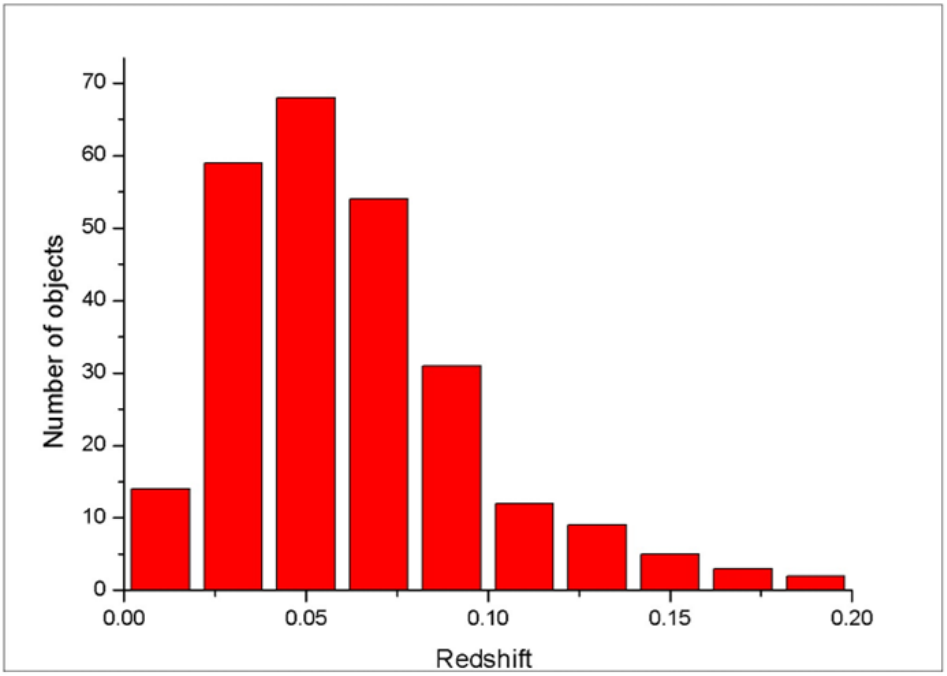


Figure 2.1. Redshift distribution of the revised 257-object BIG sample.

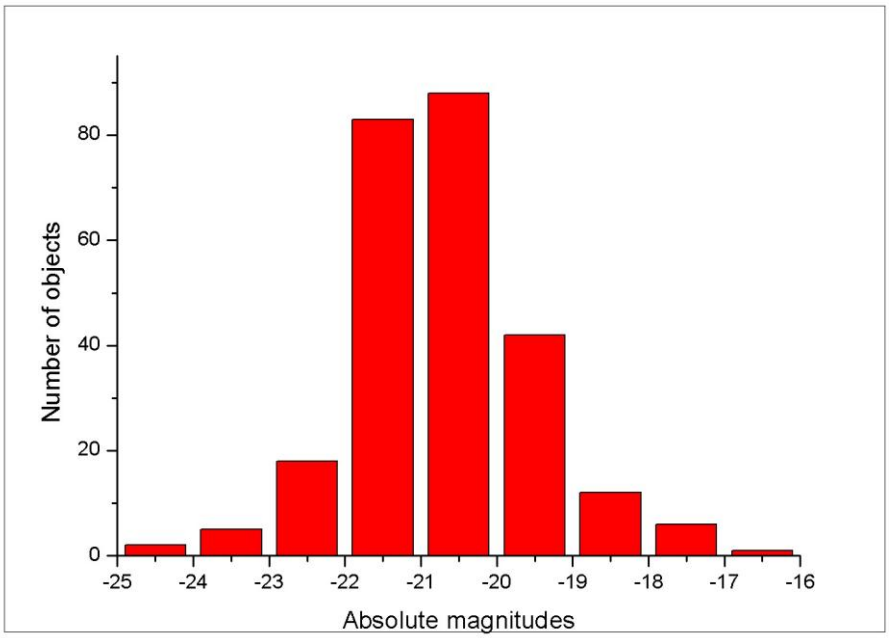


Figure 2.2. Absolute magnitude distribution of the revised 257-object BIG sample.

The same consideration applies to absolute magnitude. The observed range includes relatively faint galaxies as well as optically luminous systems. The distribution is consequently a joint result of the intrinsic population and the observational selection that produced the BIG spectroscopic dataset.

Figure 2.3 shows the M–z diagram, which makes the sample-selection pattern visually explicit. At larger redshifts, the sample is increasingly dominated by objects that are optically luminous enough to remain observable; at smaller redshifts, less luminous systems can also enter the sample. This effect should be considered when comparing activity types or apparent population fractions across distance ranges.

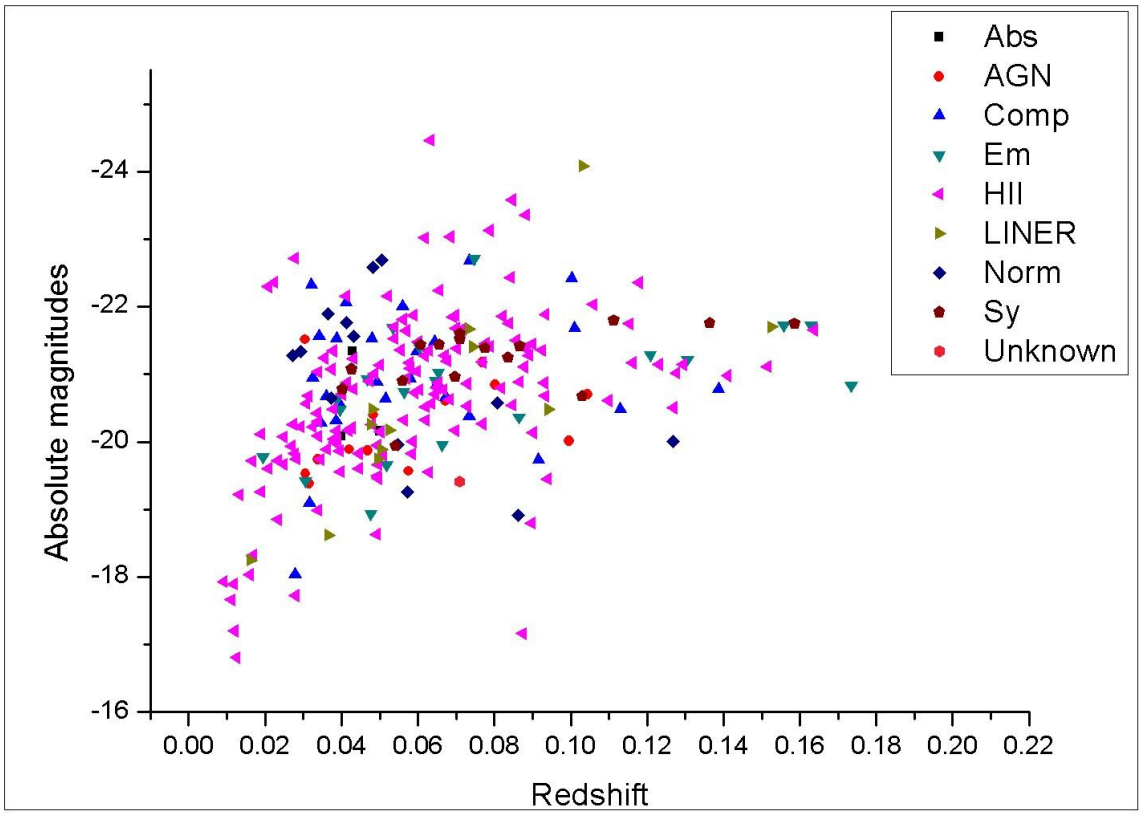


Figure 2.3. Absolute magnitude versus redshift for the revised 257-object BIG sample.

2.11. Pairs and Multiple Systems

The BIG sample includes a significant number of pairs and multiple optical systems associated with individual IRAS sources. In Table 2.3, suffixes such as “a”, “b”, and “c” identify components of the same or related systems. The existence of several optical counterparts within the positional

region of one infrared source creates an identification problem: the measured infrared emission must be associated with either one member, several members, or the interacting system as a whole [2].

Spectroscopy provides an important way to examine the physical relationship between components. Objects with similar emission-line redshifts can be considered candidate physically associated members, while different redshifts may indicate a chance projection. The individual spectra can also reveal whether the members have similar or different activity types. The BIG catalogue contains examples in which different components of a system are classified as HII, AGN, LINER, or composite.

These systems are scientifically valuable because they connect the identification problem to the physical problem of galaxy interactions. Interactions can redistribute gas, compress star-forming regions, and potentially influence nuclear activity. Binary and multiple systems provide an opportunity to establish physical connections between galaxies, determine more accurately the origin of the infrared emission, and investigate the relationship between star formation, galaxy interactions, and nuclear activity.

This topic provides a direct bridge to Chapters 3 and 4. The larger IRAS sample is designed to improve counterpart identification across the catalogue as a whole, while the high-luminosity sample emphasises the frequent occurrence of interactions and mergers among very luminous systems. The BIG sample therefore supplies an early, object-by-object demonstration of a problem that becomes increasingly important at higher infrared luminosity.

2.12. Methodological Limitations and Interpretation

Several limitations should be kept explicit when using the BIG sample for statistical interpretation. First, the sample combines spectra obtained from four different sources and is therefore not a uniform instrumental survey. Second, the spectral classification depends on the quality and wavelength coverage of the available data. Third, uncertain categories are explicitly present in the source table, indicating that not every object could be assigned a precise activity type.

The sample also contains repeated observations or multiple representations of some systems. For example, the source table includes entries with the same IRAS identifier associated with different spectra or different components. These entries are scientifically useful but require care when the sample is used for counting independent infrared sources rather than spectra or optical components.

2.13. Discussion

The BIG spectroscopic programme demonstrates the value of combining infrared selection with targeted optical follow-up. The infrared catalogue identifies an energetic population, but spectroscopy provides the information required to establish redshifts and investigate the physical processes producing the observed activity. The resulting sample is diverse, with HII galaxies dominating but AGN, Seyfert, LINER, composite, emission-line, normal, and absorption systems all represented.

The dominance of HII systems establishes star formation as a major contributor to the infrared emission in the BIG sample. At the same time, the presence of nuclear-activity classes demonstrates that a non-negligible fraction of infrared galaxies must be examined as possible AGN hosts or mixed systems. The coexistence of HII and nuclear signatures in composite or mixed classes provides a particularly direct observational motivation for the starburst–AGN connection investigated elsewhere in the Thesis.

The BIG study also illustrates why accurate source association is fundamental. An infrared beam can encompass several optical objects, and the physical interpretation of the infrared luminosity can therefore depend on whether one galaxy or a group is responsible for the emission. This observational issue is carried forward in Chapter 3 through improved catalogue cross-correlation and in Chapter 4 through the study of high-luminosity systems, where interacting and merging structures are of particular interest.

Finally, the BIG programme provides the methodological bridge to the broader data-infrastructure discussion of Chapter 5. Historical FBS and DFBS observations, dedicated spectroscopy from BAO, SAO and OHP, and modern SDSS data all contribute to the final object-

level classification. The programme consequently illustrates in practical terms how astronomical research develops from heterogeneous observations into a structured scientific sample.

2.14. Connection with the Following Chapters

The progression of the Thesis can be understood as an enlargement of scale. Chapter 2 begins with a relatively targeted sample in which optical counterparts are studied spectroscopically and classified individually. Chapter 3 extends the same scientific question—what are the optical counterparts of IRAS sources?—to the 345,163-source IRAS PSC/FSC Combined Catalogue. Chapter 4 then selects a physically extreme subset from the larger population using infrared luminosity and optical spectroscopic criteria. Chapter 5 describes the archival and computational infrastructure that can support still larger and more systematic applications.

The BIG sample is therefore not an isolated chapter. It supplies the observational foundation for the activity classifications used later, demonstrates the identification difficulties that motivate the cross-correlation work, and provides examples of interacting systems that motivate the morphological and luminosity studies of the high-luminosity population. In this sense, the BIG programme forms the observational link between individual galaxy studies and large-sample infrared astronomy.

2.15. Summary

The revised BIG dataset contains 257 spectra assembled from the BAO 2.6-m telescope, SAO 6-m telescope, OHP 1.93-m telescope, and SDSS. The combined observations provide emission-line redshifts, radial velocities, distances, absolute magnitudes, and spectral activity classifications for the studied infrared galaxies.

The sample spans $0.0092 \leq z \leq 0.1997$ and $-24.46 \leq M \leq -16.81$, with distances from 38.61 to 760.79 Mpc under the $H_0 = 71$ km s^{-1} Mpc^{-1} convention adopted by the source table. HII galaxies dominate the activity distribution, while AGN, Seyfert, LINER, composite, emission-line, normal, absorption, and uncertain systems provide additional populations.

Binary and multiple systems are an important component of the BIG programme because they connect the practical problem of infrared-source identification with the physical questions of

interaction, star formation, and nuclear activity. The spectroscopic results therefore establish the object-level observational foundation for the large-sample analysis of Chapter 3 and the high-luminosity ULIRG/HLIRG study of Chapter 4.

3. THE EXTRAGALACTIC SAMPLE OF THE IRAS PSC/FSC COMBINED CATALOGUE

3.1. Introduction

The spectroscopic study of the BIG sample established a reliable observational basis for identifying and classifying infrared galaxies. The next logical step is to extend this work to the much larger population represented by the IRAS PSC/FSC Combined Catalogue. The catalogue contains 345,163 infrared sources, of which 145,902 sources, or 42.3%, were identified in the catalogue as candidate galaxies [6] requiring further investigation through cross-correlation with optical and other astronomical catalogues.

The objective of this chapter is to describe the construction of a broad extragalactic sample from the Combined Catalogue. Because no single catalogue provides complete and unambiguous identification of all IRAS sources, several independent data sets were combined. These include catalogues of previously known galaxies, quasars and blazars, catalogues providing morphological information useful for identifying galaxy candidates, radio catalogues, and Gaia astrometric data.

3.2. The IRAS PSC/FSC Combined Catalogue

The original IRAS PSC and FSC contain many common sources, but associating the entries from the two catalogues is complicated by their different positional uncertainties. A dedicated catalogue cross-matching procedure was developed in the framework of the Armenian Virtual Observatory. Rather than applying one fixed matching radius, the procedure used a search radius corresponding to three standard deviations of the positional errors for each individual source [37]. This resulted in 73,770 PSC/FSC associations and led to the construction of the Combined Catalogue with 345,163 sources [6]. Subsequent cross-correlations with AKARI-IRC, AKARI-FIS and WISE further improved the positional information and extended the available infrared photometry.

The resulting catalogue provides a homogeneous starting point for large-scale extragalactic studies. The availability of improved positions is particularly important because the original IRAS beam sizes are substantially larger than those of modern optical surveys. Accurate positions

reduce the probability of assigning an infrared source to an unrelated optical object and are therefore essential for reliable sample construction.

3.3. Methodological Design of the Extragalactic-Sample Construction

The construction of a reliable extragalactic sample from an infrared survey is fundamentally an association problem. An IRAS detection provides infrared flux information and an approximate sky position, but the infrared measurement by itself does not uniquely specify the physical nature of the source. A robust sample therefore requires independent evidence from catalogues with different observational properties. In the present programme, the strategy proceeds hierarchically: previously known extragalactic objects are identified first; sources with optical properties characteristic of galaxies are then added as candidates; radio information is used as an additional extragalactic indicator; and Gaia astrometry is used to identify and remove stellar contaminants.

The methodology has two important characteristics. First, it uses heterogeneous evidence rather than relying on a single catalogue. Second, it separates an association from a final scientific classification. A match in a catalogue indicates a possible counterpart, whereas the final classification depends on the consistency of several observations. This distinction is essential for large infrared samples because catalogue overlap, positional uncertainty, and source confusion can produce apparently plausible but incorrect associations.

Table 3.1. Hierarchical stages used in the construction of the IRAS extragalactic sample.

Stage	Primary resource	Purpose
1	Known galaxies, QSOs, blazars	Identify known galaxies, QSOs, blazars
2	SDSS / HYPERLEDA / APM-North / Gaia Part 2	Identify galaxy candidates
3	NVSS / FIRST	Radio as an additional extragalactic indicator
4	Gaia DR3 Main	Reject likely stellar contaminants
5	All evidence combined	Final source class and confidence flag

The order in which catalogues are queried is an important component of the method. The catalogues of previously known galaxies, quasars, and blazars were placed at the highest priority.

These include the catalogue of galaxies and quasars observed in the IRAS survey [46], NGC 2000.0 [47], the Third Reference Catalogue of Bright Galaxies [48], the catalogue of quasars and active nuclei 13th edition [49], the Roma BZCAT [50], LQAC-5 [51], the SDSS DR16 quasar catalogue [52], and a QSO selection based on SDSS and WISE [53]. Search radii are selected according to the positional accuracy of the corresponding catalogue rather than imposed as one universal value.

The second level uses catalogues that provide information capable of identifying galaxy candidates. SDSS supplies morphological information distinguishing extended and point-like sources [54,55]; HYPERLEDA provides a dedicated galaxy catalogue [56]; APM-North supplies ellipticity information [57]; and Gaia DR3 Part 2 [58] provides a catalogue of extragalactic sources. The third level incorporates radio catalogues [59,60], and the final stage uses Gaia DR3 Main Source [61] information, particularly proper motions, to distinguish stellar objects.

This hierarchy prevents a simple arithmetic interpretation of the intermediate match counts. The same IRAS source can appear in several catalogues, and therefore the sums of individual cross-correlation results do not represent the number of unique extragalactic objects. The final classification must instead be produced after applying the priority rules and resolving overlapping evidence.

3.4. Choice of Search Radius and Positional Uncertainty

The choice of search radius is central to the reliability of catalogue association. The IRAS PSC/FSC Combined Catalogue was constructed using a procedure in which the positional uncertainty of each source was taken into account. For the PSC/FSC association itself, a radius corresponding to 3-sigma of the source positional errors was used, resulting in 73,770 associations and the Combined Catalogue of 345,163 sources. The same principle was applied when the combined catalogue was correlated with optical catalogues. Compact catalogues with high positional accuracy were queried with smaller radii, whereas catalogues with lower positional precision or larger positional uncertainties were queried with larger radii. The source study used 3 arcsec for SDSS and several modern catalogues, 10 arcsec for selected older extragalactic catalogues, 2 arcmin for NGC 2000.0 and RC3, 1 arcmin for HYPERLEDA, 75 arcsec for NVSS, and 15 arcsec

for FIRST. This source-dependent approach attempts to balance completeness and reliability. A radius that is unnecessarily large for a high-precision catalogue increases the probability of chance associations, while a radius that is too small for a lower-precision catalogue can reject genuine counterparts. The procedure adopted in the IRAS programme therefore treats positional uncertainty as an explicit part of the identification methodology.

3.5. Identification of Previously Known Galaxies and Active Objects

The first stage of the extragalactic-sample construction was the cross-correlation of the Combined Catalogue with catalogues of already known galaxies, quasars and blazars. The search radius was selected individually according to the positional accuracy of the corresponding catalogue, and only single matches within the adopted radius were retained as genuine associations at this stage. This provides a relatively secure subset for subsequent statistical studies and also provides reference populations against which the less certain galaxy-candidate population can be evaluated.

The matched catalogues cover several physically distinct populations. General galaxy catalogues supply conventional galaxy identifications, while quasar, active-nucleus and blazar catalogues identify sources whose optical or high-energy properties are already suggestive of nuclear activity. The inclusion of these classes is particularly relevant to the infrared-galaxy programme because the high-luminosity sample studied in Chapter 4 contains a substantial AGN component.

The resulting associations are therefore more than an inventory of known objects. They establish comparison populations that can later be used to investigate differences in infrared luminosity, optical colour, radio emission, morphology, and spectral activity.

Table 3.2. Cross-correlations with catalogues of previously known extragalactic objects

Catalogue	Catalogue size	Search radius	Single associations
Catalogued Galaxies and Quasars observed in the IRAS Survey, Version 2	11,444	10 arcsec	5,377
NGC 2000.0	13,226	2 arcmin	4,006

Catalogue	Catalogue size	Search radius	Single associations
Third Reference Catalogue of Bright Galaxies (RC3)	23,011	2 arcmin	7,956
Catalogue of quasars and active nuclei, 13th edition	168,940	10 arcsec	1,271
Roma BZCAT, 5th edition	3,561	10 arcsec	40
Large Quasar Astrometric Catalogue, 5th release (LQAC-5)	592,809	3 arcsec	1,364
SDSS quasar catalogue, DR16Q	750,414	3 arcsec	111
QSO selection from SDSS and WISE	1,604,577	3 arcsec	180

3.6. Galaxy-Candidate Selection from Optical Morphology

A large fraction of the IRAS catalogue cannot be identified simply by matching against lists of previously known galaxies. The second methodological stage therefore searches for objects whose observed properties are consistent with a galaxy even when they are not present in a pre-existing extragalactic catalogue.

SDSS is particularly useful for this purpose because it distinguishes extended objects from point-like objects. In the reported implementation, extended SDSS sources were treated as galaxy candidates. HYPERLEDA was used as a galaxy-catalogue resource, APM-North supplied an ellipticity-based criterion, and Gaia DR3 Part 2 supplied an extragalactic source catalogue. The approach therefore combines independent indicators rather than relying on a single optical catalogue.

The distinction between a candidate and a confirmed source is retained in the final population statistics. This allows the large candidate population to be carried forward for later spectroscopic or imaging confirmation rather than treating every morphological match as a secure galaxy identification.

Table 3.3. Cross-correlations used to identify galaxy candidates

Catalogue	Catalogue size	Selection information	Search radius	Associations
SDSS DR16	1,231,051,050	Extended objects treated as galaxy candidates	3 arcsec	46,396
HYPERLEDA	983,261	Galaxy catalogue / PGC2003 candidates	1 arcmin	33,309
APM-North	166,466,987	Ellipticity > 0.5 used as a galaxy-candidate criterion	3 arcsec	37,232
Gaia DR3 Part 2, Extragalactic	11,491,504	Extended sources	3 arcsec	21,599

The SDSS cross-correlation produced 46,396 single matches. HYPERLEDA yielded 33,309 associations, the APM-North catalogue produced 37,232 matches under the stated ellipticity criterion, and Gaia DR3 Part 2 provided 21,599 associations with extragalactic sources. These numbers are intermediate cross-correlation results and are not to be interpreted as independent counts of unique IRAS galaxies, because the same IRAS source may be represented in more than one catalogue.

3.7. Radio Information as an Independent Extragalactic Criterion and Gaia-based Stellar Rejection

Radio detections provide an independent observational property that can support the extragalactic interpretation of an IRAS source. The Combined Catalogue was cross-correlated with the 1.4 GHz NRAO VLA Sky Survey (NVSS) and the FIRST survey. It reported 42,072 single NVSS matches within a 75 arcsec radius and 6,640 single FIRST matches within a 15 arcsec radius.

Radio information is complementary to optical morphology because it selects sources through a different emission process. A radio association can therefore provide independent evidence that an infrared source is not a foreground star. In the hierarchy used here, radio information is best treated as a supporting criterion rather than a stand-alone physical classifier.

The final major identification stage used Gaia DR3 Main Source data. Gaia provides proper motions that can be used to recognize stellar objects and remove them from an extragalactic sample. A search radius of 3 arcsec was adopted, and 217,385 single matches were obtained. Objects with a proper-motion-to-error ratio greater than three were treated as star candidates. On this basis, 183,168 IRAS sources were identified as stellar candidates. This criterion illustrates an important methodological feature of the project: astrometric information is used not to discover galaxies directly, but to eliminate a major source of contamination before the remaining sources are interpreted as an extragalactic population.

The Gaia stage therefore complements the positive identification stages. Establishing that an IRAS source is probably Galactic prevents it from entering the extragalactic sample and improves the purity of the population that remains.

3.8. Preliminary Classification of the IRAS Population

After all correlations are completed, the source study applies a consistency rule. When independent datasets indicate the same object type, that type is treated as genuine. When the available evidence is ambiguous, the most probable type is retained with a flag. The priority order begins with catalogues of previously known extragalactic objects, followed by SDSS DR16, Gaia DR3 Part 2, APM-North, HYPERLEDA, NVSS, and FIRST.

This procedure is best understood as a hierarchical decision system rather than a single matching operation. It recognizes that catalogues have different purposes and accuracies and that a source may receive several potentially conflicting classifications. The resulting catalogue can therefore be progressively refined as new observations and catalogues become available.

The preliminary cross-correlation results provide a first large-scale classification of the 345,163 IRAS sources. At the stage reported here, 183,168 objects were identified as star candidates on the basis of Gaia proper motions with a proper-motion-to-error ratio greater than 3. A further 80,998 sources remained unknown because no convincing identification was obtained after the available cross-correlations. The sample contained 68,940 galaxy candidates and 12,057 confirmed or genuine galaxies or QSOs. These values should be treated as preliminary rather than as a final immutable catalogue.

The 12,057 confirmed or genuine extragalactic objects constitute the most secure subset at this stage, whereas the 68,940 galaxy candidates provide a substantially larger population for future verification. The 80,998 unknown sources are also scientifically relevant because they may include faint galaxies or false detections, including possible cirrus contamination. The unknown population should therefore not be interpreted automatically as an additional galaxy population.

Table 3.4. Preliminary classification of the IRAS PSC/FSC Combined Catalogue

Object category	Number of sources	Fraction of full catalogue
Confirmed galaxies or QSOs	12,057	3.5%
Galaxy candidates	68,940	20.0%
Star candidates	183,168	53.1%
Unknown sources	80,998	23.5%
Total	**345,163**	**100.0%**

The preliminary distribution demonstrates both the scale of the data set and the importance of systematic cross-correlation. Most importantly, the work shows that a substantial fraction of the Combined Catalogue can already be assigned an extragalactic or stellar nature using existing astronomical databases. At the same time, 80,998 sources remained unidentified at the stage of the publication, showing that the complete extragalactic sample requires further work.

3.9. Connection to the ULIRG and HLIRG Sample

The large extragalactic catalogue developed in this chapter is the natural parent population for the high-luminosity study of Chapter 4. The ULIRG/HLIRG work selected a much smaller set of objects by combining accurate WISE positions, SDSS spectroscopic information, and infrared luminosity calculations. Thus, Chapter 3 prioritizes broad identification and sample construction, whereas Chapter 4 prioritizes detailed physical characterization of the most luminous members.

A luminosity-selected sample can only be interpreted properly when the infrared source has a reliable optical counterpart. Conversely, the detailed ULIRG/HLIRG study demonstrates the

scientific return that can be obtained once the identification problem has been controlled. The two chapters therefore represent complementary stages of the same research programme.

3.10. Limitations and Sources of Uncertainty

Several limitations affect the preliminary extragalactic sample. The first is source confusion and positional uncertainty in the original IRAS observations. The second is catalogue incompleteness: an object absent from one optical catalogue may still be present in another. The third is overlap between catalogues, which means that individual association counts cannot simply be added. The fourth is the possibility of false or ambiguous infrared detections, especially among the remaining unknown sources.

The hierarchical procedure reduces but does not eliminate these problems. It is therefore appropriate to regard the candidate and unknown populations as subject to further refinement. This is also why the source study retains flags for ambiguous classifications instead of forcing every object into a single secure class.

These limitations explain the importance of the archival and virtual-observatory methodology developed in Chapter 5. Historical spectra and images can provide independent evidence for an IRAS source, while a VO framework makes it possible to add new catalogues without reconstructing the entire sample from the beginning.

3.11. Methodological Perspective

The main methodological achievement of the present work is the transformation of a large heterogeneous infrared catalogue into a structured population in which different levels of identification confidence can be distinguished. The workflow begins with a homogeneous infrared survey, incorporates improved positional information, applies independent optical and radio criteria, removes likely stellar contaminants, and retains flags where the available information is ambiguous.

This approach provides a general strategy for constructing reliable multiwavelength samples from surveys in which the original measurements have lower angular resolution than modern optical

observations. The method is particularly appropriate when the scientific goal is to construct large statistically useful samples rather than investigate only a small number of individual sources.

Within this Thesis, the extragalactic-sample construction is the methodological bridge between the detailed BIG spectroscopy of Chapter 2 and the high-luminosity population analysis of Chapter 4. The ArVO and BAO archive infrastructure of Chapter 5 then provides a framework for extending the same philosophy toward historical and computationally intensive datasets.

3.12. Discussion

The principal strength of the extragalactic-sample methodology is the use of independent information with different observational sensitivities. Optical morphology identifies extended sources, existing catalogues provide previously established object classifications, radio emission can support an extragalactic identification, and Gaia proper motions provide an efficient means of rejecting stars. The combination reduces the dependence on any one observational criterion.

The methodology also illustrates an important limitation. Cross-correlation results from individual catalogues cannot simply be summed because of overlaps among catalogues. The final sample must therefore be constructed through a hierarchy of identifications and consistency checks. The source study explicitly retains the most probable classification when ambiguity remains, indicating that the resulting catalogue should be considered a progressively refined data product rather than an automatically perfect classification.

3.13. Summary

The IRAS PSC/FSC Combined Catalogue provides a homogeneous all-sky starting point for the construction of a large extragalactic infrared sample. The methodology combines previously known extragalactic catalogues, optical morphology and galaxy-candidate information, radio surveys, and Gaia astrometry. The use of source-dependent search radii and a hierarchical priority scheme is central to the reliability of the process.

At the reported stage, 12,057 sources were classified as confirmed or genuine galaxies or QSOs, 68,940 as galaxy candidates, 183,168 as star candidates, and 80,998 as unknown. These populations provide the basis for further spectroscopic confirmation, luminosity calculations, IR-

to-optical studies, morphological analysis, and searches for new ULIRGs and HLIRGs. In this sense, the large extragalactic sample represents the bridge between the detailed spectroscopic study of BIG objects and the focused analysis of ULIRGs and HLIRGs.

4. HIGH-LUMINOSITY INFRARED GALAXIES: ULIRGs AND HLIRGs

4.1. Introduction

The most luminous infrared galaxies represent an extreme and scientifically important population within infrared-selected samples. Their large infrared output allows them to be detected at substantial distances and provides a way to identify systems in which substantial amounts of energy are being emitted in dust-reprocessed form. The physical origin of this emission may be dominated by vigorous star formation, by an active galactic nucleus (AGN), or by a mixture of both processes.

The present chapter develops the high-luminosity component of the IRAS research programme described in the preceding chapters. Chapter 3 established the IRAS PSC/FSC Combined Catalogue as the parent population and described the cross-correlation procedures used to identify extragalactic counterparts. Here, the same catalogue is filtered through additional optical, photometric, spectroscopic, and infrared-luminosity requirements to obtain a controlled sample of ULIRGs and HLIRGs.

The two principal papers [22,23] form a continuous research sequence. The first establishes the sample-selection procedure and identifies 114 high-luminosity objects. The second performs the detailed optical spectral classification of those objects and investigates the distribution of activity types, colours, and infrared luminosities.

4.2. Physical and Scientific Motivation

Following Sanders and Mirabel (1996), the luminosity ranges adopted in the source studies are $10^{11}\ L_\odot < L_{IR} < 10^{12}\ L_\odot$ for LIRGs, $10^{12}\ L_\odot < L_{IR} < 10^{13}\ L_\odot$ for ULIRGs, and $L_{IR} > 10^{13}\ L_\odot$ for HLIRGs [5]. These definitions provide a convenient observational framework for separating the most luminous part of the infrared-galaxy population.

The scientific motivation for isolating this population is closely related to the possible connection between starburst and nuclear activity. The studies emphasise that high infrared luminosity may be associated with intense star formation, while the most luminous systems may contain a

substantial AGN contribution [24]. They also emphasise the frequent occurrence of double, multiple, interacting, or merging structures among luminous infrared galaxies [36].

The resulting sample can therefore be used to investigate several linked observables: infrared luminosity, redshift, optical colour, spectral activity, and morphology. The key advantage is that these properties are measured or derived for the same selected population.

4.3. Parent Catalogue and Positional Refinement

The parent data set is the IRAS PSC/FSC Combined Catalogue [6] containing 345,163 sources. The catalogue combines the IRAS Point Source and Faint Source Catalogues and incorporates improved positional information from later infrared catalogues, including AKARI and WISE.

For the high-luminosity search, sources with WISE-based positional information were selected because the WISE coordinates provide substantially higher positional accuracy than the original IRAS positions. The accuracy is approximately 1 arcsec for WISE and 0.1 arcsec for SDSS. Combining the uncertainties in quadrature gives about 1.005 arcsec, and a 3-sigma search radius was adopted for matching to SDSS DR14 [18].

$$\sigma = \sqrt{\sigma_{WISE}^2 + \sigma_{SDSS}^2}$$

$$\sigma = \sqrt{1^2 + 0.1^2} \approx 1.005$$

This positional-refinement step is critical because an IRAS source can be associated with multiple optical objects within the relatively large original infrared positional uncertainty. The requirement of a unique optical counterpart within the adopted radius reduces this ambiguity before spectroscopy and luminosity analysis are performed.

4.4. Optical Counterpart and Spectroscopic Selection

The WISE-refined IRAS catalogue was cross-correlated with SDSS DR14. Only objects with spectroscopic information and a unique optical counterpart within the adopted 3-sigma radius were retained. This procedure produced 6,346 objects with confident optical counterparts.

The selection also used the SDSS spectroscopic magnitude limit, r = 17.77 mag. The resulting high-luminosity sample should therefore be understood as a spectroscopically accessible subset of the IRAS population rather than as a complete census of all ULIRGs and HLIRGs.

The use of a unique counterpart and an optical spectroscopic limit has an important methodological consequence. The sample is well suited to homogeneous comparative analysis because all objects have an optical redshift measurement, but optically faint, strongly obscured, or otherwise spectroscopically inaccessible infrared galaxies can be missed.

4.5. Four-Band IRAS Photometry Requirement

The infrared luminosity estimator requires flux densities in all four IRAS bands, 12, 25, 60, and 100 μm [21]. Out of the 6,346 objects with unique SDSS spectroscopic counterparts, 6,270 had measurements available in all four bands and therefore entered the luminosity calculation.

Requiring all four flux densities keeps the luminosity estimator homogeneous. Because the adopted expression assigns different weights to the four bands, incomplete photometry would alter the estimator and make the resulting luminosities less directly comparable.

The sequence from the full Combined Catalogue to the final luminosity-analysis sample demonstrates the progressive tightening of quality criteria used in the study.

4.6. Infrared and Far-Infrared Luminosities

Infrared and far-infrared luminosities have been calculated using the relations of Duc et al. (1997) [21]. For distance R in Mpc and IRAS flux densities f_{12}, f_{25}, f_{60}, and f_{100}, the adopted relations are:

$$L_{ir} = 5.6 \times 10^5 R^2 (13.56 f_{12} + 5.26 f_{25} + 2.54 f_{60} + f_{100}) L_{\odot}$$

$$L_{fir} = 5.6 \times 10^5 R^2 (2.58 f_{60} + f_{100}) L_{\odot}$$

The first expression uses all four IRAS bands and was used to determine whether an object crossed the ULIRG threshold. The second is dominated by the 60 and 100 μm measurements and provides a corresponding far-infrared luminosity.

Because luminosity depends on the square of the adopted distance, the redshift measurement entering the distance determination directly affects the derived infrared luminosity. The use of SDSS spectroscopy provides a homogeneous source of redshifts for the selected objects.

4.7. Construction of the High-Luminosity Sample

Among the 6,270 sources with complete four-band IRAS flux information, objects with infrared luminosities above 10^{12} $L_{\odot}$ have been retained, the lower boundary adopted for ULIRGs. The SDSS spectroscopic magnitude limit r = 17.77 mag was retained as part of the optical selection.

The complete selection pathway combines catalogue construction, positional refinement, unique optical identification, spectroscopic availability, complete four-band photometry, and an infrared-luminosity threshold. Each step removes a distinct source of uncertainty and progressively converts a very large infrared source catalogue into a scientifically controlled high-luminosity sample.

As a result, 114 extragalactic objects were selected that have high IR luminosity and correspond to ULIRG/HLIRG classes.

Table 4.1. The list of 114 ULIRGs/HLIRGs found from IRAS PSC/FSC Combined Catalogue based on SDSS redshifts; Byurakan ULIRGs.

	IRAS coordinates			IRAS fluxes (Jy)				SDSS model magnitudes						
#	RA J2000	DEC J2000	Δ (″)	F12	F25	F60	F100	u (mag)	g (mag)	r (mag)	i (mag)	z (mag)	z	LOG (L_{ir}/ L_{sun})
1	00 20 00.270	+32 08 02.82	0.19	0.29	0.19	0.33	4.76	18.97	17.53	16.70	16.16	15.87	0.110	12.059
2	00 30 25.608	-09 26 17.39	1.29	1.03	1.61	11.40	47.70	19.93	18.46	17.74	17.31	17.04	0.129	13.157
3	00 33 21.837	-05 05 44.63	2.42	0.46	5.66	27.60	99.80	19.31	17.12	16.12	15.65	15.27	0.053	12.740
4	00 44 59.123	+19 21 41.45	0.26	0.14	0.18	0.32	0.82	16.91	16.78	16.65	16.36	16.38	0.179	12.073
5	01 14 27.481	+27 21 38.98	1.33	1.34	0.76	15.60	115.00	19.33	18.07	17.16	16.55	16.13	0.109	13.277
6	01 19 56.727	+13 34 31.51	0.29	0.10	0.16	0.31	0.80	19.23	18.00	17.29	16.92	16.68	0.191	12.052
7	01 27 09.340	+08 40 15.52	2.84	0.57	0.34	2.64	39.00	19.59	17.35	16.34	15.80	15.36	0.045	12.026
8	01 37 58.267	-08 34 30.44	0.04	0.16	0.18	0.29	0.76	20.43	18.39	17.38	16.77	16.41	0.189	12.128
9	01 59 50.248	+00 23 40.98	0.12	0.12	0.54	2.22	2.16	15.93	15.70	15.52	15.26	15.30	0.163	12.444
10	02 04 22.321	-09 37 57.65	0.46	0.15	0.17	0.34	0.76	19.77	18.40	17.63	17.04	16.82	0.192	12.133
11	02 24 23.087	-02 10 30.55	0.23	0.12	0.16	0.30	0.97	20.82	19.01	17.63	17.06	16.70	0.173	12.015
12	02 28 40.542	-09 35 05.68	2.02	0.86	0.88	5.87	80.10	17.52	16.58	16.18	15.79	15.63	0.068	12.682
13	02 44 17.066	-00 30 38.22	0.26	0.11	0.16	0.48	0.83	21.51	19.00	17.65	17.11	16.68	0.200	12.147
14	03 23 22.858	-07 56 15.21	0.16	0.10	0.13	1.00	1.69	18.95	18.00	17.37	16.93	16.67	0.166	12.166
15	03 29 02.594	-00 55 39.10	0.39	0.46	0.54	1.78	12.80	17.30	15.83	15.04	14.61	14.33	0.067	12.043
16	07 37 35.222	+35 36 21.36	0.32	0.12	0.15	1.88	4.45	15.10	13.86	13.11	12.72	12.40	0.819	13.523
17	08 04 07.412	+39 19 27.69	0.20	0.23	0.13	0.27	0.73	19.73	18.38	17.54	17.05	16.76	0.163	12.064
18	08 13 31.289	+25 45 03.15	0.08	0.11	0.15	0.28	0.54	16.42	16.00	15.61	15.27	15.23	1.510	13.259
19	08 24 45.643	+15 39 43.83	0.30	0.13	0.18	0.36	0.80	20.11	18.64	17.69	17.21	16.94	0.220	12.229

20	08 31 48.881	+04 29 39.07	0.05	0.17	0.26	0.43	0.77	16.34	15.80	15.35	15.05	14.78	0.174	12.147
21	08 34 57.652	+48 45 15.63	0.19	0.09	0.09	0.34	2.43	19.19	17.60	16.70	16.21	15.93	0.175	12.101
22	08 54 48.878	+20 06 30.75	0.11	0.28	0.43	0.89	1.16	16.57	15.93	15.42	15.04	14.71	0.778	13.411
23	08 55 17.068	+13 05 09.65	0.22	0.11	0.21	0.24	1.63	19.32	18.13	17.40	16.90	16.65	0.157	12.002
24	08 57 06.358	+19 08 53.69	0.19	0.11	0.12	0.33	0.52	17.25	17.19	16.99	17.12	16.61	0.331	12.440
25	08 58 41.758	+10 41 22.04	0.16	0.11	0.19	1.12	1.95	18.93	17.49	16.62	16.08	15.81	0.148	12.134
26	08 59 23.600	+47 36 10.53	0.12	0.09	0.12	0.44	0.88	20.80	18.76	17.60	17.02	16.69	0.180	12.024
27	09 00 25.376	+39 03 54.38	0.65	0.32	1.70	7.43	4.59	18.72	17.34	16.85	16.50	16.25	0.058	12.069
28	09 05 54.723	+24 18 29.03	0.32	0.05	0.09	0.36	0.59	19.98	18.61	17.72	17.24	16.87	0.232	12.057
29	09 14 38.146	+09 33 22.95	0.12	0.11	0.15	0.28	1.19	19.54	18.27	17.26	16.78	16.43	0.246	12.300
30	09 21 59.400	+45 09 12.61	0.22	0.06	0.13	0.23	0.57	18.58	18.05	17.73	17.46	17.34	0.235	12.054
31	09 26 08.518	+28 24 02.59	0.15	0.08	0.21	0.25	0.62	19.19	18.24	17.62	17.17	17.00	0.203	12.062
32	09 58 46.856	+50 54 56.54	0.06	0.05	0.06	0.22	1.37	19.57	18.10	17.25	16.80	16.56	0.214	12.042
33	10 01 49.527	+28 47 09.01	0.10	0.16	0.26	0.23	0.34	18.54	17.96	16.83	16.63	16.81	0.185	12.100
34	10 05 32.126	+30 53 18.27	0.14	0.09	0.10	0.20	0.63	18.89	17.81	17.07	16.61	16.35	0.234	12.099
35	10 05 41.874	+43 32 40.60	0.25	0.13	0.19	0.56	0.87	16.39	16.38	16.38	16.18	16.28	0.179	12.122
36	10 26 17.480	+30 36 43.24	0.19	0.06	0.16	0.26	0.67	18.75	18.04	17.54	17.40	16.87	0.340	12.380
37	10 33 43.863	+62 11 06.27	0.56	0.08	0.08	0.20	0.57	19.76	18.49	17.57	17.12	16.84	0.229	12.038
38	10 33 59.475	+35 55 09.18	0.14	0.09	0.16	0.52	0.79	17.10	16.97	16.66	16.35	16.28	0.169	12.002
39	10 40 00.549	+48 57 44.44	0.19	0.08	0.11	0.64	0.66	19.16	18.34	17.76	17.27	17.20	0.176	12.013
40	10 52 23.573	+44 08 47.17	0.73	0.12	0.17	3.53	5.41	19.37	17.59	16.83	16.32	16.04	0.092	12.114
41	10 53 30.955	+33 13 42.39	0.14	0.11	0.13	0.21	0.54	19.77	18.56	17.63	17.20	16.90	0.244	12.170
42	10 58 39.325	+38 29 06.77	0.40	0.05	0.09	0.62	0.75	19.78	18.35	17.53	17.10	16.83	0.208	12.086
43	11 08 32.952	+40 37 31.72	0.07	0.10	0.13	0.68	1.02	19.39	18.07	17.36	16.92	16.66	0.166	12.045
44	11 12 11.118	+27 32 56.18	0.64	0.07	0.25	0.25	0.83	18.46	17.82	17.42	17.13	17.04	0.235	12.210
45	11 14 38.892	+32 41 33.60	0.10	0.17	0.35	1.59	1.52	21.22	18.94	17.13	16.09	15.80	0.188	12.447
46	11 16 03.133	+02 08 52.30	0.04	0.17	0.19	0.40	0.33	16.81	16.83	16.68	16.44	16.50	0.211	12.219
47	11 21 14.226	+03 25 46.82	0.07	0.17	0.22	0.36	1.16	17.46	17.29	17.05	16.66	16.66	0.152	12.039
48	11 22 33.084	+50 15 27.98	0.21	0.09	0.12	0.19	0.95	19.98	18.52	17.69	17.12	16.96	0.195	12.010
49	11 41 22.043	+40 59 50.39	0.16	0.20	0.14	1.02	1.51	19.99	18.33	17.42	16.89	16.55	0.149	12.156
50	11 44 30.922	+25 46 34.98	0.38	0.15	0.15	0.29	0.65	18.64	17.44	16.56	16.12	15.84	0.185	12.082
51	11 53 14.237	+13 14 27.74	0.19	0.10	0.30	2.58	3.32	19.88	18.30	17.59	17.04	16.82	0.127	12.259
52	11 57 44.359	+12 07 50.97	0.26	0.17	0.13	0.22	0.55	18.78	17.57	16.78	16.35	16.06	0.183	12.048
53	11 57 53.218	-03 15 37.03	0.21	0.12	0.30	0.33	0.63	17.91	17.74	17.20	16.69	16.82	0.214	12.231
54	11 58 16.724	+13 26 24.23	0.09	0.10	0.16	0.23	0.53	17.95	17.59	17.40	17.26	17.12	0.439	12.606
55	12 01 16.658	+23 34 52.85	0.34	0.12	0.14	0.26	0.60	19.19	18.28	17.53	17.22	16.90	0.263	12.281
56	12 02 05.612	+11 28 12.13	0.29	0.11	0.16	0.95	1.15	18.74	17.93	17.43	17.03	16.93	0.194	12.253
57	12 02 26.765	-01 29 15.35	0.13	0.36	0.54	2.41	2.72	17.28	16.96	16.62	16.27	16.10	0.150	12.506
58	12 04 24.554	+19 25 09.99	0.32	0.11	0.37	1.76	1.78	19.50	18.15	17.46	17.02	16.79	0.168	12.360
59	12 05 26.443	+22 43 37.56	0.15	0.62	0.54	0.51	21.00	19.42	17.78	16.77	16.16	15.74	0.076	12.250
60	12 11 19.098	+05 55 23.85	0.21	0.15	0.17	0.24	0.55	19.67	18.38	17.53	16.98	16.72	0.201	12.128
61	12 15 09.822	+09 27 09.37	0.27	0.09	0.27	0.21	0.45	20.15	18.66	17.61	17.17	16.90	0.263	12.286
62	12 29 06.906	+32 24 17.75	0.30	0.09	0.15	0.57	0.76	19.98	17.79	17.05	16.64	16.40	0.173	12.022
63	12 31 21.374	+27 55 23.93	0.10	0.08	0.13	0.35	0.67	20.18	18.41	17.44	16.96	16.69	0.212	12.090
64	12 37 16.002	+04 01 48.67	0.57	0.10	0.32	0.66	1.03	19.18	17.67	16.81	16.30	15.95	0.146	12.025
65	12 38 20.209	+17 50 39.22	0.16	0.14	0.13	0.26	0.39	17.22	16.61	16.43	16.29	16.24	0.452	12.658
66	12 45 26.988	+65 23 58.40	0.05	0.06	0.07	0.21	0.65	17.56	17.48	17.30	17.36	16.95	0.320	12.235
67	12 51 00.438	+66 03 26.87	0.07	0.06	0.07	0.17	0.58	16.78	16.75	16.62	16.79	16.41	0.282	12.117
68	12 51 20.052	+02 19 02.43	0.25	0.09	0.13	0.47	0.89	19.40	18.42	17.58	17.18	16.82	0.253	12.290
69	12 54 10.444	+03 59 51.70	0.20	0.12	0.29	0.25	0.82	19.65	18.30	17.47	16.89	16.58	0.168	12.032
70	12 56 04.931	+02 38 31.96	0.22	0.11	0.24	0.29	1.49	19.81	18.45	17.37	16.85	16.50	0.219	12.272
71	12 57 39.331	+08 09 31.74	0.04	0.09	0.12	0.59	0.78	17.76	17.47	17.17	17.24	16.91	0.272	12.362
72	13 01 18.171	+11 06 01.88	0.09	0.12	0.14	0.29	0.83	19.97	18.70	17.55	17.06	16.88	0.174	12.002
73	13 06 18.579	+08 52 41.43	0.21	0.13	0.26	0.29	0.62	17.78	17.55	17.29	16.92	16.95	0.172	12.041
74	13 22 11.965	+01 30 34.31	0.17	0.10	0.32	0.33	0.72	18.13	17.74	17.37	17.22	17.59	0.184	12.114
75	13 22 44.041	+63 37 25.16	0.58	0.06	0.07	0.59	0.91	19.85	18.41	17.50	17.04	16.83	0.200	12.064
76	13 24 19.895	+05 37 04.79	0.15	0.26	0.40	1.17	0.71	18.73	18.25	17.55	17.04	16.98	0.203	12.494
77	13 37 15.938	+03 09 36.79	0.33	0.11	0.16	0.22	0.69	19.11	18.12	17.33	16.81	16.56	0.192	12.035
78	13 44 42.108	+55 53 13.31	0.53	0.24	2.28	21.70	21.40	16.59	14.99	14.31	13.84	13.71	0.037	12.090
79	13 47 33.379	+12 17 24.35	0.30	0.14	0.67	1.92	2.06	18.16	16.37	15.35	14.79	14.59	0.120	12.201

80	13 48 40.199	+58 18 52.16	0.96	0.05	0.07	1.27	1.73	19.90	18.44	17.57	17.10	16.78	0.158	12.105
81	13 50 43.182	+37 24 34.11	0.18	0.11	0.09	0.30	0.94	19.46	17.97	17.00	16.60	16.30	0.248	12.239
82	13 55 50.208	+20 46 14.67	0.15	0.19	0.21	0.26	0.54	16.16	15.94	15.66	15.31	15.48	0.197	12.184
83	14 04 38.810	+43 27 07.52	0.11	0.12	0.29	0.62	0.99	16.18	15.63	15.30	15.33	14.92	0.323	12.625
84	14 18 42.223	+42 33 43.39	0.02	0.10	0.09	0.31	0.57	16.54	16.40	16.43	16.44	16.37	0.420	12.556
85	14 22 31.386	+26 02 05.42	0.39	0.18	0.15	1.49	1.99	18.28	17.49	17.08	16.70	16.54	0.159	12.287
86	14 25 16.043	+16 23 07.97	0.16	0.10	0.25	0.54	0.98	19.47	17.76	16.89	16.47	16.25	0.185	12.154
87	14 27 27.266	-00 58 41.16	0.19	0.12	0.32	0.91	1.28	20.05	18.44	17.60	17.05	16.69	0.162	12.183
88	14 35 35.554	+01 28 34.73	0.18	0.13	0.22	0.35	0.65	20.10	18.70	17.77	17.33	17.00	0.232	12.281
89	14 40 12.775	+61 56 33.19	0.20	0.06	0.09	0.32	0.68	17.20	17.01	16.83	16.97	16.55	0.276	12.198
90	14 41 04.407	+53 20 08.94	0.33	0.07	0.35	1.95	2.39	18.27	17.08	16.47	15.75	15.72	0.105	12.002
91	14 49 20.716	+42 21 01.45	0.23	0.11	0.23	0.40	0.56	19.10	18.05	17.28	16.95	16.81	0.179	12.063
92	14 50 54.177	+35 08 37.94	0.17	0.07	0.09	0.64	1.21	18.48	17.57	16.91	16.38	16.19	0.206	12.162
93	14 55 49.433	+48 54 36.34	0.08	0.08	0.20	0.53	0.54	18.82	18.22	17.36	17.28	17.07	0.247	12.281
94	15 02 59.016	+03 40 04.00	0.16	0.09	0.08	0.23	1.09	24.25	18.30	17.49	17.05	16.83	0.218	12.098
95	15 05 39.529	+57 43 07.33	0.30	0.12	0.07	1.02	1.50	19.07	18.30	17.76	17.31	17.18	0.151	12.075
96	15 13 18.050	+24 46 55.40	0.19	0.06	0.17	0.52	0.56	19.19	18.07	17.13	16.61	16.34	0.220	12.134
97	15 22 38.091	+33 31 36.10	0.25	0.08	0.35	1.77	1.89	17.31	16.79	16.45	15.89	16.06	0.125	12.109
98	15 24 43.874	+23 40 10.41	0.20	0.07	0.18	1.30	1.48	19.31	17.94	17.11	16.58	16.32	0.139	12.045
99	15 34 57.252	+23 30 11.56	1.79	0.48	7.91	104.00	112.00	15.63	13.96	13.08	12.61	12.29	0.018	12.149
100	15 43 16.002	+14 07 31.93	0.21	0.07	0.10	0.20	1.11	19.19	18.24	17.58	17.18	16.94	0.221	12.080
101	15 43 51.941	+16 24 22.26	0.21	0.08	0.12	0.23	0.55	16.93	16.45	16.25	16.12	15.90	0.848	12.924
102	15 45 30.246	+48 46 08.97	0.09	0.07	0.11	0.28	0.45	16.43	16.32	16.30	16.30	16.06	0.400	12.460
103	16 18 49.253	+44 25 17.28	0.07	0.05	0.11	0.28	0.50	17.34	16.98	16.64	16.67	16.15	0.335	12.301
104	16 32 21.419	+15 51 45.77	0.64	0.07	0.12	1.48	1.99	19.67	18.52	17.70	17.26	16.96	0.242	12.527
105	16 42 58.823	+39 48 37.07	0.17	0.11	0.29	0.60	1.38	16.16	15.68	15.39	15.10	14.86	0.594	13.047
106	17 04 41.377	+60 44 30.56	0.07	0.07	0.13	0.30	0.70	15.60	15.43	15.25	15.19	14.82	0.372	12.458
107	20 55 19.286	+00 32 26.16	0.40	1.53	1.82	44.00	58.90	16.58	15.15	14.42	14.04	13.74	0.023	12.023
108	21 48 33.443	+11 31 47.78	0.03	0.05	0.07	0.45	1.80	19.93	18.13	17.19	16.69	16.37	5.225?	13.556?
109	21 52 36.547	-07 42 09.35	0.26	0.35	0.47	3.77	82.70	17.13	15.54	14.77	14.35	14.03	0.035	12.071
110	22 36 06.403	+12 18 43.84	0.20	0.38	0.53	3.57	49.10	18.57	17.25	16.54	16.09	15.83	0.071	12.492
111	22 53 02.457	+17 20 18.57	0.16	0.10	0.14	0.58	2.36	19.62	18.07	17.13	16.52	16.27	0.179	12.196
112	22 58 32.466	+07 49 40.59	0.20	0.12	0.22	0.61	1.14	18.76	18.00	17.58	17.40	16.99	0.396	12.755
113	23 08 33.953	+05 21 29.87	0.10	0.20	0.43	1.15	0.83	18.38	17.42	16.53	15.91	15.51	0.173	12.335
114	23 56 13.870	-00 32 10.00	1.51	0.14	0.12	0.24	0.59	19.19	16.90	16.07	15.68	15.38	0.189	12.041

4.8. Composition of the Sample and Newly Identified Objects

The final high-luminosity population contains 114 objects: 107 ULIRGs and 7 HLIRGs. The objects were checked against the NASA/IPAC Extragalactic Database [62]. Of the 107 ULIRGs, 59 were already known, while 2 of the 7 HLIRGs had previously been classified as ULIRGs. Consequently, 48 ULIRGs were newly identified and all 7 HLIRGs were newly classified as HLIRGs in the selected sample.

These newly identified objects were designated “Byurakan ULIRGs”. Their significance lies not only in the number of new objects but also in the homogeneous route by which they were selected from the Combined Catalogue using improved positions and SDSS spectroscopy.

Table 4.2. Composition of the high-luminosity infrared-galaxy sample.

Class	Number	Status reported in source study
ULIRGs	107	59 previously known; 48 new
HLIRGs	7	7 new in the selected sample
Total	**114**	**High-luminosity extragalactic IR objects**

4.9. Luminosity-Redshift Distribution and Selection Effects

The selected objects were examined using luminosity-redshift and optical-colour diagrams. The ULIRG distribution occupies approximately $\log(L_{IR}/L_{\odot}) = 12.0–12.8$ and redshifts of roughly $z < 0.5$. The upper-redshift region is affected by the SDSS optical magnitude limit, while the small number of HLIRGs prevents a homogeneous description of their distribution.

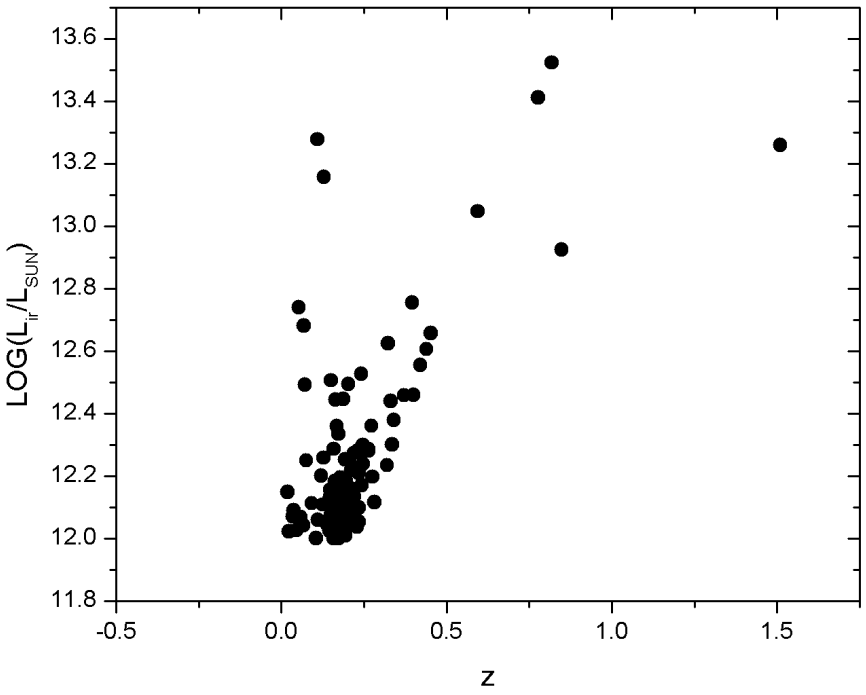


Figure 4.1. IR luminosity vs. redshift for 114 ULIRGs and HLIRGs (object #108 is left out from the Figure because of its high redshift).

The apparent increase of luminosity toward larger redshift must therefore be interpreted as a combination of detectability and intrinsic luminosity. More distant objects must be intrinsically

brighter to satisfy the flux and optical selection requirements. Thus, the luminosity-redshift distribution is a useful description of the sample but is not, by itself, evidence for an evolutionary relation.

This selection effect is particularly important when comparing the 114-object sample with the broader IRAS population. The high-luminosity sample is deliberately optimized for reliable luminosity and spectroscopic characterization, not for volume completeness.

4.10. Spectroscopic Classification: General Method

The second study analyzed SDSS spectra for the 114 high-luminosity infrared galaxies and assigned activity classifications to 113 of them; one remained unclassified because of poor spectral quality.

The classification was performed primarily by direct visual inspection of the spectra. Many of the spectra contain weak lines affected by noise, making a completely uniform application of narrow-line diagnostic diagrams difficult. Visual inspection also makes it possible to recognize broad components that are important for Seyfert subclasses 1.0–1.9 [63].

This approach is especially important when broad and narrow emission components overlap. If the broad component is not separated, the measured total intensity does not correspond to the narrow component alone and can distort the diagnostic ratios. Line widths have been considered and complete spectral appearance together with the available line ratios.

4.11. Classification Criteria and Sample Spectra

Different types of active galaxies with narrow lines have been distinguished according to the following criteria:

- Seyfert galaxies versus LINERs: $[OIII]/H_\beta > 3$
- AGN versus HII regions: $[NII]/H_\alpha > 0.6$ and $[OI]/H_\alpha > 0.1$

The following activity types are observed among our objects:

S1.0 – Seyfert 1 galaxies with broad lines. These objects have broad permitted Balmer lines of hydrogen (HI) and narrow forbidden lines. The narrow hydrogen lines are completely swamped

by the strong broad components. Physically, these objects are similar to quasars, but they have lower luminosities ($M_{abs} > -23$) [64] and H_β/[OIII]5007 > 5 [65].

NLS1.0 – Narrow-line Seyfert 1 galaxies. These are defined as soft X-ray sources whose permitted broad lines are relatively narrow (approximately 2,000 km/s), with widths only slightly greater than those of the forbidden lines. Various emission lines of FeI, FeII, and FeIII are present, as well as highly ionized iron lines, such as [FeVII] and [FeX] [66].

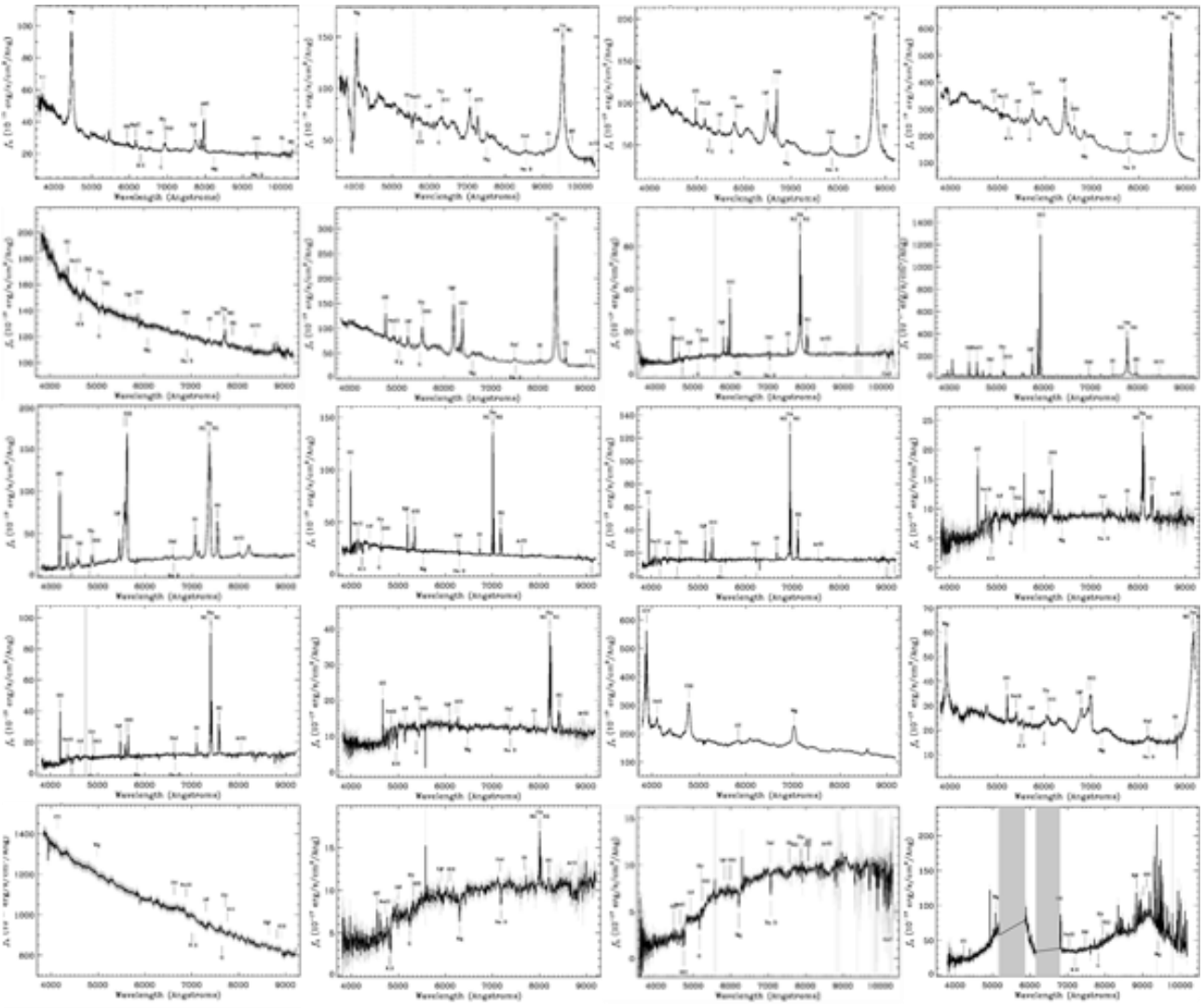


Figure 4.2. Typical examples of each activity type (from left to right and from top to bottom): *S1.0, NLS1.0, S1.2, NLS1.2, S1.5, NLS1.5, S1.8, S2.0, LINER, HII, S1.9/HII, S1.9/LINER, LINER/HII, NLAGN, QSO, NLQSO, BL Lac, Em, Abs, Unkn.*

S1.2 – Seyfert 1.2 galaxies. These are AGNs whose spectra have intermediate characteristics between those of classical Seyfert 1 and Seyfert 2 galaxies. Both broad and narrow components

of permitted lines are present; in our sample, these profiles are seen in the H_α and H_β lines [64]. Nevertheless, the broad lines are stronger, and $2.0 < H_\beta/[OIII]5007 < 5.0$ [65].

NLS1.2 – Narrow-line Seyfert 1.2 galaxies. These are defined as soft X-ray sources whose permitted broad lines are relatively narrow (approximately 2,000 km/s), with widths only slightly greater than those of the forbidden lines. Various emission lines of FeI, FeII, and FeIII are present, together with lines of highly ionized iron, such as [FeVII] and [FeX] [66].

S1.5 – Seyfert 1.5 galaxies. These are AGNs with spectra intermediate between those of classical Seyfert 1 and Seyfert 2 galaxies. A narrow HI line profile is clearly visible superimposed on the broad wings [64], and $0.333 < H_\beta/[OIII]5007 < 2.0$ [65]. The intensities of the broad and narrow components are approximately equal.

NLS1.5 – Narrow-line Seyfert 1.5 galaxies. These are defined as soft X-ray sources whose permitted broad lines are relatively narrow (approximately 2,000 km/s), with widths only slightly greater than those of the forbidden lines. Various emission lines of FeI, FeII, and FeIII are present, together with lines of highly ionized iron, such as [FeVII] and [FeX] [66].

S1.8 – Seyfert 1.8 galaxies. These are AGNs with spectra intermediate between those of classical Seyfert 1 and Seyfert 2 galaxies. They have relatively weak broad H_α and H_β components superimposed on stronger narrow lines, with $H_\beta/[OIII]5007 < 0.333$ [65].

S1.9 – Seyfert 1.9 galaxies. These are AGNs with spectra intermediate between those of classical Seyfert 1 and Seyfert 2 galaxies. They have a relatively weak broad H_α component superimposed on a stronger narrow line. The broad H_β component is not observed [64], and $H_\beta/[OIII]5007 < 0.333$ [65].

S2.0 – Seyfert 2 galaxies. Their spectra show relatively narrow permitted Balmer and forbidden lines compared with Seyfert 1 galaxies, with nearly identical full widths at half maximum (FWHM), typically in the range of 300–1,000 km/s. No broad component is visible. A secondary classification criterion is $[OIII]5007/H_\beta \geq 3$, which makes it possible to distinguish these objects from Sy1 or NLS1 galaxies [67].

LINER – Low-Ionization Nuclear Emission-line Regions. This term was introduced by Heckman [68] for galactic nuclei with emission lines formed in low-ionization regions. Their spectra

resemble those of Seyfert 2 galaxies, with relatively strong low-ionization lines, such as [OI] and [OII]. The line ratios are [OII]3727 / [OIII]5007 ≥ 1 and [OI]6300 / [OIII]5007 ≥ 1/3. According to Ho et al. [69], there are two classes of LINER galaxies: Type 1 exhibits broad Balmer emission similar to that of Seyfert 1 galaxies, whereas Type 2 shows no broad H_α component, resembling Seyfert 2 galaxies.

HII – Extragalactic HII regions. Their spectra resemble those of starburst (SB) galaxies, with strong, narrow emission lines (FWHM ≤ 300 km/s), but with [OIII] / H_β ≥ 3 and [NII]6583 / H_α < 0.6, together with a blue continuum [67, 70]. Essentially, HII galaxies and starburst galaxies are similar; however, classifying an object as a starburst galaxy requires information about its star-formation rate.

Composite objects (HII / LINER, HII / Sy, or LINER / Sy). These objects have complex spectra containing features characteristic of HII regions and LINERs, HII regions and Seyfert galaxies, or LINERs and Seyfert galaxies. In some cases, features of all three types are present [71]. Previously, these objects were classified as galaxies with transition spectra.

AGN – Active Galactic Nuclei. These are relatively low-quality spectra in which only a few emission lines are observed, mainly H_α together with [NII] lines. The [NII] / H_α ratio is typical of AGNs, namely Seyfert galaxies or LINERs. Of course, LINER-like emission spectra are not necessarily associated with nuclear activity, but they can still be useful for preliminary classification.

Em – Emission-line spectra. These are relatively low-quality spectra in which one or more emission lines are observed, but precise classification is not possible. Such spectra usually have a strong stellar component, including continuum emission and absorption features, making the emission lines difficult to distinguish from the continuum and absorption lines.

Abs – Absorption-line spectra. These spectra show only absorption lines superimposed on a continuum. The most prominent features are generally stellar lines such as MgI 5175, NaI 5890–5896, and the Balmer lines. Infrared, radio, or X-ray sources with such spectra may often contain hidden AGNs.

In the high-luminosity sample, the criteria were not applied mechanically. Visual inspection was necessary to account for broad components, noisy lines, and cases in which several activity signatures may coexist. This combined procedure is a methodological feature of the study and should be preserved when the sample is used for subsequent statistical analysis.

4.12. Detailed Spectral Activity Types

The detailed classification contains BL Lac (1), QSO (1), narrow-line QSO (1), S1.0 (1), NLS1.0 (2), S1.2 (1), NLS1.2 (15), S1.5 (3), NLS1.5 (6), S1.8 (1), S2.0 (5), LINER (14), HII (36), S1.9/LINER (3), S1.9/HII (3), LINER/HII (8), NLAGN (6), emission-line (3), absorption-line (3), and unclassified (1) objects. For statistical analysis, the study groups these detailed types into nine broad categories.

The grouped population is dominated by HII galaxies and broad-line Seyferts, followed by LINERs and composite systems. The distribution is shown in Table 4.3.

Table 4.3. Grouped spectral activity types of the 114-object sample.

Activity group	Number	Subclasses included
BLL/QSO	3	BL Lac, QSO, NLQSO
BLS	28	S1.0, NLS1.0, S1.2, NLS1.2, S1.5, NLS1.5
NLS	6	S1.8, S2.0
NLAGN	9	NLAGN, S1.9/LINER
LINER	14	LINER
Composite	11	S1.9/HII, LINER/HII
HII	36	HII / star-forming
Em	3	Emission-line
Abs	3	Absorption-line
Unclassified	1	Without classification
Total	**114**	

The distribution demonstrates that high infrared luminosity does not correspond to a single optical activity mechanism. Star-forming systems remain the largest individual class, while AGN-related systems collectively represent a substantial fraction of the population.

4.13. Mean Properties by Activity Type

Table 4.4 lists the mean redshift, mean SDSS u-g and g-r colours, and mean log infrared luminosity for the grouped activity classes. The BLL/QSO group has the highest mean redshift, followed by the broad-line Seyfert group, while the HII, LINER, composite, NLS, and NLAGN groups have lower mean redshifts. Em group high mean redshift depends on object #108 (z=5.225) which is not reliable. All objects lie at Galactic latitudes $|b| > 30$ degrees and therefore the colour indices are reported without Galactic-extinction correction. Also 108 of the 114 objects have $z \leq 0.45$ and do not require a k-correction.

Table 4.4. Mean redshift, SDSS colours, and infrared luminosity by activity class.

Activity group	N	Mean z	Mean u-g	Mean g-r	Mean log $L_{IR}/L_{\odot}$
BLL/QSO	3	0.89	0.60	0.44	13.14
BLS	28	0.30	0.34	0.31	12.37
NLS	6	0.19	1.20	0.90	12.10
NLAGN	9	0.17	1.41	0.85	12.14
LINER	14	0.16	1.60	0.86	12.12
Composite	11	0.19	1.23	0.78	12.21
HII	36	0.17	1.31	0.74	12.19
Em	3	1.83	1.78	1.01	12.86
Abs	3	0.19	2.20	1.18	12.07

The mean properties listed in Table 4.4 suggest a general tendency for several lower-activity groups to have redder optical colours and lower mean infrared luminosities than the more powerful AGN groups. However, these trends are not strictly monotonic across all activity classes, and the small number of objects in some groups limits the strength of the conclusions.

The SDSS DR14 (u–g) versus (g–r) colour–colour diagram is shown in Figure 4.3. Table 4.4 also indicates differences in mean redshift and infrared luminosity among the activity types. Because more luminous objects can be detected at greater distances, these differences must be interpreted in the context of observational selection rather than as direct evidence of an evolutionary sequence.

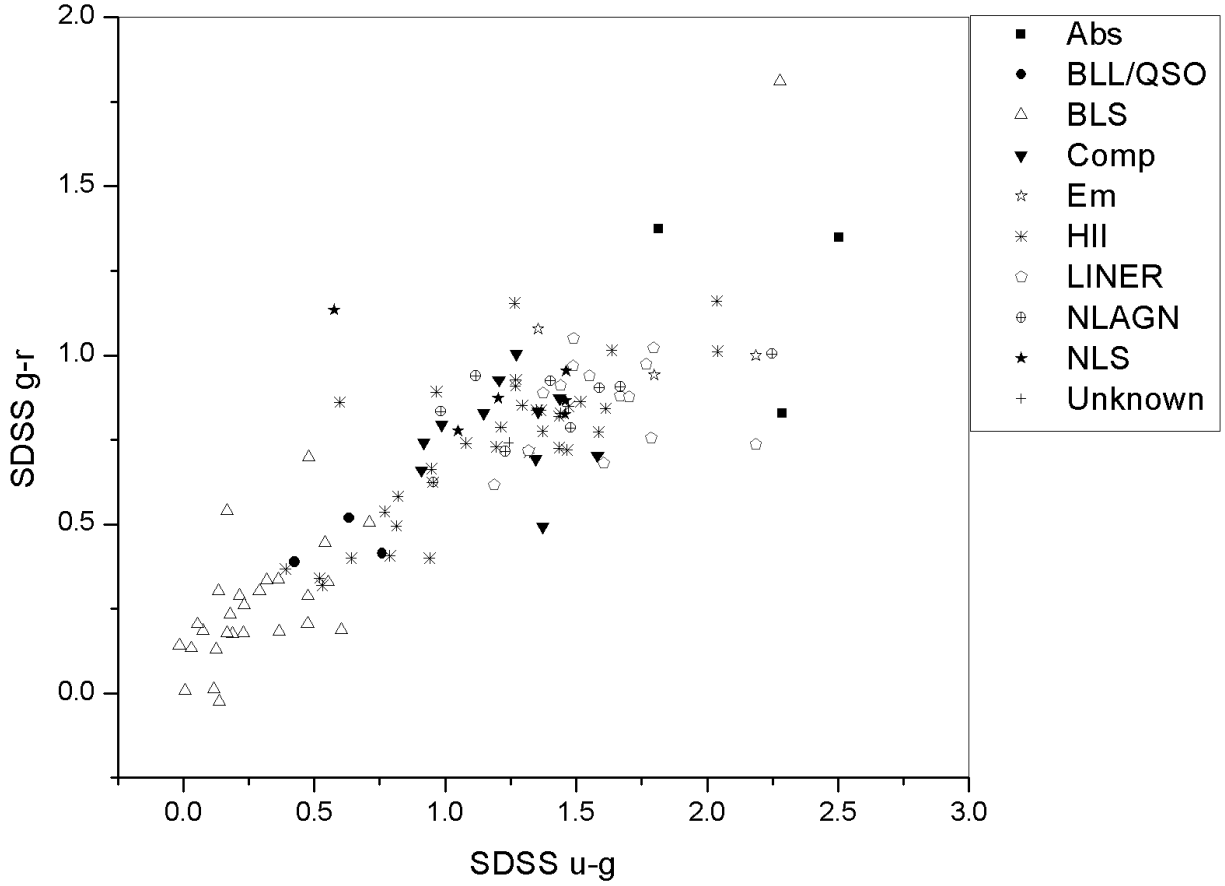


Figure 4.3. SDSS colour-colour diagram u-g / g-r.

It is noticeable that emission-line galaxies occupy a region of the diagram close to that of high-luminosity objects in terms of their properties. This may indicate that a more detailed classification could reveal Seyfert galaxies with broad lines or quasars among them. The colour analysis also shows a systematic transition from blue to red when moving from high-luminosity objects toward lower-luminosity ones.

4.14. Starburst and AGN Activity

The simultaneous occurrence of HII, Seyfert, LINER, quasar, and composite classes demonstrates that the infrared luminosity of a galaxy is not sufficient to determine the dominant mechanism producing the observed energy output. Spectroscopy is required to identify whether the optical spectrum is more consistent with star formation, nuclear activity, or mixed excitation.

The source studies explicitly raise the question of whether starburst activity triggers AGN activity, whether nuclear activity affects star formation, or whether both arise from a common physical process. The present sample does not establish a causal sequence. It does, however, provide a well-defined population in which such relations can be investigated statistically.

The coexistence of star-forming and AGN-related objects also motivates combining optical spectral information with infrared colours, X-ray data, radio information, and morphology in future work.

4.15. Comparison with Existing Activity Information

The classification results have been compared with activity classifications available in the NASA/IPAC Extragalactic Database. Only 44 of the 114 objects had activity information in that database, and only 6 of those 44 were regarded as correctly or completely classified.

This comparison illustrates the importance of examining the original SDSS spectra rather than relying exclusively on database labels. This study attributes its improved classification to the explicit consideration of broad spectral components and the overall appearance of the emission-line profiles.

4.16. Comparison with the BIG Population

The 114-object sample provides an important comparison with the 257-object BIG sample analyzed in Chapter 2. HII galaxies dominate the broader BIG sample, representing 58.0% of the classified population, whereas the high-luminosity sample contains a much larger relative contribution from broad-line and other AGN-related classes.

The two populations are not identical selection samples and therefore should not be compared as though they were drawn from the same completeness function. Nevertheless, the difference is scientifically informative: selecting the high-luminosity end of the infrared population changes the relative prevalence of the activity classes.

This comparison establishes a clear progression within the Thesis from the general IRAS-selected galaxy population to a deliberately isolated high-luminosity subset.

4.17. Relation to the IRAS Extragalactic Sample

The 114-object sample should be regarded as a physically selected subset of the much larger IRAS extragalactic population described in Chapter 3. The broad sample provides the parent distribution of infrared sources, while the ULIRG/HLIRG selection isolates objects above a defined luminosity threshold for detailed investigation.

This hierarchical relationship is important scientifically. It means that the 114 objects were not selected independently from unrelated literature compilations but were extracted from the same Combined Catalogue that serves as the basis for the larger extragalactic-sample work. The high-luminosity analysis therefore represents an extension of the same catalogue-integration methodology.

The connection also makes it possible to plan future analyses in which activity, luminosity, colour, and morphology can be investigated continuously across the infrared-luminosity range rather than only at the ULIRG threshold.

4.18. Relation to the ArVO and BAO Archive Framework

The selection of the 114-object sample depends on the same general data-integration philosophy that underlies the Armenian Virtual Observatory work described in Chapter 5. Improved infrared positions, optical catalogue matching, spectroscopic retrieval, and the combination of heterogeneous measurements are all examples of an integrated astronomical-data workflow.

The BAO archive adds a historical dimension to this methodology. For selected high-luminosity infrared galaxies, archival optical observations could provide independent historical evidence, additional spectral information, or morphological comparisons. The DFBS and its automated-

classification developments also demonstrate how large historical datasets can increasingly be processed computationally.

Thus, the ULIRG/HLIRG sample is both a scientific result and a demonstrator of the broader methodological framework developed throughout this Thesis.

4.19. Summary

The high-luminosity infrared-galaxy study begins with the 345,163-source IRAS PSC/FSC Combined Catalogue. WISE-based positional refinement and cross-correlation with SDSS DR14 produced 6,346 unique spectroscopic optical counterparts. Requiring complete 12, 25, 60, and 100 μm IRAS fluxes reduced the luminosity-analysis population to 6,270 objects.

Using the infrared-luminosity relations [21], objects above 10^{12} $L_{\odot}$ and within the adopted SDSS spectroscopic limit were selected. The resulting sample contains 107 ULIRGs and 7 HLIRGs, including 48 newly identified ULIRGs and 7 newly identified HLIRGs.

The subsequent spectral classification assigned activity types to 113 of the 114 objects. HII galaxies are the largest individual class, but the population also contains a substantial contribution from broad- and narrow-line Seyferts, LINERs, composites, NLAGN, BLL/QSO objects, emission-line systems, and absorption-line systems.

The results connect directly with the broader Thesis programme. Chapter 2 establishes the activity distribution of BIG galaxies, Chapter 3 constructs the much larger IRAS extragalactic population, Chapter 4 isolates and characterises the most luminous systems, and Chapter 5 provides the archival and virtual-observatory methodology through which these investigations can be extended.

5. ARMENIAN VIRTUAL OBSERVATORY AND BAO PLATE ARCHIVE: DATA INFRASTRUCTURE AND METHODOLOGICAL APPLICATIONS

5.1. Introduction: ArVO, the BAO Plate Archive and DFBS

Modern astronomical research increasingly depends on the integration of observations obtained by different instruments, at different wavelengths, and in different formats. For infrared-galaxy research, an infrared detection is only the starting point: its optical counterpart must be identified, its nature established, and its properties compared with spectroscopy and other available observations. Digital archives and Virtual Observatory systems provide a way to make heterogeneous datasets accessible and to support their systematic analysis.

The Armenian Virtual Observatory (ArVO) is a joint initiative of the Byurakan Astrophysical Observatory (BAO) and the Institute for Informatics and Automation Problems of the National Academy of Sciences of Armenia. The International Virtual Observatory Alliance (IVOA), founded in 2002, has supported the development of interoperability and technical standards among Virtual Observatory initiatives. ArVO has participated in this international environment since 2005 and provides a methodological context for connecting astronomical databases, analysis software, and computational resources [72–78].

A complementary resource is the BAO Plate Archive, which preserves photographic plates and films obtained during decades of observations with different telescopes and observing techniques. Digitization converts this historical material into image data that can be indexed, inspected, calibrated, and searched alongside modern catalogues. The archive therefore has potential value beyond preservation: it allows earlier observations to contribute to current source-identification and classification work [25]. A central design principle of the project is that the archive should not become a passive collection of scans. Its objectives include preservation, digitization, database construction, creation of an interactive sky map, scientific exploitation, and integration with broader astronomical databases. These goals are consistent with the archival-science approach discussed in the Astroplate literature [79–84].

The Digitized First Byurakan Survey (DFBS) is the best-developed example of this transformation. Based on the Markarian Survey (the First Byurakan Survey), it contains data for about 40 million extracted low-dispersion spectra for approximately 20 million astronomical objects and from 1,874 digitized plates. The survey covers approximately 17,000 square degrees and includes a wide variety of objects, among them ultraviolet-excess galaxies, stars, quasars, and other sources with distinguishable spectral features [26, 27].

The principal scientific content of this chapter is provided by two studies. Astsatryan et al. [28] investigate automated classification of DFBS low-dispersion spectra using supervised and unsupervised machine-learning methods [28]. Astsatryan et al. [29] extend this work to the classification of subtypes within selected object groups and introduce a cloud-based service supporting extraction, model training, fine-tuning, evaluation, and inference [29]. The following sections describe these two contributions in detail and then discuss their methodological relationship to the infrared-galaxy research presented in the preceding chapters.

5.2. The DFBS as a Digital Astronomical Resource

The DFBS is a joint project of BAO in collaboration with Università di Roma La Sapienza (Italy), Cornell University (Ithaca, N.Y., USA) and Hamburger Sternwarte (Hamburg, Germany). The DFBS is based on objective-prism observations, in which spectra of many objects are recorded across the same photographic plate. This observational design provides a large and diverse spectral dataset, but it also creates specific technical challenges. An individual plate may contain thousands of sources, and the appearance of each spectrum depends on source brightness, spectral energy distribution, photographic response, background conditions, and possible blending with nearby objects. Consequently, the extraction of a reliable spectral image from a plate is a distinct task that must be addressed before automated classification can be applied [26, 27].

The scientific information in a low-dispersion spectrum is carried by its overall shape, colour distribution, continuum, and recognizable emission or absorption features. These characteristics can provide initial discrimination among broad classes of objects, although they do not always permit a definitive physical classification. In particular, some active galactic nuclei and compact galaxies may be difficult to distinguish from other sources at low dispersion. The DFBS should

therefore be understood as a large archival resource for candidate identification and classification, not as a replacement for detailed follow-up spectroscopy where that is required [26, 27, 85].

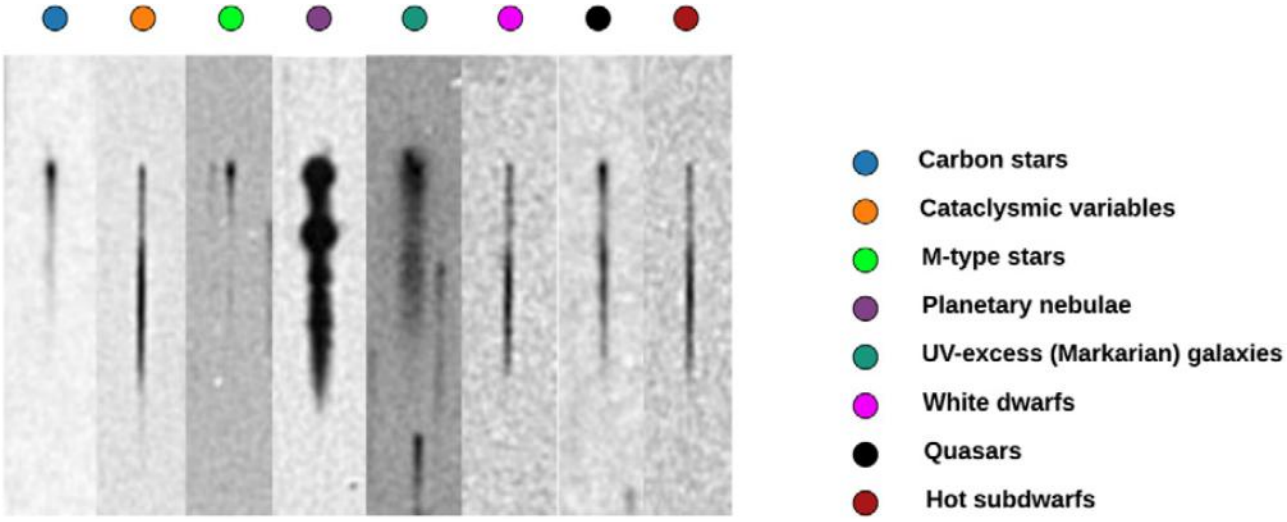


Figure 5.1. DFBS spectra for various types of objects.

The two core studies address complementary stages of this process. The first one develops an initial automated classification framework for several object groups represented in the DFBS. The second work refines both image extraction and classification, with emphasis on subclassifying Markarian galaxies, carbon stars, and subdwarfs. Table 5.1 summarizes their principal differences.

Table 5.1. Overview of the two core DFBS classification studies [28, 29].

Study	Main objective	Principal methods	Reported outcome
Astsatryan et al. [28]	Automated classification of broad groups of astronomical objects in DFBS low-dispersion spectra.	Three-step image preprocessing; supervised and unsupervised CNN-based approaches; augmentation of labeled images.	Supervised classification achieved about 87% overall accuracy on the evaluated dataset; the tested unsupervised approach performed substantially less well.
Astsatryan et al. [29]	Classification of subtypes within Markarian galaxies, carbon stars, and subdwarfs; provision of a reusable ML service.	Improved four-step extraction; revised CNN; focal loss and self-transfer learning; cloud-hosted notebooks and evaluation tools.	Reported 96% accuracy for the tested sub-object task and 93% for the evaluated broad-group task.

5.3. Automated Classification of Low-Dispersion Spectra

A study by Astsatryan et al. [28] addressed the challenge of classifying the large number of spectra available in the DFBS. Its objective was to develop a service capable of recognizing selected astronomical object types from extracted spectral images. The study considered ultraviolet-excess galaxies, including Markarian-type galaxies, quasars and compact galaxies, together with several

stellar and nebular categories. Automated classification is necessary because direct visual inspection alone does not scale to millions of spectra [28].

The work evaluates supervised and unsupervised learning strategies. In supervised learning, a dataset of labeled examples provides the reference classes used to train the model. The network learns image features associated with these classes and then predicts labels for previously unseen examples. In unsupervised learning, the algorithm attempts to identify structures and similarities in the input data without requiring a class label for every training example. The distinction is relevant to astronomical datasets, where labels may be incomplete and clustering can be used to explore similarities among objects [86–89].

The target classification categories include hot subdwarfs, carbon stars, ultraviolet-excess galaxies, M-type stars, and planetary nebulae. These categories differ in the shape of their continua, the presence and strength of spectral lines, and their overall colour. The paper [28] emphasises, however, that objects from distinct physical classes can sometimes resemble one another in low-dispersion spectra. The purpose of the automated system is therefore to provide a scalable first-level classification of the available imagery; the reported performance must be interpreted in relation to the classes and training examples actually evaluated [28].

5.4. Preprocessing and Spectral-Image Extraction

The method begins with image preprocessing designed to separate the spectral signal from the background. A Gaussian blur is applied to reduce high-frequency noise, following the general objective of image-denoising methods [90]. The image is then processed using adaptive mean thresholding, which estimates a threshold in local image regions rather than imposing one threshold over the entire plate. Pixels that differ sufficiently from the local background are treated as foreground candidates. This step creates a binary representation suitable for detecting the object boundary.

The next step determines the spatial boundary of the spectrum. Contour-tracing methods, including the algorithmic approaches described by Pavlidis, provide a way to determine an object's boundary and bounding rectangle [91, 92]. Astronomical coordinates, including right ascension and declination, are converted into image-pixel coordinates using information from

the relevant plate and FITS header; FITS is the standard format used to store and exchange this type of astronomical data [93]. When a target has known coordinates, these provide a starting position for identifying the corresponding spectral image. The object image can then be extracted from the original plate for subsequent classification.

This extraction stage is important because classification accuracy cannot be separated from the quality of the input images. If the extracted region contains background fragments, excludes part of the spectrum, or includes another object, the network receives misleading information. The challenge is particularly significant for faint objects and crowded fields, in which the distinction between source and background may be unclear. Image processing is therefore an integral part of the automated classification pipeline rather than a trivial preparation step.

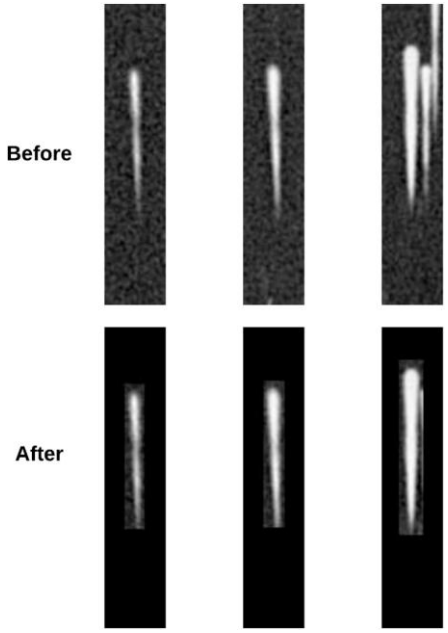


Figure 5.2. The results of the bounding-rectangle-finding algorithm for spectral images.

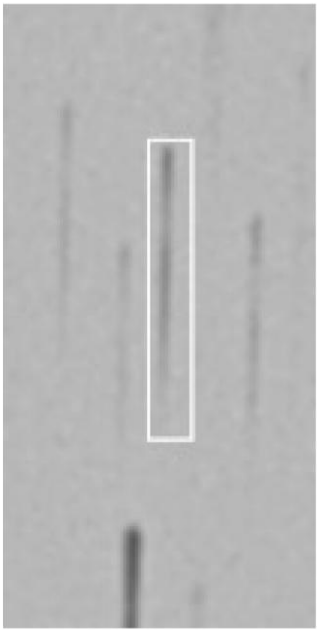


Figure 5.3. An illustration of the extraction of an object's 2D spectrum from the original plate image.

5.5. Supervised and Unsupervised Learning

For the supervised experiments, spectral images are represented in a standardized form and normalized before being supplied to the CNN. Normalization rescales pixel values into a consistent range, helping the network training procedure operate on comparable numerical inputs. Data augmentation, including transformations such as flipping, rotation, shifting, and

related image variations, is a commonly used strategy for increasing training-set diversity in astronomical image classification [94].

The CNN learns image features through convolutional layers and uses these features to generate a predicted class. Convolutional layers are suitable for this task because they identify local patterns in the spectral image while reusing the same filters across different image positions [86]. Pooling operations reduce the size of intermediate feature maps, and the final classification layers map the learned representation to the candidate categories. The eight-layer architecture and its TensorFlow/Keras implementation are those reported in this study [28].

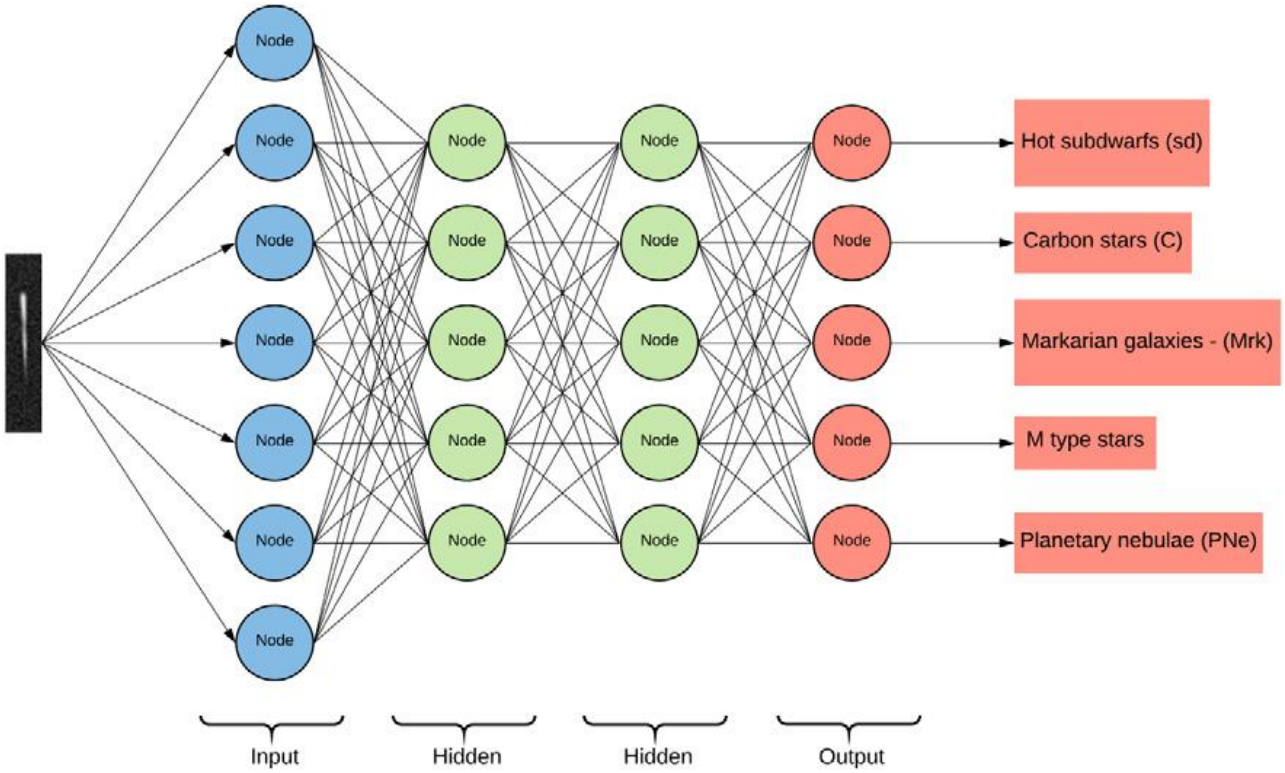


Figure 5.4. Supervised learning scheme.

The unsupervised experiments use an autoencoder to construct a reduced representation of the spectral images. The encoder compresses the input into a lower-dimensional feature space, while the decoder attempts to reconstruct the original input. The encoded features are then examined using clustering methods, including DBSCAN and K-means, to group similar patterns [87–89]. This provides a way to examine the structure of the dataset without specifying every class in advance.

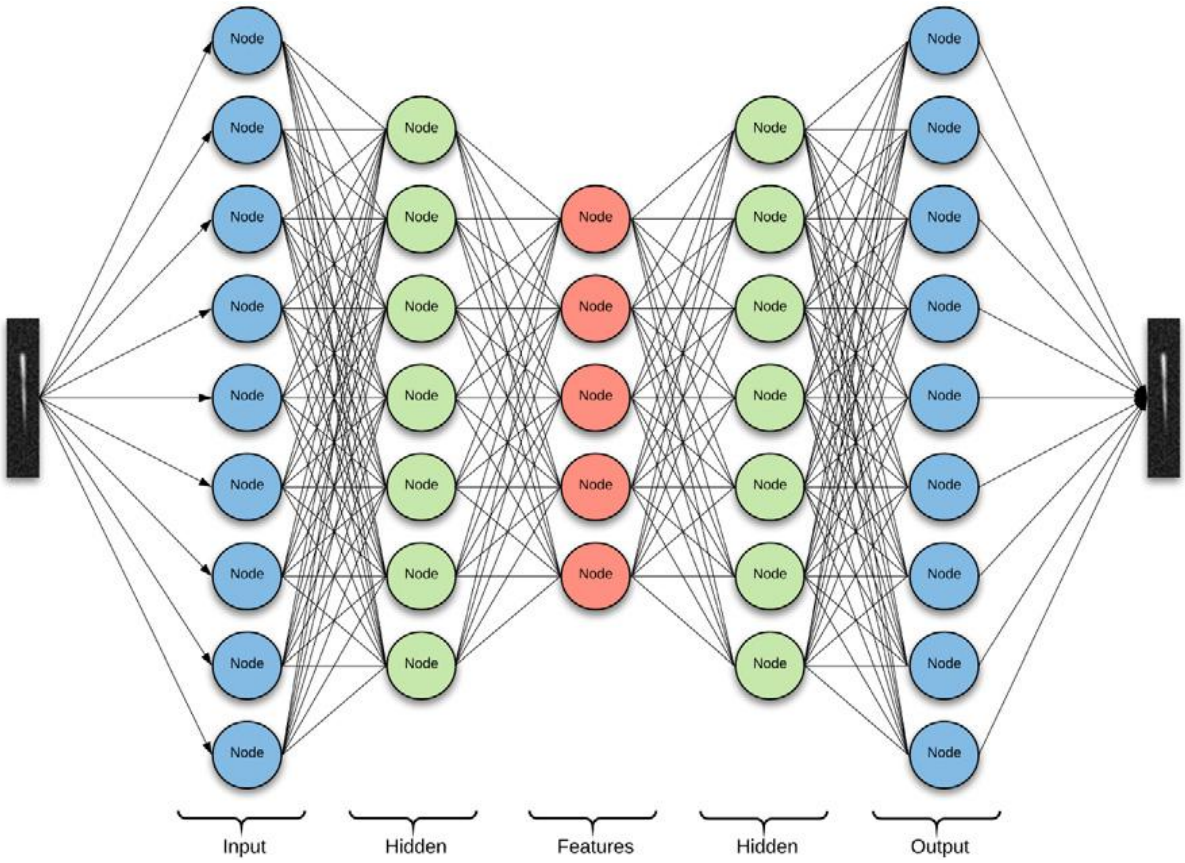


Figure 5.5. Unsupervised learning scheme.

5.6. Experimental Results and Limitations

Ref. [28] reports an overall accuracy of approximately 87% for the supervised CNN classification. The evaluated classes show different precision and recall values, illustrating that performance varies with the nature of the class and the size and quality of its labeled sample. Hot subdwarfs achieved reported precision, recall, and F1-score values of approximately 95%, 96%, and 96%, respectively. Carbon stars achieved approximately 86%, 84%, and 85%, while ultraviolet-excess galaxies achieved approximately 89%, 91%, and 90%. The M-type-star category had a notably lower recall of about 49%, corresponding to an F1-score of approximately 61%.

The reported scores for planetary nebulae were very high, but the corresponding evaluation sample was extremely small. The apparent perfect scores should therefore not be interpreted as proof that planetary nebulae can always be classified without error. More generally, the uneven results across the classes show why overall accuracy alone is not sufficient to describe a classifier's performance.

Table 5.2. Supervised ML model accuracy and datasets

Objects	Initial Train	Initial Test	Generated Train	Generated Test	Precision (%)	Recall (%)	F1-score (%)
Hot subdwarfs (sd)	550	157	2,750	785	95	96	96
Carbon stars (C)	331	94	1,650	470	86	84	85
Ultraviolet-excess galaxies	305	86	1,525	430	89	91	90
M-type stars	104	44	770	220	80	49	61
Planetary nebulae (PNe)	11	4	131	48	100	100	100
Total	**1,301**	**385**	**6,826**	**1,953**	—	—	—

For the unsupervised experiments, the study [28] gives an accuracy of approximately 38% and a mean squared error of 0.0012. The conclusion is that the supervised approach is more effective for the studied datasets. The supervised model was also applied to larger sets of extracted objects, including datasets of about 2.5 million (Table 5.3) and four million spectra (Table 5.4). The number of predicted objects in each target class varied with the confidence threshold: lowering the threshold increased the number of assigned labels, while applying a higher threshold retained fewer but more confident predictions. We used the resulting predictions and regression analyses to estimate that the full DFBS could contain at least six million carbon stars, three million hot subdwarfs, and one million Markarian galaxies. These are model-based estimates, not independent visual confirmations of every predicted object. Their reliability depends on how representative the training and inference sets are, how classification errors behave for the full survey, and how a prediction-confidence threshold is chosen. The reported variation in performance between object classes means that such estimates should be treated as indicative rather than as definitive census values.

Table 5.3. CNN model classification results for two and a half million objects

Accuracy	C Class	SD Class	PN Class	M Class	Mrk Class	Other
95%	403,556	154,623	6	6	29,336	1,929,075
90%	579,906	244,217	16	21	44,342	1,648,100
85%	697,535	322,233	278	62	61,782	1,434,712
80%	789,450	394,554	348	129	81,463	1,250,622

Table 5.4. CNN model classification results for about four million objects

Accuracy	C Class	SD Class	PN Class	M Class	Mrk Class	Other
95%	556,309	286,717	14	11	61,858	3,346,654
90%	829,577	440,737	38	46	91,670	2,889,531
85%	1,023,669	573,130	312	129	127,578	2,526,781
80%	1,181,754	695,987	392	265	167,559	2,205,606

The main contribution of the study is the demonstration that supervised CNNs can be applied to DFBS spectral images at a scale far beyond manual inspection. It also identifies challenges that motivate the later work: the noise and low contrast of photographic data, class imbalance, limited labeled samples, and the difficulty of obtaining reliable predictions for every object type. The need for improved extraction and finer classification therefore follows directly from the limitations of the initial approach [28].

5.7. Cloud-Based Machine Learning Service for Astronomical Sub-Object Classification

This study by Astsatryan et al. [29] extends the earlier DFBS classification framework by concentrating on sub-object classification. Rather than stopping at a broad category such as Markarian galaxy, carbon star, or subdwarf, the study attempts to distinguish subclasses within each group. This matters because objects sharing a broad spectral type can have different physical properties and scientific significance, and because an internally consistent classification can make large archival samples more useful for targeted research [29].

The study considers three principal object groups. Markarian galaxies are subdivided into Mrk AGN, Mrk SB (starburst), and Mrk Abs (absorption-line) categories. Carbon stars are subdivided into C-R, C-N, and C-H types, while subdwarfs are subdivided into sdO, sdB, and sdA types. The experimental evaluation, however, focuses on the subclasses for which sufficient labeled examples were available [29].

An important observational issue is that the apparent shape of a photographic spectrum depends on its brightness and the response of the photographic emulsion. The study therefore discusses

magnitude-dependent differences in the appearance of spectral traces. Very faint objects may resemble plate defects or background fragments, while bright sources can be overexposed. The spectra were grouped by magnitude and the authors note that reliable classification is expected mainly within intermediate brightness ranges; objects near the faint limit or strongly overexposed are more difficult to classify [29].

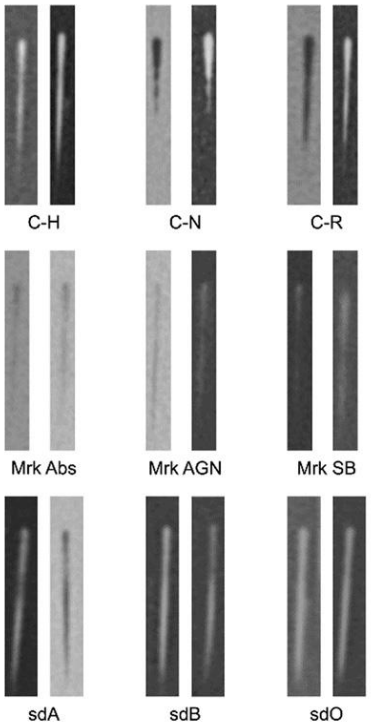


Figure 5.6. Examples of the subtypes of objects from the three most common object groups in the DFBS.

5.8. Revised Extraction, CNN Architecture and Transfer Learning

The study [29] improves the extraction process by developing a four-step procedure that builds on the earlier pipeline. A key change is the replacement of adaptive mean thresholding with adaptive Gaussian thresholding. The latter is reported to be more robust against noise and bright background artefacts, although it is computationally more expensive. Additional size constraints are used to reject extracted regions that are too small to represent valid spectral images. A fallback procedure is also introduced for cases where the primary extraction fails, allowing the system to examine and filter alternative contours around a known starting point. The use of Gaussian-based processing and contour-based fallback extraction is consistent with the broader image-processing methods [29, 90, 95].

The revised procedure can extract up to 30% more objects than the previous method when applied to the same coordinate-labeled dataset. This improvement is important because an enhanced extraction pipeline can increase the number of usable training examples and inference candidates before the classifier itself is changed. It also reinforces the dependence of classification quality on the reliability of preprocessing and extraction.

The revised CNN accepts images of 160 × 50 pixels and uses four convolutional blocks with 32, 64, 128, and 256 filters. The blocks combine convolution, batch normalization, and pooling. The service also includes architectures such as ResNet and MobileNetV2, which are established CNN approaches for image recognition [96, 97]. The training strategy addresses class imbalance through focal loss [98], uses the Adam optimizer [99], and applies data augmentation to increase variation in the training inputs [87]. The specific architecture and training configuration were evaluated, as described in our work [29].

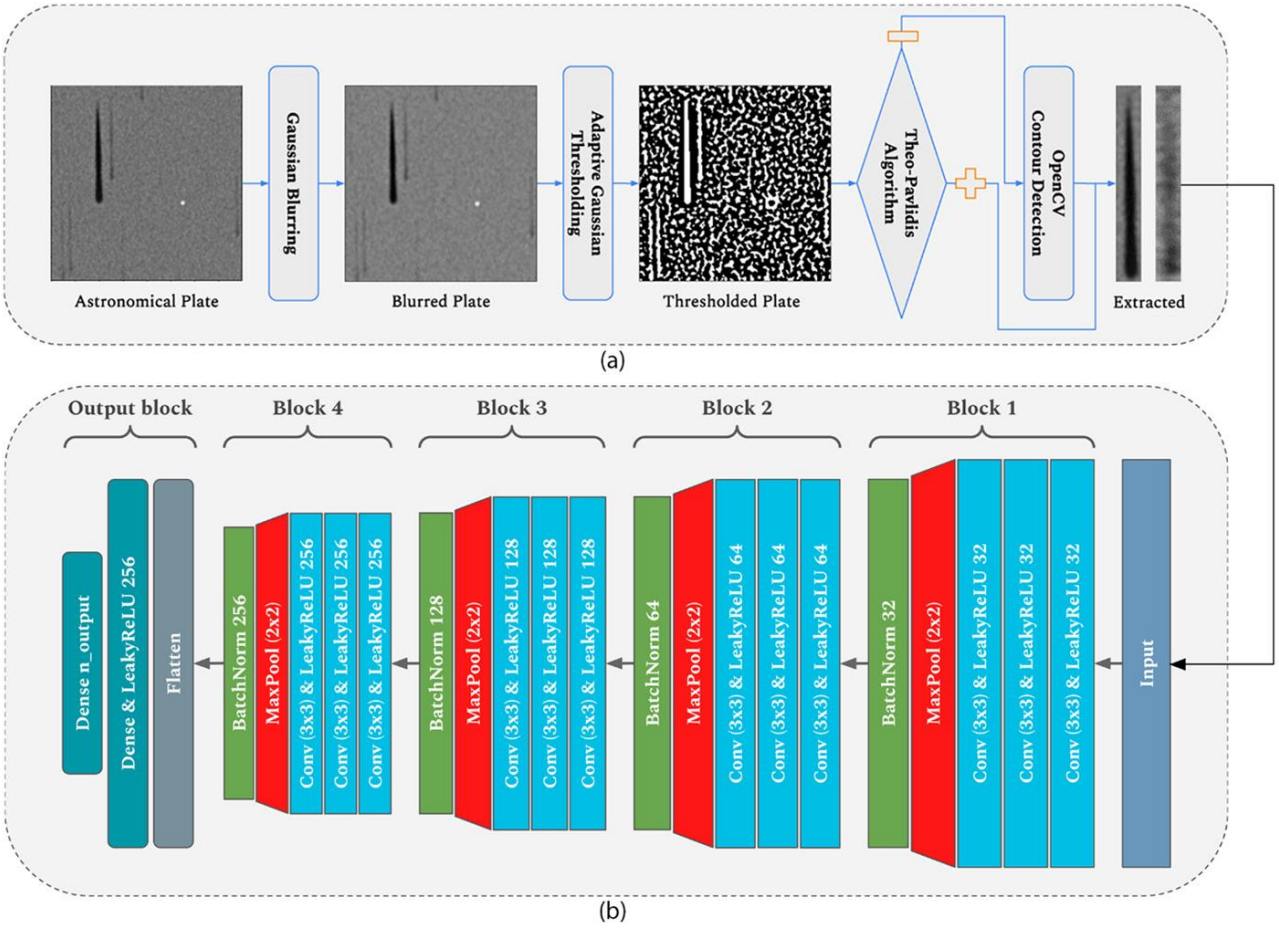


Figure 5.7. (a) four-step image processing, (b) the proposed CNN architecture.

A further element is self-transfer learning. The model is first trained on a broader classification task and is then fine-tuned on the smaller dataset containing the target subtypes. This allows the model to reuse features learned from a larger collection of spectra and adapt them to a more specific classification problem. Transfer learning is particularly relevant when the target task has fewer labeled examples than the broader task used for initial training [100].

5.9. Cloud-Based Service and Workflow

The study addresses not only model performance but also how the classification procedure can be made usable by other researchers. The proposed cloud-based service relies on Google Colaboratory and Jupyter notebooks, with the training and evaluation code maintained in a GitHub repository. Published analysis of Google Colaboratory describes its use as an accessible environment for accelerating deep-learning applications [101], while the associated Computational Astrophysics library provides the implementation resources [102].

The service is organized around the principal stages of a machine-learning workflow: data loading and preparation, training, fine-tuning, testing, inference, and visualization. Users can set parameters such as image dimensions, batch sizes, and model checkpoints. Several network architectures, including ResNet and MobileNetV2, are available for experiments [96, 97], and saved checkpoints can be used for testing or fine-tuning. Evaluation is supported by classification reports and TensorBoard-based displays of training accuracy and loss.

An important feature of the service is that it is not designed solely to execute one fixed pretrained model. It provides an environment for researchers to run their own training and evaluation pipelines and to apply trained models to new data. The study also describes an extraction mode that searches plates for unlabeled objects, allowing the framework to be used for exploratory analysis of plate content. In this mode, the absence of object identifiers limits the interpretation of individual detections, but the results can still provide information about the distribution of candidate spectra across the examined material [29].

By integrating image extraction, learning, evaluation, and inference in a single environment, the service represents a practical extension of the earlier DFBS classification work. It moves the framework closer to a reusable analysis tool for the astronomical community and makes the

computational workflow more transparent and repeatable. At the same time, the service cannot eliminate the need for suitable training labels, sufficient examples for rare classes, and independent checks of scientific classifications.

5.10. Evaluation, Large-Scale Inference and Limitations

Our study [29] evaluates the revised framework on a sub-object dataset containing 2,107 training examples and 405 validation examples, and on a broad-group dataset with 1,478 training and 258 validation examples. For sub-object classification, the tested classes are C-H carbon stars, Markarian starburst galaxies, and sdB subdwarfs. The reported overall accuracy is 96%. Class-specific F1-scores are approximately 0.95 for C-H stars, 0.95 for Markarian starburst galaxies, and 0.97 for sdB subdwarfs. For broad-group classification, there is an accuracy of 93% for the revised framework, compared with 88% for the earlier framework. The individual metrics again vary by class. For example, the reported F1-score is about 0.97 for the Markarian-galaxy class, while the M-type-star class has a lower recall and F1-score. The planetary-nebula results are very high, but they are based on a small number of examples. The metrics therefore show both the potential of the improved method and the continuing dependence of performance on class-specific sample sizes and spectral characteristics.

Table 5.5. Comparison of classification reports for two frameworks on the sub-objects' dataset. Support denotes the number of samples in the corresponding class (train + validation).

Class	Support	Precision	Recall	F1-score
C-H	626 + 117	0.89 / 0.93	0.91 / 0.97	0.90 / 0.95
Mrk SB	664 + 132	0.94 / 0.95	0.89 / 0.95	0.91 / 0.95
sdB	817 + 156	0.94 / 0.98	0.96 / 0.96	0.95 / 0.97
Accuracy	2,107 + 405	—	—	0.93 / 0.96
Macro avg	2,107 + 405	0.92 / 0.96	0.92 / 0.96	0.92 / 0.96
Weighted avg	2,107 + 405	0.93 / 0.96	0.93 / 0.96	0.93 / 0.96

Table 5.6. Comparison of classification reports for two frameworks on the group classification dataset.

Class	Support	Precision	Recall	F1-score
C	362 + 63	0.82 / 0.87	0.89 / 0.87	0.85 / 0.87
M	169 + 29	0.70 / 0.74	0.48 / 0.69	0.57 / 0.71
Mrk	333 + 58	0.91 / 0.94	0.88 / 1.00	0.89 / 0.97
PN	13 + 2	1.00 / 1.00	1.00 / 1.00	1.00 / 1.00
sd	601 + 106	0.93 / 1.00	0.98 / 0.98	0.95 / 0.99
Accuracy	1,478 + 258	—	—	0.88 / 0.93
Macro avg	1,478 + 258	0.87 / 0.91	0.85 / 0.91	0.86 / 0.91
Weighted avg	1,478 + 258	0.87 / 0.93	0.88 / 0.93	0.87 / 0.93

The study also tests the trained classifier on inference datasets containing approximately one million (Table 5.7) and four million (Table 5.8) extracted objects, assembled from 65 and 266 plates, respectively. The predicted numbers change with the confidence threshold. For the four-million-object dataset, a threshold of 0.95 produced predictions of 37,355 C-H objects, 48,695 Markarian starburst galaxies, and 57,278 sdB subdwarfs. At a threshold of 0.80, the corresponding counts increased to 204,382, 184,918, and 133,209. These values demonstrate the effect of confidence selection: a high threshold reduces the number of predicted members of each target subclass but retains only the model's more confident assignments. The large-scale counts must not be interpreted as a verified census of the full DFBS. They depend on the extraction method, the plate sample used for inference, the model's performance on data that may differ from the training set, and the chosen threshold.

Table 5.7. Model inference results on one million astronomical objects

Threshold	C-H	Mrk SB	sdB	Other
0.95	10,557	10,237	15,084	964,123
0.90	23,823	19,044	21,691	935,443
0.85	38,633	29,163	28,265	903,940
0.80	56,026	40,609	35,231	868,135

Table 5.8. Model inference results on four million astronomical objects

Threshold	C-H	Mrk SB	sdB	Other
0.95	37,355	48,695	57,278	3,856,673
0.90	85,379	89,872	82,183	3,742,567
0.85	139,925	134,719	106,725	3,618,632
0.80	204,382	184,918	133,209	3,477,492

Taken together, the results show that improved extraction, class-aware training, and the provision of a cloud-based workflow can increase the practical utility of an archival classification system. At the same time, high accuracy on a selected evaluation dataset is not equivalent to guaranteed performance across every plate, magnitude range, or spectral class. Independent validation, careful attention to imbalanced categories, and preservation of confidence information remain essential when using automated classifications as the basis for astronomical samples [28, 29].

Table 5.9. Selected performance and inference results reported in the two core studies.

Study / task	Evaluation data	Reported result	Interpretation
2021: supervised broad-group classification [28]	Five tested classes; initial labeled samples were augmented for training and testing.	Overall accuracy approximately 87%. Class results varied, with lower recall for M-type stars.	Demonstrates scale-up potential but also class-dependent performance.
2021: unsupervised classification [28]	The evaluated low-dispersion spectral-image datasets.	Accuracy approximately 38%; reported MSE 0.0012.	The tested unsupervised approach was less effective than supervised learning for this task.
2024: sub-object classification [29]	2,107 training and 405 validation samples; C-H, Mrk SB, and sdB classes.	Overall accuracy 96%; class F1-scores approximately 0.95–0.97.	Strong performance for the three evaluated subclasses; not evidence for all possible subtypes.
2024: broad-group classification [29]	1,478 training and 258 validation samples.	Accuracy 93%, compared with 88% for the previous framework in the paper's comparison.	Indicates an improvement under the evaluated comparison setting.
2024: large-scale inference [29]	Datasets of about 1 million and 4 million extracted objects.	Predicted counts varied substantially as confidence thresholds changed from 0.95 to 0.80.	Outputs are model predictions, not individually confirmed population counts.

5.11. Comparative Discussion and Relevance to Infrared-Galaxy Research

The two studies form a clear methodological sequence. Our study [28] establishes the feasibility of using CNNs to classify low-dispersion DFBS spectral images and evaluates supervised and unsupervised approaches [28]. Our next work [29] builds on that foundation by improving the extraction pipeline, focusing on subtype classification, addressing class imbalance, applying self-transfer learning, and providing a cloud-based service for training and inference [29]. The later work therefore expands both the technical capabilities of the classifier and the means through which researchers can use it.

For the present Thesis, the principal connection is methodological rather than a claim that the described models have already classified the complete IRAS PSC/FSC Combined Catalogue or the 114-object ULIRG/HLIRG sample. The DFBS is an archival optical resource, while the main Thesis samples originate from infrared selection and are subsequently investigated through optical identification, spectroscopy, and multiwavelength information. A future use of the DFBS workflow could help locate archival spectral material for selected sources, support the identification of candidate active galaxies, or provide additional context for optical counterparts of infrared detections. These applications would require explicit cross-matching, validation of the resulting associations, and careful consideration of the spectral information available for each object.

The two studies also complement the emphasis in earlier chapters on selection criteria and classification limitations. Chapter 2 demonstrates the value of source-level spectroscopy for BIG objects; Chapter 3 shows the importance of controlled catalogue cross-correlation when constructing a large extragalactic sample; and Chapter 4 uses spectroscopic classifications to investigate the high-luminosity infrared population. The DFBS machine-learning work addresses a related problem at the level of digitized plate images: how to make the extraction and classification of very large astronomical datasets more systematic and reproducible.

Several limitations should be retained in any scientific use of these techniques. First, training data may not represent the full range of brightness, plate conditions, or spectral diversity in the survey. Second, rare subclasses can be underrepresented and confused with more common classes. Third, extraction errors may propagate into the classifier, while duplicated observations across plates

may produce multiple predictions for the same object. Fourth, confidence thresholds change the balance between the number of selected candidates and the reliability expected of their labels. Finally, low-dispersion spectral appearance alone cannot always distinguish physical classes that require line measurements or higher-quality follow-up spectra. These limitations are reflected in the results and scope of the two evaluated approaches [28, 29].

Accordingly, the most appropriate role for the DFBS classification methods within the broader Thesis programme is as enabling infrastructure. They can assist with exploratory classification and candidate selection and can guide the prioritization of objects for more detailed investigation. The scientific conclusions concerning infrared luminosity, AGN activity, star formation, and galaxy interactions must continue to be grounded in the relevant infrared measurements, reliable counterpart associations, spectroscopy, and explicit treatment of selection effects. The DFBS methods may support this work, but they do not replace the observational evidence required for each individual scientific claim.

5.12. Summary

This chapter has examined two connected studies that develop automated classification methods for the Digitized First Byurakan Survey. The study [28] introduced image preprocessing, spectral extraction, and supervised and unsupervised CNN approaches for classifying selected DFBS object groups, reporting approximately 87% overall accuracy for the supervised model and lower performance for the tested unsupervised approach [28].

The study [29] extended the framework to sub-object classification and a cloud-based machine-learning service. Its main contributions include an improved four-step extraction pipeline, a revised CNN architecture, data augmentation and focal loss, self-transfer learning, and notebooks supporting training, evaluation, fine-tuning, and large-scale inference. The paper reported 96% accuracy for the evaluated sub-object task and 93% for broad-group classification in its comparison framework, while also showing that inferred population counts depend strongly on confidence thresholds and require cautious interpretation [29].

The two studies demonstrate how the DFBS archive can be developed from a digitized historical survey into an increasingly systematic resource for large-scale astronomical analysis. Their

relevance to infrared-galaxy research lies in the potential to support optical-source identification, candidate classification, and archival follow-up. They provide a methodological complement to the BIG, large IRAS extragalactic, and ULIRG/HLIRG investigations developed in the preceding chapters, while leaving direct cross-matching and validation for specific infrared samples as tasks for future application.

6. OVERALL DISCUSSION AND SCIENTIFIC SYNTHESIS

6.1. Scope of the Research Programme

This Thesis addresses a central problem in infrared-galaxy research: an infrared detection identifies a source of energy, but does not by itself establish the nature of the emitting object, its distance, the mechanism powering its emission, or the contribution of possible companions. A reliable interpretation requires several connected types of evidence: positional identification, optical or multiwavelength counterpart information, spectroscopy, infrared photometry, and an assessment of the source environment. The studies presented in Chapters 2–5 approach these connected tasks at different scales and should be read as complementary components of one research programme.

Chapter 2 examines the Byurakan-IRAS Galaxy (BIG) sample through optical identification and spectroscopy. Chapter 3 extends the work to the 345,163-source IRAS PSC/FSC Combined Catalogue and develops a preliminary extragalactic classification through hierarchical catalogue cross-correlation. Chapter 4 uses refined positions, SDSS spectroscopy, and four-band IRAS photometry to select and classify a high-luminosity sample of ULIRGs and HLIRGs. Chapter 5 provides the archival and computational context, focusing on the DFBS and two studies of automated classification of low-dispersion spectral images.

The common methodological sequence is therefore: infrared selection, positional refinement, counterpart identification, spectral characterization, luminosity determination, activity classification, population-level interpretation, and—where suitable data exist—archival extension. The main contribution of the Thesis is the connection of these stages while keeping their different selection functions and evidential limits explicit.

6.2. Infrared Selection as the Starting Point

The IRAS PSC/FSC Combined Catalogue provides the common parent population for the large-scale identification work and the high-luminosity selection. The catalogue contains 345,163 sources and was assembled through associations between the Point Source Catalogue and Faint Source Catalogue, followed by positional improvement using additional infrared catalogues [6,

20]. Its four IRAS bands, especially the 60 and 100 μm measurements, provide the mid- and far-infrared information needed for identifying and characterizing dust-emitting systems [3, 4, 6].

Infrared and optical catalogues have different positional accuracies, IR being much worse. A nearest-neighbour association alone can therefore assign the wrong optical counterpart to an IRAS detection. The catalogue-construction study addresses this through source-dependent search criteria and a hierarchy of known-object catalogues, optical morphology, radio information, and Gaia astrometry [20]. The BIG programme provides a complementary source-level approach in which candidate identification is supported by inspection of low-dispersion spectra and follow-up spectroscopy [19].

This is an astrophysical rather than merely technical issue. An incorrect counterpart can lead to an incorrect redshift, distance, luminosity, and activity classification. Positional refinement and explicit treatment of ambiguous associations are therefore integral to the reliability of the scientific conclusions drawn from infrared-selected samples.

6.3. BIG as the Detailed Source-Level Component

The BIG dataset provides detailed spectroscopic information for a selected set of IRAS-associated objects. The revised sample used in this Thesis contains 257 spectra: 56 from the BAO 2.6-m telescope, 54 from the 6-m telescope of the Special Astrophysical Observatory, 64 from the 1.93-m telescope of the Observatoire de Haute-Provence, and 83 from SDSS. In cases where multiple spectra of the same object were available from different telescopes or observations, the spectrum with the best overall quality was selected for the final dataset. The reported redshift range is $0.0092 \leq z \leq 0.1997$, while the absolute magnitudes span approximately –24.46 to –16.81 [2, 19]. Because the records include repeated spectral observations and systems with more than one optical component, the number of spectra or catalogue records should not automatically be interpreted as the number of unique IRAS sources.

In the detailed classification, 149 records are assigned to HII/star-forming galaxies. The remaining records include 17 Seyfert galaxies, 12 LINERs, 22 composite objects, 4 HII/LINER, 2 HII/Seyfert, 13 AGN (without detailed activity type), 13 normal galaxies, 21 emission-line objects without a more specific class, 3 absorption-line systems, and 1 unclassified object. Grouped for synthesis,

the categories comprise 42 AGN/Seyfert/LINER records, 28 composite or mixed-class records, and the remaining emission-line, normal, absorption, and unclassified categories [2, 19].

The predominance of HII/star-forming spectra is a meaningful property of this selected BIG dataset, but it should not be generalized to all infrared galaxies: the sample reflects its particular identification history, spectroscopic coverage, and observational accessibility. The accompanying AGN, Seyfert, LINER, and composite classes demonstrate that strong infrared emission can be associated with more than one physical mechanism [2, 19].

6.4. Spectroscopic Classification as a Physical Diagnostic

Spectroscopy supplies information that infrared luminosity alone cannot provide. Emission-line strengths, line ratios, line widths, and continuum features help distinguish HII-region emission from nuclear activity and separate broad-line from narrow-line AGN. In the BIG sample, spectroscopy also provides redshifts and supports the derivation of distances and absolute magnitudes. In the high-luminosity sample, the quality and profile of individual SDSS lines are especially important because broad permitted components can overlap narrow components, while weak diagnostic lines may be affected by noise [19, 23, 63–71].

The classification of the 114-object ULIRG/HLIRG sample assigned activity types to 113 objects, with one remaining unclassified because of poor spectral quality. The source study used visual inspection alongside line-width and line-ratio criteria where the available spectra allowed it; no single line ratio is sufficient to replace examination of the full spectrum. The study also compared its results with the activity information available from NED: 44 objects had database activity information, but only six were judged by this study to be correctly or completely classified [23, 62]. This comparison illustrates the value of direct spectral examination, particularly for unusual or highly luminous objects, without implying that every new classification is free of uncertainty.

The broader lesson is that classification labels should be interpreted with their data quality and decision rules in view. A preliminary catalogue label, a visual spectral classification, and a subtype supported by several diagnostic lines do not necessarily have the same evidential strength. Preserving uncertain and mixed categories is preferable to forcing every source into an overconfident physical interpretation.

6.5. From the BIG Population to ULIRGs and HLIRGs

The BIG sample and the ULIRG/HLIRG sample answer related but different questions. BIG offers detailed source-level information for a selected IRAS-associated population. The high-luminosity study isolates the extreme end of infrared luminosity from the same Combined Catalogue and applies stricter requirements for optical spectroscopy and complete infrared photometry [19, 22, 23]. The two samples are therefore complementary, but their raw class fractions should not be compared as though the samples had identical selection functions.

The high-luminosity selection began with 6,346 sources having unique SDSS spectroscopic counterparts after positional refinement. Of these, 6,270 had flux measurements in all four IRAS bands required by the adopted luminosity calculation. Applying the infrared-luminosity threshold $L_{IR} > 10^{12} L_{\odot}$ together with the study's SDSS spectroscopic magnitude criterion, $r < 17.77$, produced 114 objects: 107 ULIRGs and 7 HLIRGs [21, 22].

In the comparison with NED reported by the source study, 59 of the 107 ULIRGs were already known, leaving 48 newly identified ULIRGs. All seven objects in the selected sample were newly recognized as HLIRGs in that study, although two had previously been classified as ULIRGs. This distinction matters: a source can be newly assigned to a more luminous classification without being a previously unknown astronomical object. The sample's contribution is the combination of a defined selection pathway, complete four-band IRAS measurements for the luminosity calculation, and spectroscopic characterization of the resulting high-luminosity population [22, 23].

6.6. Infrared Luminosity and Its Observational Interpretation

The luminosity analysis in Chapter 4 uses the IRAS flux densities and the adopted distance relations based on Duc et al. (1997) [21]. The total infrared luminosity combines measurements from the 12, 25, 60, and 100 μm bands, whereas the far-infrared estimate emphasises the 60 and 100 μm measurements. Requiring all four IRAS flux densities therefore defines the photometric completeness of the 6,270-object parent subset used for the high-luminosity selection. The final count of 114 is a result of the luminosity and optical-spectroscopic criteria applied to this subset, not a count of every ULIRG or HLIRG in the full Combined Catalogue [21, 22].

The luminosity–redshift diagram (Figure 4.1) must be read in the context of these selection limits. The apparent concentration of ULIRGs around $\log(L_{IR}/L_{\odot}) \approx 12.0–12.8$ and redshifts of roughly $z < 0.5$ describes the selected sample. Because the optical spectroscopic selection has a magnitude limit, more distant objects generally need to be brighter to remain observable and enter the sample. The resulting trend is therefore shaped by detectability and sample construction as well as by intrinsic luminosity; it is not, by itself, evidence for luminosity evolution [22].

This distinction is central to all population-level conclusions in the Thesis. Luminosity is a physical property derived from measured fluxes and adopted distances, but which objects enter a sample is governed by additional photometric, positional, and spectroscopic criteria. Claims about the abundance or evolution of ULIRGs and HLIRGs require a fuller treatment of completeness and selection effects than the 114-object sample alone provides.

6.7. Star Formation, AGN, and the Activity Sequence

The activity distribution of the high-luminosity sample shows that high infrared luminosity does not correspond to one unique excitation mechanism. Under the grouping used by the spectroscopic study, the 114 objects include 36 HII galaxies, 28 broad-line Seyferts, 14 LINERs, 11 composite objects, 9 narrow-line AGN, 6 narrow-line Seyferts, 3 BLL/QSO objects, 3 emission-line objects, 3 absorption-line objects, and 1 unclassified source [23]. This mixture supports the interpretation that both star formation and nuclear activity contribute to the observed properties of highly luminous infrared galaxies, with some sources showing composite or ambiguous signatures.

Compared with the HII-dominated BIG classification, the high-luminosity sample contains a more prominent contribution from broad-line and other AGN-related classes (i.e. high optical luminosity objects). This is an important observational difference between the selected samples, but not a controlled measurement of a single evolutionary sequence: the selection criteria, luminosity ranges, optical limits, and sample sizes differ. The results are consistent with a greater prominence of AGN-related activity among these selected high-luminosity objects, but do not establish that starburst activity necessarily triggers an AGN or that one phase must follow the other [2, 19, 22, 23].

The SDSS u–g and g–r colour analysis provides an additional observational dimension. The reported means suggest a broad tendency for some of the more AGN-dominated groups to be bluer and to have higher mean infrared luminosities than lower-activity groups. The trend is not strictly monotonic across every class, and the number of objects in some classes is small. Redshift, dust extinction, the optical magnitude limit, and the selection of objects with usable spectra can all affect the observed colours. The colour diagrams should therefore be interpreted as properties of the selected sample and possible indicators for follow-up, not as a universal one-parameter activity or evolutionary track [23].

6.8. Interactions and the Nature of the Infrared Source

An infrared detection does not always correspond to one isolated optical galaxy. A single IRAS source can have more than one plausible optical counterpart, and interacting or multiple systems may contribute jointly to the measured infrared flux. The broader BIG review reports 361 groups containing 907 counterparts, with at least two spectroscopically observed members in 31 groups [2]. These catalogue-level groupings establish the importance of treating multiplicity as part of the identification problem, while individual systems still require evidence to distinguish physical association from chance projection.

This issue matters directly for luminosity interpretation. If the infrared flux of a multi-component system is assigned entirely to one galaxy, the inferred luminosity of that component may be overstated; if the emission is shared, the physical unit relevant to the infrared measurement may be the interacting system rather than an individual member. The BIG and high-luminosity studies therefore motivate a systematic comparison of isolated galaxies, pairs, multiples, and mergers, with particular care in connecting infrared positions to optical morphology and spectroscopic redshifts [2, 19, 22, 36].

The material assembled for this Thesis supports the conclusion that interacting and multiple structures are important features of infrared-galaxy research. It does not establish a universal interaction-driven evolutionary sequence, nor does the presence of a companion alone prove that an interaction caused the observed luminosity or nuclear activity. Those questions require controlled samples and direct physical evidence.

6.9. The Large Extragalactic Sample and the Problem of Completeness

The classification of the IRAS PSC/FSC Combined Catalogue provides the broad population framework for the source-level and high-luminosity studies. At the stage reported in Chapter 3, the 345,163 catalogue entries were divided into 183,168 star candidates, 80,998 unknown sources, 68,940 galaxy candidates, and 12,057 confirmed or genuine galaxies or QSOs [6, 20]. These categories distinguish levels of identification confidence rather than final, immutable physical identities.

The method combines positive identification from known extragalactic catalogues with optical morphology, radio information, and Gaia-based stellar rejection. The results demonstrate the value of using several independent datasets, but catalogue matches can overlap and individual cross-correlation match counts cannot be added as if each represented a disjoint set of sources. Likewise, the remaining unknown category can include faint galaxies, other astronomical sources, unreliable detections, or contaminated measurements; it should not be interpreted as an additional galaxy population without further verification [20].

The sample is therefore preliminary, but valuable. It separates secure identifications from candidates that require further examination and supplies a large parent population for future luminosity and morphology studies. Before deriving population frequencies or luminosity functions, however, the catalogue will need further source validation, consistent treatment of ambiguous matches, and an explicit assessment of completeness and selection effects. Nevertheless, the sample of some 80,997 extragalactic objects may serve as the biggest sample of IR galaxies and QSOs for further statistical analysis and studies [20].

6.10. Role of ArVO and the BAO Archive in the Overall Methodology

The Armenian Virtual Observatory (ArVO), the BAO Plate Archive, and the Digitized First Byurakan Survey (DFBS) provide the archival and computational context for extending the use of historical observations. The BAO Plate Archive project reports approximately 37,000 photographic plates and films, of which 30,121 had been digitized at the reported stage. The DFBS is based on 1,874 plates and contains approximately 20 million objects and about 40 million extracted low-dispersion spectra [25–27]. These resources complement modern IRAS, WISE,

AKARI, SDSS, Gaia, and radio catalogues by retaining historical optical information in digital form, including thousands of objects related to this study: Markarian galaxies and IR galaxies.

The two machine-learning studies discussed in Chapter 5 address different stages of automated spectral-image analysis. The study [28] applied supervised and unsupervised CNN methods to selected DFBS object classes. It reported approximately 87% overall accuracy for its supervised model and substantially lower performance for the tested unsupervised approach, with class-dependent performance and lower recall for M-type stars [28]. The study [29] extended the work to selected subtypes and a cloud-based service. It reported 96% accuracy for the evaluated sub-object task and 93% for broad-group classification under its comparison setting; the study also showed that large-scale predicted counts varied with the adopted confidence threshold [29]. These values describe the evaluated datasets and tasks, not guaranteed performance across every DFBS plate or class.

The role of these methods in this Thesis is methodological. They demonstrate a way to extract, classify, and prioritize very large sets of archival spectral images, but the studies have not directly classified the complete IRAS PSC/FSC Combined Catalogue or the 114-object ULIRG/HLIRG sample. Any future connection would require explicit positional cross-matching, attention to duplicate observations, validation of the associations, and independent evaluation on spectra that represent the intended target population. Automated classification can assist candidate discovery and prioritization, but does not replace the measurements and spectral evidence needed for a physical conclusion [28, 29].

6.11. Unified Methodological Framework

The main research components can be understood as stages that reduce different uncertainties. The BIG study supplies detailed spectroscopy for selected IRAS-associated records; the larger catalogue study uses multiple catalogues to classify a much larger source population; the ULIRG/HLIRG work applies stricter photometric and spectroscopic requirements to a high-luminosity subset; and ArVO/DFBS methods support the preservation, retrieval, extraction, and potential scalable analysis of archival observations. The components are related, but they do not share identical samples or selection functions.

Table 6.1 summarizes the main evidence and contribution of each component, together with the most important qualification needed when interpreting its results.

Table 6.1. Integrated synthesis of the principal research components.

Research component	Data and method	Contribution and interpretive limit
BIG sample	IRAS identification and optical spectroscopy; 257 medium-resolution spectra from four facilities.	Redshifts, optical magnitudes, activity types, and evidence for pairs and multiple systems. The sample is selected and combines heterogeneous observations.
Large IRAS extragalactic sample	345,163-source Combined Catalogue; hierarchical catalogue matching, optical morphology, radio data, and Gaia astrometry.	12,057 confirmed/genuine galaxies or QSOs, 68,940 galaxy candidates, 183,168 star candidates, and 80,998 unknown sources at the reported stage. The classification remains preliminary.
ULIRG/HLIRG sample	WISE-based positional refinement, SDSS spectroscopy, four-band IRAS photometry, and infrared-luminosity selection.	114 selected objects: 107 ULIRGs and 7 HLIRGs. Results depend on spectroscopic, photometric, and magnitude limits and are not a volume-complete census.
ArVO, BAO archive, and DFBS classification	Digitized photographic plates and low-dispersion spectra; automated extraction, CNN classification, and cloud-based workflows.	Supports preservation and scalable exploration of archival data. The models of [28, 29] have not been directly applied to the complete IRAS sample in the work presented here.

6.12. Methodological Strengths and Limitations

A principal strength of the research programme is the use of complementary information. Infrared photometry identifies dust-emitting sources, higher-precision catalogue positions improve counterpart associations, spectroscopy supplies redshift and activity diagnostics, radio data provide independent extragalactic evidence, Gaia astrometry helps identify stellar

contaminants, and archival images and spectra extend the available observational history. Agreement among independent data sources strengthens identification, while explicit candidate and unknown categories retain uncertainty instead of masking it.

A second strength is the combination of scales. The 257-spectrum BIG dataset permits detailed source-level examination; the 345,163-source Combined Catalogue provides a broad parent population; the 114-object ULIRG/HLIRG sample isolates the most luminous systems under defined criteria; and the DFBS studies address scalable archival image extraction and classification. This combination connects object-level diagnostics with catalogue-level population analysis and with methods intended to exploit historical data more efficiently.

The limitations follow from the same design. BIG combines spectra obtained with different facilities and is not a homogeneous volume-limited survey. The large IRAS extragalactic classification is preliminary and retains a substantial unknown category. The ULIRG/HLIRG sample is constrained by the availability of SDSS spectroscopy, complete IRAS photometry, and the adopted magnitude limit. Spectroscopic classifications depend on signal-to-noise, line visibility, and the classification criteria used. Interacting systems complicate the association of an infrared flux with an individual optical component. Finally, the DFBS classification results are class- and dataset-dependent; the high-confidence performance on selected samples cannot automatically be extended to the full archive or to infrared-selected samples.

These limitations do not negate the value of the results, but define their proper scope. The Thesis supports observational conclusions about the samples actually selected and examined; it does not establish a complete census of all IRAS galaxies, a universal activity sequence, or a causal evolutionary relation between star formation, AGN activity, and interactions.

6.13. Scientific Implications and Future Research

The **first priority** is to continue validating the large IRAS extragalactic catalogue. Secure counterparts, ambiguous associations, galaxy candidates, star candidates, and unknown sources should remain distinguishable, with new evidence incorporated in a way that preserves the history and confidence of each classification. A validated parent sample would provide a firmer basis for statistical comparison with BIG and for selecting further ULIRG and HLIRG candidates.

The **second priority** is to extend the luminosity and multiwavelength analysis to a larger and better-characterised population. The available IRAS fluxes can be combined with WISE, AKARI, SDSS, and other measurements to investigate infrared-to-optical properties by activity class. Such comparisons should account for redshift, photometric limits, data availability, and uncertainty, and should not treat observed colour or flux-ratio trends as direct evidence of causation.

The **third priority** is a systematic study of morphology and multiplicity. A consistently defined sample could compare isolated objects, pairs, multiple systems, and merger candidates across infrared luminosity and spectroscopic activity. This should include assessment of whether one optical component or the system as a whole is the most appropriate counterpart to the IRAS flux, and whether available spectra establish a physical association rather than a projected alignment.

The **fourth priority** is to test the application of DFBS extraction and machine-learning tools to sources selected from the IRAS programme. A practical workflow would cross-match IRAS/WISE/SDSS positions with archival coverage, extract relevant DFBS images, retain extraction and classification confidence measures, remove duplicate detections where appropriate, and validate the output against independently classified spectra. Independent test data, class-specific metrics, and explicit threshold choices would be needed before using model predictions in population statistics.

Finally, follow-up spectroscopy and reproducible data products would strengthen all four lines of work. Higher-quality spectra can resolve uncertain activity classes, while documented selection rules, calibrated catalogue fields, clear source provenance, and accessible data products would enable other researchers to reproduce the construction of samples and test the resulting interpretations.

6.14. Chapter Summary

The Thesis brings together four closely connected research components: detailed optical and spectroscopic study of BIG objects; preliminary extragalactic classification of the IRAS PSC/FSC Combined Catalogue; selection and spectral characterization of a high-luminosity sample containing 107 ULIRGs and 7 HLIRGs; and archival infrastructure and machine-learning methods based on ArVO, the BAO Plate Archive, and DFBS. Each component addresses a

different part of the path from infrared detection to physical interpretation, and each contributes evidence at a different scale.

Taken together, the results demonstrate the value of combining infrared selection, positional refinement, spectroscopy, luminosity analysis, activity classification, morphological context, and archival data methods. They also show why uncertainty must remain visible: the large extragalactic classification is preliminary, the high-luminosity sample is selection-limited, the interpretation of activity and colour distributions is not causal, and the DFBS machine-learning methods have not yet been directly applied to the complete IRAS sample. The resulting framework provides a basis for further validated sample construction and for more systematic studies of infrared luminosity, star formation, nuclear activity, and galaxy interactions.

CONCLUSION

This Thesis has investigated infrared galaxies selected from the IRAS PSC/FSC Combined Catalogue through four connected research components: optical identification and spectroscopy of the Byurakan-IRAS Galaxy (BIG) sample; construction of a large preliminary extragalactic classification of the Combined Catalogue; selection and spectroscopic study of high-luminosity infrared galaxies; and the development of archival and computational methods associated with the Armenian Virtual Observatory (ArVO), the Byurakan Astrophysical Observatory (BAO) Plate Archive, and the Digitized First Byurakan Survey (DFBS). Taken together, these studies show that the physical interpretation of an infrared detection depends on the coordinated use of positional, photometric, spectroscopic, and contextual information rather than on infrared flux alone.

The BIG component provides source-level spectroscopic information for the revised dataset of 257 spectra, assembled from observations with the BAO 2.6-m telescope, the 6-m telescope of the Special Astrophysical Observatory, the 1.93-m telescope of the Observatoire de Haute-Provence, and the Sloan Digital Sky Survey. The reported redshift interval is $0.0092 \leq z \leq 0.1997$, and the absolute magnitudes span approximately –24.46 to –16.81. HII/star-forming objects form the largest activity category, while AGN, Seyfert, LINER, Composite, mixed, emission-line, normal, absorption-line, and unclassified records are also represented. Because the dataset combines observations from different facilities and includes repeated records and systems with multiple optical components, its counts should be interpreted as a spectroscopic sample rather than as a homogeneous census of unique IRAS sources.

At the larger catalogue scale, the IRAS PSC/FSC Combined Catalogue provides a parent population of 345,163 sources. The hierarchical cross-correlation analysis produced 12,057 confirmed or genuine galaxies or QSOs, 68,940 galaxy candidates, 183,168 star candidates, and 80,998 sources that remained unknown at the reported preliminary stage. These categories provide a structured basis for continued investigation, but they do not constitute a final, fully verified classification of every source. In particular, the candidate and unknown groups require further validation, and overlapping catalogue matches must not be interpreted as independent objects.

The high-luminosity study illustrates how positional refinement and multiwavelength selection can lead from a very large infrared catalogue to a well-defined sample for detailed analysis. Of 6,346 sources with unique SDSS spectroscopic counterparts, 6,270 had measurements in all four IRAS bands required by the adopted luminosity calculation. Applying the infrared-luminosity threshold and the SDSS spectroscopic magnitude criterion yielded 114 objects: 107 ULIRGs and 7 HLIRGs. In the source study's comparison with NED, 48 ULIRGs were newly identified, while all seven objects were newly recognized as HLIRGs in the selected sample; two of those seven had previously been classified as ULIRGs. Thus, "newly identified as HLIRGs" does not necessarily mean that every such object was previously unknown as an astronomical source.

Spectroscopic analysis assigned activity classifications to 113 of the 114 high-luminosity objects; one remained unclassified because of poor spectral quality. The resulting population includes HII galaxies, broad- and narrow-line Seyferts, LINERs, composite systems, narrow-line AGN, BLL/QSO objects, and other emission- and absorption-line types. Comparison with the activity information available in NED also showed the value of direct spectral inspection for this selected sample. These classifications nevertheless depend on spectral quality and on the availability of diagnostic lines, so mixed and uncertain categories remain important parts of the scientific interpretation.

Across the samples, high infrared luminosity cannot be equated automatically with either starburst activity or AGN activity. The observations reveal systems dominated by star formation, systems with evidence of nuclear activity, and composite or ambiguous cases. The high-luminosity sample contains a substantial contribution from AGN-related types, but the observed differences between activity types, luminosity, redshift, and optical colour do not establish a causal evolutionary sequence between starbursts and AGN. These trends must be interpreted in the context of optical magnitude limits, spectroscopic availability, photometric requirements, and other selection/limitation effects. The high-luminosity sample is suited to the characterization of selected objects, not to a volume-complete census of all ULIRGs and HLIRGs.

Interacting, binary, and multiple systems provide an additional reason to examine the optical environment of infrared detections. When more than one optical component lies within the relevant infrared positional region, the measured infrared emission may arise from one galaxy or

from the system as a whole. This affects counterpart assignment and can influence the inferred distance, luminosity, and physical interpretation. The identification of such systems therefore motivates future work that combines reliable counterpart association with resolved imaging and, where possible, spectroscopy of multiple components.

The archival component extends the Thesis beyond modern survey catalogues. The BAO Plate Archive preserves a substantial historical observational record, while the DFBS provides digitized low-dispersion spectra derived from 1,874 plates and covering approximately 20 million objects with about 40 million extracted spectra. Our work [28] reported an overall supervised-classification accuracy of approximately 87% for its evaluated object classes, whereas the study [29] reported 96% accuracy for its tested sub-object task and 93% for its broad-group comparison. These figures belong to different experiments and datasets and should not be treated as a direct like-for-like performance comparison. The studies demonstrate the potential of automated extraction and classification for archival data, while also showing that class imbalance, training-set coverage, image quality, and confidence thresholds affect interpretation.

The principal contribution of the Thesis is the integration of these complementary approaches into a coherent methodology for infrared-galaxy research. The BIG sample supplies detailed spectroscopic characterization; the Combined Catalogue analysis establishes a broad but still preliminary classification framework; the ULIRG/HLIRG study examines the high-luminosity end of the selected IRAS population; and ArVO, the BAO Plate Archive, and DFBS provide resources and methods for extending the analysis to historical data. The principal limitations are likewise clear: the large-catalogue classifications require further verification, the spectroscopic samples are not uniformly selected, optical and photometric limits shape the high-luminosity sample, some spectra do not support secure fine classification, and the DFBS machine-learning methods have not yet been directly applied to the complete IRAS Combined Catalogue or to the 114-object ULIRG/HLIRG sample.

Future work should therefore prioritize validation and expansion rather than treating the present samples as final censuses. Important next steps include refining ambiguous counterpart associations, verifying candidate galaxies and unclassified infrared sources, consistently documenting the selection functions of the different samples, and studying pairs and multiple

systems with resolved multiwavelength observations. A further methodological direction is to cross-match archival DFBS coverage with selected IRAS sources and evaluate automated extraction and classification against independently classified spectra before using model outputs in population statistics. Within these limits, the Thesis establishes an observational and methodological foundation for more systematic studies of infrared luminosity, star formation, nuclear activity, and the role of galaxy interactions in shaping the observed properties of infrared galaxies.

BIBLIOGRAPHY

[1] A.M. Mickaelian, G. Harutyunyan, "Infrared Astronomy", *World of Science*, vol. 2, pp. 32–38, 2012.

[2] G.A. Mikayelyan, "Recent Results of Studies of IR Galaxies", *Communications of the Byurakan Astrophysical Observatory*, vol. 72, pp. 409–415, 2025.

[3] Joint IRAS Science Working Group, "IRAS Catalog of Point Sources, Version 2.0", *NASA RP-1190*, 1988.

[4] M. Moshir, G. Kopan, T. Conrow, H. McCallon, P. Hacking, D. Gregorich, G. Rohrbach, M. Melnyk, W. Rice, L. Fullmer, J. White, T. Chester, "The IRAS Faint Source Catalog, Version 2", *NASA/IPAC Infrared Science Archive*, 1990.

[5] D. B. Sanders, I. F. Mirabel, "Luminous Infrared Galaxies", *Annual Review of Astronomy & Astrophysics*, vol. 34, pp. 749–792, 1996.

[6] H.V. Abrahamyan, A.M. Mickaelian, A.V. Knyazyan, "The IRAS PSC/FSC Combined Catalogue", *Astronomy and Computing*, vol. 10, pp. 99–106, 2015.

[7] D. Ishihara, T. Onaka, H. Kataza, A. Salama, C. Alfageme, A. Cassatella, N. Cox, P. García-Lario, C. Stephenson, M. Cohen, N. Fujishiro, H. Fujiwara, S. Hasegawa, Y. Ita, W. Kim, H. Matsuhara, H. Murakami, T.G. Müller, T. Nakagawa, Y. Ohyama, S. Oyabu, J. Pyo, I. Sakon, H. Shibai, S. Takita, T. Tanabé, K. Uemizu, M. Ueno, F. Usui, T. Wada, H. Watarai, I. Yamamura, C. Yamauchi, "The AKARI/IRC mid-infrared all-sky survey", *Astronomy & Astrophysics*, vol. 514, id A1, 14 pp., 2010.

[8] I. Yamamura, S. Makiuti, N. Ikeda, Y. Fukuda, S. Oyabu, T. Koga, G.J. White, "AKARI/FIS All-Sky Survey Bright Source Catalogue", *VizieR On-line Data Catalog: II/298. Originally published in: ISAS/JAXA*, 2010.

[9] R.M. Cutri, E.L. Wright, T. Conrow, J.W. Fowler, P.R.M. Eisenhardt, C. Grillmair, J.D. Kirkpatrick, F. Masci, H.L. McCallon, S.L. Wheelock, S. Fajardo-Acosta, L. Yan, D. Benford, M. Harbut, T. Jarrett, S. Lake, D. Leisawitz, M.E. Ressler, S.A. Stanford, C.W. Tsai, F. Liu, G. Helou, A. Mainzer, D. Gettings, A. Gonzalez, D. Hoffman, K.A. Marsh, D. Padgett, M.F. Skrutskie, R.P. Beck, M. Papin, M. Wittman, "Explanatory Supplement to the AllWISE Data Release Products", IPAC/Caltech, 2013.

[10] A.M. Mickaelian, "Optical identification of the IRAS point sources based on the first Byurakan survey", *Astrophysics*, vol. 38, pp. 349–351, 1995.

[11] A.M. Mickaelian, L.A. Sargsyan, "Optical Identification of IRAS Point Sources. Galaxies. X.", *Astrophysics*, vol. 47, pp. 213–228, 2004.

[12] A.M. Mickaelian, L.R. Hovhannisyan, L.A. Sargsyan, "Spectroscopic Study of Multiple IRAS Galaxies", *Astrophysics*, vol. 46, pp. 177–190, 2003.

[13] L.A. Sargsyan, A.M. Mickaelian, "Spectral study of BIG objects on the 2.6-m telescope at the Byurakan Astrophysical Observatory", *Astrophysics*, vol. 49, pp. 19–35, 2006.

[14] A.M. Mickaelian, S.A. Akopian, S.K. Balayan, A.N. Burenkov, "Optical studies of infrared sources", *Astronomy Letters*, vol. 24, pp. 635–642, 1998.

[15] S.K. Balayan, S.A. Akopian, A.M. Mickaelian, A.N. Burenkov, "Optical Studies of Extragalactic Infrared Sources", *Astronomy Letters*, vol. 27, pp. 284–291, 2001.

[16] A.M. Mickaelian, "Spectral Observations of BIG Objects. III", *Astrophysics*, vol. 47, pp. 361–377, 2004.

[17] A.M. Mickaelian, G.S. Harutyunyan, A. Sarkissian, "Study and Classification of SDSS Spectra for Byurakan-IRAS Galaxies", *Astronomy Letters*, vol. 44, pp. 351–361, 2018.

[18] B. Abolfathi, D. S. Aguado, G. Aguilar, C. Allende Prieto, A. Almeida, T. T. Ananna, F. Anders, S. F. Anderson, B. H. Andrews, B. Anguiano, A. Aragón-Salamanca, M. Argudo-Fernández, E. Armengaud, M. Ata, E. Aubourg, V. Avila-Reese, C. Badenes, S. Bailey, C. Balland, K. A. Barger, J. Barrera-Ballesteros, C. Bartosz, F. Bastien, D. Bates, F. Baumgarten, J. Bautista, R. Beaton, T. C. Beers, F. Belfiore, C. F. Bender, M. Bernardi, M. A. Bershady, F. Beutler, J. C. Bird, D. Bizyaev, G. A. Blanc, M. R. Blanton, M. Blomqvist, A. S. Bolton, M. Boquien, J. Borissova, J. Bovy, C. A. Bradna Diaz, W. N. Brandt, J. Brinkmann, J. R. Brownstein, K. Bundy, A. J. Burgasser, E. Burtin, N. G. Busca, C. I. Cañas, M. Cano-Díaz, M. Cappellari, R. Carrera, A. R. Casey, B. Cervantes Sodi, Y. Chen, B. Cherinka, C. Chiappini, P. D. Choi, D. Chojnowski, C.-H. Chuang, H. Chung, N. Clerc, R. E. Cohen, J. M. Comerford, J. Comparat, J. Correa do Nascimento, L. da Costa, M.-C. Cousinou, K. Covey, J. D. Crane, I. Cruz-Gonzalez, K. Cunha, G. da Silva Ilha, G. J. Damke, J. Darling, J. W. Davidson Jr., K. Dawson, M. A. C. de Icaza Lizaola, A. de la Macorra, S. de la Torre, N. De Lee, V. de Sainte Agathe, A. Deconto Machado, F. Dell'Agli, T. Delubac, A. M. Diamond-Stanic, J. Donor, J. J. Downes, N. Drory, H. du Mas des Bourboux, C. J. Duckworth, T. Dwelly, J. Dyer, G. Ebelke, A. D. Eigenbrot, D. J. Eisenstein, Y. P. Elsworth, E. Emsellem, M. Eracleous, G. Erfanianfar, S. Escoffier, X. Fan, E. Fernández Alvar, J. G. Fernandez-Trincado, R. F. Cirolini, D. Feuillet, A. Finoguenov, S. W. Fleming, A. Font-Ribera, G. Freischlad, P. Frinchaboy, H. Fu, Y. Gómez Maqueo Chew, L. Galbany, A. E. García Pérez, R. Garcia-Dias, D. A. García-Hernández, L. A. Garma Oehmichen, P. Gaulme, J. Gelfand, H. Gil-Marín, B. A. Gillespie, D. Goddard, J. I. González Hernández, V. Gonzalez-Perez, K.

Grabowski, P. J. Green, C. J. Grier, A. Gueguen, H. Guo, J. Guy, A. Hagen, P. Hall, P. Harding, S. Hasselquist, S. Hawley, C. R. Hayes, F. Hearty, S. Hekker, J. Hernandez, H. Hernandez Toledo, D. W. Hogg, K. Holley-Bockelmann, J. A. Holtzman, J. Hou, B.-C. Hsieh, J. A. S. Hunt, T. A. Hutchinson, H. S. Hwang, C. E. Jimenez Angel, J. A. Johnson, A. Jones, H. Jönsson, E. Jullo, F. S. Khan, K. Kinemuchi, D. Kirkby, C. C. Kirkpatrick IV, F.-S. Kitaura, G. R. Knapp, J.-P. Kneib, J. A. Kollmeier, I. Lacerna, R. R. Lane, D. Lang, D. R. Law, J.-M. Le Goff, Y.-B. Lee, H. Li, C. Li, J. Lian, Y. Liang, M. Lima, L. Lin, D. Long, S. Lucatello, B. Lundgren, J. T. Mackereth, C. L. MacLeod, S. Mahadevan, M. A. Geimba Maia, S. Majewski, A. Manchado, C. Maraston, V. Mariappan, R. Marques-Chaves, T. Masseron, K. L. Masters, R. M. McDermid, I. D. McGreer, M. Melendez, S. Meneses-Goytia, A. Merloni, M. R. Merrifield, S. Meszaros, A. Meza, I. Minchev, D. Minniti, E.-M. Mueller, F. Muller-Sanchez, D. Muna, R. R. Muñoz, A. D. Myers, P. Nair, K. Nandra, M. Ness, J. A. Newman, R. C. Nichol, D. L. Nidever, C. Nitschelm, P. Noterdaeme, J. O'Connell, R. J. Oelkers, A. Oravetz, D. Oravetz, E. Aquino Ortíz, Y. Osorio, Z. Pace, N. Padilla, N. Palanque-Delabrouille, P. A. Palicio, H.-A. Pan, K. Pan, T. Parikh, I. Pâris, C. Park, S. Peirani, M. Pellejero-Ibanez, S. Penny, W. J. Percival, I. Perez-Fournon, P. Petitjean, M. M. Pieri, M. Pinsonneault, A. Pisani, F. Prada, A. Prakash, A. B. de Andrade Queiroz, M. J. Raddick, A. Raichoor, S. B. Rembold, H. Richstein, R. A. Riffel, R. Riffel, H.-W. Rix, A. C. Robin, S. Rodríguez Torres, C. Román-Zúñiga, A. J. Ross, G. Rossi, J. Ruan, R. Ruggeri, J. Ruiz, M. Salvato, A. G. Sánchez, S. F. Sánchez, J. Sanchez Almeida, J. R. Sánchez-Gallego, F. A. Santana Rojas, B. X. Santiago, R. P. Schiavon, J. S. Schimoia, E. Schlafly, D. Schlegel, D. P. Schneider, W. J. Schuster, A. Schwope, H.-J. Seo, A. Serenelli, S. Shen, Y. Shen, M. Shetrone, M. Shull, V. Silva Aguirre, J. D. Simon, M. Skrutskie, A. Slosar, R. Smethurst, V. Smith, J. Sobeck, G. Somers, B. J. Souter, D. Souto, A. Spindler, D. V. Stark, K. Stassun, M. Steinmetz, D. Stello, T. Storchi-Bergmann, A. Streblyanska, G. S. Stringfellow, G. Suárez, J. Sun, L. Szigeti, M. Taghizadeh-Popp, M. S. Talbot, B. Tang, C. Tao, J. Tayar, M. Tembe, J. Teske, A. R. Thakar, D. Thomas, P. Tissera, R. Tojeiro, C. Tremonti, N. W. Troup, M. Urry, O. Valenzuela, R. van den Bosch, J. Vargas-González, M. Vargas-Magaña, J. A. Vazquez, S. Villanova, N. Vogt, D. Wake, Y. Wang, B. A. Weaver, A.-M. Weijmans, D. H. Weinberg, K. B. Westfall, D. G. Whelan, E. Wilcots, V. Wild, R. A. Williams, J. Wilson, W. M. Wood-Vasey, D. Wylezalek, T. Xiao, R. Yan, M. Yang, J. E. Ybarra, C. Yèche, N. Zakamska, O. Zamora, P. Zarrouk, G. Zasowski, K. Zhang, C. Zhao, G.-B. Zhao, Z. Zheng, Z. Zheng, Z.-M. Zhou, G. Zhu, J. C. Zinn, H. Zou, "The Fourteenth Data Release of the Sloan Digital Sky Survey: First Spectroscopic Data from the Extended Baryon Oscillation Spectroscopic Survey and from the Second Phase of the Apache Point Observatory Galactic Evolution Experiment", *The Astrophysical Journal Supplement Series*, vol. 235, id 42, 19 pp., 2018.

[19] G. A. Mikayelyan, A. M. Mickaelian, H. V. Abrahamyan, G. M. Paronyan, M. V. Gyulzadyan, "Observations and Study of Byurakan-IRAS Galaxies: Summary", *Communications of the Byurakan Astrophysical Observatory*, vol. 66, pp. 31–40, 2019.

[20] G. A. Mikayelyan, A. M. Mickaelian, H. V. Abrahamyan, G. M. Paronyan, "Building the Largest Sample of IR Galaxies: Preliminary Results", *Communications of the Byurakan Astrophysical Observatory*, vol. 70, pp. 321–327, 2023.

[21] P.-A. Duc, I.F. Mirabel, J. Maza, "Southern ultraluminous infrared galaxies: an optical and infrared database", *Astronomy & Astrophysics Supplement Series*, vol. 124, pp. 533–557, 1997.

[22] G. A. Mikayelyan, A. M. Mickaelian, H. V. Abrahamyan, G. M. Paronyan, "Discovery of New Bright ULIRGs from the IRAS PSC/FSC Combined Catalogue", *Communications of the Byurakan Astrophysical Observatory*, vol. 65, pp. 13–20, 2018.

[23] G. A. Mikayelyan, A. M. Mickaelian, H. V. Abrahamyan, G. M. Paronyan, M. V. Gyulzadyan, "Spectral Classification of Ultrabright IR Galaxies from the IRAS PSC/FSC Combined Catalog", *Astrophysics*, vol. 62, pp. 452–466, 2019.

[24] S. Veilleux, "Spectroscopic Diagnostics for AGNs", *Proceedings of IAU Colloquium 184, ASP Conference Series, held 18-22 June, 2001 in Byurakan, Armenia. Edited by R.F. Green, E.Ye. Khachikian, and D.B. Sanders. San Francisco, CA: Astronomical Society of the Pacific*, vol. 284, pp. 111–125, 2002.

[25] G.A. Mikayelyan, G.M. Paronyan, N.M. Azatyan, G.R. Kostandyan, A.L. Samsonyan, H.R. Andreasyan, H.V. Abrahamyan, D.H. Andreasyan, A.G. Suqiasyan, A.M. Mickaelian, "BAO Plate Archive Project: Recent Results", *Communications of the Byurakan Astrophysical Observatory*, vol. 68, pp. 97–104, 2021.

[26] A. M. Mickaelian, R. Nesci, C. Rossi, D. Weedman, G. Cirimele, L. A. Sargsyan, L. K. Erastova, K. S. Gigoyan, G. A. Mikayelyan, E. Massaro, S. Gaudenzi, J. Houck, D. Barry, L. D'Amante, P. Germano, "The Digitized First Byurakan Survey – DFBS", *Astronomy & Astrophysics*, vol. 464, pp. 1177–1180, 2007.

[27] E. Massaro, A.M. Mickaelian, R. Nesci, D. Weedman, "The Digitized First Byurakan Survey", *ARACNE Editrice, Rome, Italy*, 78 pages, 2008.

[28] H. Astsatryan, G. Gevorgyan, A. Knyazyan, A. Mickaelian, G. Mikayelyan, "Astronomical Objects Classification Based on the Digitized First Byurakan Survey Low-Dispersion Spectra", *Astronomy and Computing*, vol. 34, id 100442, 2021.

[29] H. Astsatryan, S. Babayan, A. Mickaelian, G. Mikayelyan, M. Astsatryan, "Cloud-Based Machine Learning Service for Astronomical Sub-Object Classification: Case Study on the First Byurakan Survey Spectra", *Data Science Journal*, vol. 23 (1), id 6, 14 pp., 2024.

[30] M.F. Kessler, J.A. Steinz, M.E. Anderegg, J. Clavel, G. Drechsel, P. Estaria, J. Faelker, J.R. Riedinger, A. Robson, B.G. Taylor, S. Ximénez de Ferrán, "The Infrared Space Observatory (ISO) mission", *Astronomy & Astrophysics*, vol. 315, pp. L27–L31, 1996.

[31] M. W. Werner, T. L. Roellig, F. J. Low, G. H. Rieke, M. Rieke, W. F. Hoffmann, E. Young, J. R. Houck, B. Brandl, G. G. Fazio, J. L. Hora, R. D. Gehrz, G. Helou, B. T. Soifer, J. Stauffer, J. Keene, P. Eisenhardt, D. Gallagher, T. N. Gautier, W. Irace, C. R. Lawrence, L. Simmons, J. E. Van Cleve, M. Jura, E. L. Wright, D. P. Cruikshank, "The Spitzer Space Telescope Mission", *The Astrophysical Journal Supplement Series*, vol. 154, pp. 1–9, 2004.

[32] G. L. Pilbratt, J. R. Riedinger, T. Passvogel, G. Crone, D. Doyle, U. Gageur, A. M. Heras, C. Jewell, L. Metcalfe, S. Ott, M. Schmidt, "Herschel Space Observatory. An ESA facility for far-infrared and submillimetre astronomy", *Astronomy & Astrophysics*, vol. 518, id L1, 6 pp., 2010.

[33] J. P. Gardner, J. C. Mather, R. Abbott, J. S. Abell, M. Abernathy, F. E. Abney, J. G. Abraham, R. Abraham, Y. M. Abul-Huda, S. Acton, C. K. Adams, E. Adams, D. S. Adler, M. Adriaensen, J. A. Aguilar, M. Ahmed, N. S. Ahmed, T. Ahmed, R. Albat, L. Albert, S. Alberts, D. Aldridge, M. M. Allen, S. S. Allen, M. Altenburg, S. Altunc, J. L. Alvarez, J. Álvarez-Márquez, C. Alves de Oliveira, L. L. Ambrose, S. M. Anandakrishnan, G. C. Andersen, H. J. Anderson, J. Anderson, K. Anderson, S. M. Anderson, J. Aprea, B. J. Archer, J. W. Arenberg, I. Argyriou, S. Arribas, É. Artigau, A. R. Arvai, P. Atcheson, C. B. Atkinson, J. Averbukh, C. Aymergen, J. J. Bacinski, W. E. Baggett, G. Bagnasco, L. L. Baker, V. A. Balzano, K. A. Banks, D. A. Baran, E. A. Barker, L. K. Barrett, B. O. Barringer, A. Barto, W. Bast, P. Baudoz, S. Baum, T. G. Beatty, M. Beaulieu, K. Bechtold, T. Beck, M. M. Beddard, C. Beichman, L. Bellagama, P. Bely, T. W. Berger, L. E. Bergeron, A.-D. Bernier, M. D. Bertch, C. Beskow, L. E. Betz, C. P. Biagetti, S. Birkmann, K. F. Bjorklund, J. D. Blackwood, R. P. Blazek, S. Blossfeld, M. Bluth, A. Boccaletti, M. E. Boegner Jr., R. C. Bohlin, J. J. Boia, T. Böker, N. Bonaventura, N. A. Bond, K. A. Bosley, R. A. Boucarut, P. Bouchet, J. Bouwman, G. Bower, A. S. Bowers, C. W. Bowers, L. A. Boyce, C. T. Boyer, M. L. Boyer, M. Boyer, R. Boyer, L. D. Bradley, G. R. Brady, B. R. Brandl, J. L. Brannen, D. Breda, H. G. Bremmer, D. Brennan, P. A. Bresnahan, S. N. Bright, B. J. Broiles, A. Bromenschenkel, B. H. Brooks, K. J. Brooks, B. Brown, B. Brown, T. M. Brown, B. W. Bruce, J. G. Bryson, E. D. Bujanda, B. M. Bullock, A. J. Bunker, R. Bureo, I. J. Burt, J. A. Bush, H. A. Bushouse, M. C. Bussman, O. Cabaud, S. Cale, C. D. Calhoon, H. Calvani, A. M. Canipe, F. M. Caputo, M. Cara, L. Carey, M. E. Case, T. Cesari, L. D. Cetorelli, D. R. Chance, L. Chandler, D. Chaney,

G. N. Chapman, S. Charlot, P. Chayer, J. I. Cheezum, B. Chen, C. H. Chen, B. Cherinka, S. C. Chichester, Z. S. Chilton, D. Chittiraibalan, M. Clampin, C. R. Clark, K. W. Clark, S. M. Clark, E. E. Claybrooks, K. A. Cleveland, A. L. Cohen, L. M. Cohen, K. D. Colón, B. L. Coleman, L. Colina, B. J. Comber, T. M. Comeau, T. Comer, A. Conde Reis, D. C. Connolly, K. E. Conroy, A. R. Contos, J. Contreras, N. J. Cook, J. L. Cooper, R. A. Cooper, M. F. Correia, M. Correnti, C. Cossou, B. F. Costanza, A. Coulais, C. R. Cox, R. T. Coyle, M. M. Cracraft, K. A. Crew, G. J. Curtis, B. Cusveller, C. Da Costa Maciel, C. T. Dailey, F. Daugeron, G. S. Davidson, J. E. Davies, K. A. Davis, M. S. Davis, R. Day, D. de Chambure, P. de Jong, G. De Marchi, B. H. Dean, J. E. Decker, A. S. Delisa, L. C. Dell, G. Dellagatta, F. Dembinska, S. Demosthenes, N. M. Dencheva, P. Deneu, W. W. DePriest, J. Deschenes, N. Dethienne, Ö. H. Detre, R. I. Diaz, D. Dicken, A. S. DiFelice, M. Dillman, M. O. Disharoon, W. Dixon V., J. B. Doggett, K. L. Dominguez, T. S. Donaldson, C. M. Doria-Warner, T. D. Santos, H. Doty, R. E. Douglas Jr., R. Doyon, A. Dressler, J. Driggers, P. A. Driggers, J. L. Dunn, K. C. DuPrie, J. Dupuis, J. Durning, S. B. Dutta, N. M. Earl, P. Eccleston, P. Ecobichon, E. Egami, R. Ehrenwinkler, J. D. Eisenhamer, M. Eisenhower, D. J. Eisenstein, Z. El Hamel, M. L. Elie, J. Elliott, K. W. Elliott, M. Engesser, N. Espinoza, O. Etienne, M. Etxaluze, L. Evans, L. Fabreguettes, M. Falcolini, P. R. Falini, C. Fatig, M. Feeney, L. D. Feinberg, R. Fels, N. Ferdous, H. C. Ferguson, L. Ferrarese, M.-H. Ferreira, P. Ferruit, M. Ferry, J. C. Filippazzo, D. Firre, M. Fix, N. Flagey, K. A. Flanagan, S. W. Fleming, M. Florian, J. R. Flynn, L. Foiadelli, M. R. Fontaine, E. M. Fontanella, P. R. Forshay, E. A. Fortner, O. D. Fox, A. P. Framarini, J. I. Francisco, R. Franck, M. Franx, D. E. Franz, S. D. Friedman, K. E. Friend, J. R. Frost, H. Fu, A. W. Fullerton, L. Gaillard, S. Galkin, B. Gallagher, A. D. Galyer, M. García Marín, L. E. Gardner, D. Garland, B. A. Garrett, D. Gasman, A. Gáspár, R. Gastaud, D. Gaudreau, P. T. Gauthier, V. Geers, P. H. Geithner, M. Gennaro, J. Gerber, J. C. Gereau, R. Giampaoli, G. Giardino, P. C. Gibbons, K. Gilbert, L. Gilman, J. H. Girard, M. E. Giuliano, K. Gkountis, A. Glasse, K. Z. Glassmire, A. M. Glauser, S. D. Glazer, J. Goldberg, D. A. Golimowski, S. P. Gonzaga, K. D. Gordon, S. J. Gordon, P. Goudfrooij, M. J. Gough, A. J. Graham, C. M. Grau, J. D. Green, G. R. Greene, T. P. Greene, P. E. Greenfield, M. A. Greenhouse, T. R. Greve, E. M. Greville, S. Grimaldi, F. E. Groe, A. Groebner, D. M. Grumm, T. Grundy, M. Güdel, P. Guillard, J. Guldalian, C. A. Gunn, A. Gurule, I. M. Gutman, P. D. Guy, B. Guyot, W. J. Hack, P. Haderlein, J. B. Hagan, A. Hagedorn, K. Hainline, C. Haley, M. Hami, F. C. Hamilton, J. Hammann, H. B. Hammel, C. J. Hanley, C. A. Hansen, B. Hardy, B. Harnisch, M. H. Harr, P. Harris, J. A. Hart, G. F. Hartig, H. Hasan, K. M. Hashim, R. Hashimoto, S. J. Haskins, R. E. Hawkins, B. Hayden, W. L. Hayden, M. Healy, K. Hecht, V. J. Heeg, R. Hejal, K. A. Helm, N. J. Hengemihle, T. Henning, A. Henry, R. L. Henry, K. Henshaw, S. Hernandez, D. C. Herrington, A. Heske, B. E. Hesman, D. L. Hickey, B. N. Hilbert, D. C. Hines, M. R. Hinz, M. Hirsch, R. S. Hitcho, K. Hodapp, P. E. Hodge, M. Hoffman, S. T. Holfeltz, B. J. Holler, J. R. Hoppa, S. Horner, J. M. Howard, R. J. Howard, J. M. Huber, J. S. Hunkeler, A. Hunter, D. G.

Hunter, S. W. Hurd, B. J. Hurst, J. B. Hutchings, J. E. Hylan, L. I. Ignat, G. Illingworth, S. M. Irish, J. C. Isaacs III, W. C. Jackson Jr., D. T. Jaffe, J. Jahic, A. Jahromi, P. Jakobsen, B. James, J. C. James, L. R. James, W. B. Jamieson, R. D. Jandra, R. Jayawardhana, R. Jedrzejewski, B. S. Jeffers, P. Jensen, E. Joanne, A. T. Johns, C. A. Johnson, E. L. Johnson, P. Johnson, P. S. Johnson, T. K. Johnson, T. W. Johnson, D. Johnstone, D. Jollet, D. P. Jones, G. S. Jones, O. C. Jones, R. A. Jones, V. Jones, I. J. Jordan, M. E. Jordan, R. Jue, M. H. Jurkowski, G. Justis, K. Justtanont, C. C. Kaleida, J. S. Kalirai, P. C. Kalmanson, L. Kaltenegger, J. Kammerer, S. K. Kan, G. C. Kanarek, S.-H. Kao, D. M. Karakla, H. Karl, S. A. Kassin, D. D. Kauffman, P. Kavanagh, L. L. Kelley, D. M. Kelly, S. Kendrew, H. Kennedy V., D. A. Kenny, R. A. Keski-Kuha, C. D. Keyes, A. Khan, R. C. Kidwell, R. A. Kimble, J. S. King, R. C. King, W. M. Kinzel, J. R. Kirk, M. E. Kirkpatrick, P. Klaassen, L. Klingemann, P. U. Klintworth, B. A. Knapp, S. Knight, P. J. Knollenberg, D. M. Knutsen, R. Koehler, A. M. Koekemoer, E. T. Kofler, V. L. Kontson, A. R. Kovacs, V. Kozhurina-Platais, O. Krause, G. A. Kriss, J. Krist, M. R. Kristoffersen, C. Krogel, A. P. Krueger, B. A. Kulp, N. Kumari, S. W. Kwan, M. Kyprianou, A. G. Labador, Á. Labiano, D. Lafrenière, P.-O. Lagage, V. G. Laidler, B. Laine, S. Laird, C.-P. Lajoie, M. D. Lallo, M. Y. Lam, S. M. LaMassa, S. D. Lambros, R. J. Lampenfield, M. E. Lander, J. H. Langston, K. Larson, M. Larson, R. J. LaVerghetta, D. R. Law, J. F. Lawrence, D. W. Lee, J. Lee, Y.-N. P. Lee, J. Leisenring, M. D. Leveille, N. A. Levenson, J. S. Levi, M. B. Levine, D. Lewis, J. Lewis, N. Lewis, M. Libralato, N. Lidon, P. L. Liebrecht, P. Lightsey, S. Lilly, F. C. Lim, P. L. Lim, S.-K. Ling, L. J. Link, M. N. Link, J. L. Lipinski, X. Liu, A. S. Lo, L. Lobmeyer, R. M. Logue, C. A. Long, D. R. Long, I. D. Long, K. S. Long, M. López-Caniego, J. M. Lotz, J. M. Love-Pruitt, M. Lubskiy, E. B. Luers, R. A. Luetgens, A. J. Luevano, S. M. G. F. Lui, J. M. Lund III, R. A. Lundquist, J. Lunine, N. Lützgendorf, R. J. Lynch, A. J. MacDonald, K. MacDonald, M. J. Macias, K. I. Macklis, P. Maghami, R. Y. Maharaja, R. Maiolino, K. G. Makrygiannis, S. G. Malla, E. M. Malumuth, E. Manjavacas, A. Marini, A. Marrione, A. Marston, A. R. Martel, D. Martin, P. G. Martin, K. L. Martinez, M. Maschmann, G. L. Masci, M. E. Masetti, M. Maszkiewicz, G. Matthews, J. E. Matuskey, G. A. McBrayer, D. W. McCarthy, M. J. McCaughrean, L. A. McClare, M. D. McClare, J. C. McCloskey, T. D. McClurg, M. McCoy, M. W. McElwain, R. D. McGregor, D. B. McGuffey, A. G. McKay, W. K. McKenzie, B. McLean, M. McMaster, W. McNeil, W. De Meester, K. L. Mehalick, M. Meixner, M. Meléndez, M. P. Menzel, M. T. Menzel, M. Merz, D. D. Mesterharm, M. R. Meyer, M. L. Meyett, L. E. Meza, C. Midwinter, S. N. Milam, J. T. Miller, W. C. Miller, C. L. Miskey, K. Misselt, E. P. Mitchell, M. Mohan, E. E. Montoya, M. J. Moran, T. Morishita, A. Moro-Martín, D. L. Morrison, J. Morrison, E. C. Morse, M. Moschos, S. H. Moseley, G. E. Mosier, P. Mosner, M. Mountain, J. S. Muckenthaler, D. G. Mueller, M. Mueller, D. Muhiem, P. Mühlmann, S. E. Mullally, S. M. Mullen, A. J. Munger, J. Murphy, K. T. Murray, J. C. Muzerolle, M. Mycroft, A. Myers, C. R. Myers, F. R. R. Myers, R. Myers, K. Myrick, A. F. Nagle IV, O. Nayak, B. Naylor, S. G. Neff, E. P. Nelan, J. Nella, D. T. Nguyen,

M. N. Nguyen, B. Nickson, J. J. Nidhiry, M. B. Niedner, M. Nieto-Santisteban, N. K. Nikolov, M. A. Nishisaka, A. Noriega-Crespo, A. Nota, R. C. O'Mara, M. Oboryshko, M. B. O'Brien, W. R. Ochs, J. D. Offenberg, P. M. Ogle, R. G. Ohl, J. H. Olmsted, S. B. Osborne, B. P. O'Shaughnessy, G. Östlin, B. O'Sullivan, O. J. Otor, R. Ottens, N. N.-Q. Ouellette, D. J. Outlaw, B. A. Owens, C. Pacifici, J. C. Page, J. G. Paranilam, S. Park, K. A. Parrish, L. Paschal, P. Patapis, J. Patel, K. Patrick, R. A. Pattishall Jr., D. W. Paul, S. J. Paul, T. A. Pauly, C. M. Pavlovsky, M. Peña-Guerrero, A. H. Pedder, M. W. Peek, P. A. Pelham, K. Penanen, B. A. Perriello, M. D. Perrin, R. F. Perrine, C. Perrygo, M. Peslier, M. Petach, K. A. Peterson, T. Pfarr, J. M. Pierson, M. Pietraszkiewicz, G. Pilchen, J. L. Pipher, N. Pirzkal, J. T. Pitman, D. M. Player, R. Plesha, A. Plitzke, J. A. Pohner, K. K. Poletis, J. A. Pollizzi, E. Polster, J. T. Pontius, K. Pontoppidan, S. C. Porges, G. D. Potter, S. Prescott, C. R. Proffitt, L. Pueyo, I. A. Quispe Neira, A. Radich, R. T. Rager, J. Rameau, D. D. Ramey, R. Ramos Alarcon, R. Rampini, R. Rapp, R. A. Rashford, B. J. Rauscher, S. Ravindranath, T. Rawle, T. N. Rawlings, T. Ray, M. W. Regan, B. Rehm, K. D. Rehm, N. Reid, C. A. Reis, F. Renk, T. B. Reoch, M. Ressler, A. W. Rest, P. J. Reynolds, J. G. Richon, K. Richon V., M. Ridgaway, A. R. Riedel, G. H. Rieke, M. J. Rieke, R. E. Rifelli, J. R. Rigby, C. S. Riggs, N. J. Ringel, C. E. Ritchie, H.-W. Rix, M. Robberto, G. L. Robinson, M. S. Robinson, O. Robinson, F. W. Rock, D. R. Rodriguez, B. Rodríguez del Pino, T. Roellig, S. O. Rohrbach, A. J. Roman, F. J. Romelfanger, F. P. Romo Jr., J. J. Rosales, P. Rose, A. F. Roteliuk, M. N. Roth, B. Q. Rothwell, S. Rouzaud, J. Rowe, N. Rowlands, A. Roy, P. Royer, C. Rui, P. Rumler, W. Rumpl, M. L. Russ, M. B. Ryan, R. M. Ryan, K. Saad, M. Sabata, R. Sabatino, E. Sabbi, P. A. Sabelhaus, S. Sabia, K. C. Sahu, B. N. Saif, J.-C. Salvignol, P. Samara-Ratna, B. S. Samuelson, F. A. Sanders, B. Sappington, B. A. Sargent, A. Sauer, B. J. Savadkin, M. Sawicki, T. M. Schappell, C. Scheffer, S. Scheithauer, R. Scherer, C. Schiff, E. Schlawin, O. Schmeitzky, T. S. Schmitz, D. J. Schmude, A. Schneider, J. Schreiber, H. Schroeven-Deceuninck, J. J. Schultz, R. Schwab, C. H. Schwartz, D. Scoccimarro, J. F. Scott, M. B. Scott, B. L. Seaton, B. S. Seely, B. Seery, M. Seidleck, K. Sembach, C. E. Shanahan, B. Shaughnessy, R. A. Shaw, C. M. Shay, E. Sheehan, K. Sheth, H.-Y. Shih, I. Shivaei, N. Siegel, M. G. Sienkiewicz, D. D. Simmons, B. P. Simon, M. Sirianni, A. Sivaramakrishnan, J. E. Slade, G. C. Sloan, C. E. Slocum, S. E. Slowinski, C. T. Smith, E. P. Smith, E. C. Smith, K. Smith, R. Smith, S. J. Smith, J. L. Smolik, D. R. Soderblom, S. T. Sohn, J. Sokol, G. Sonneborn, C. D. Sontag, P. R. Sooy, R. Soummer, D. M. Southwood, K. Spain, J. Sparmo, D. T. Speer, R. Spencer, J. D. Sprofera, S. S. Stallcup, M. K. Stanley, J. A. Stansberry, C. C. Stark, C. W. Starr, D. Y. Stassi, J. A. Steck, C. D. Steeley, M. A. Stephens, R. J. Stephenson, A. C. Stewart, M. Stiavelli, H. Stockman Jr., P. Strada, A. N. Straughn, S. Streetman, D. K. Strickland, J. F. Strobele, M. Stuhlinger, J. E. Stys, M. Such, K. Sukhatme, J. F. Sullivan, P. C. Sullivan, S. M. Sumner, F. Sun, B. D. Sunnquist, D. A. Swade, M. S. Swam, D. F. Swenton, R. A. Swoish, O. I. Tam Litten, L. Tamas, A. Tao, D. K. Taylor, J. M. Taylor, M. te Plate, M. Van Tea, K. K. Teague, R. C. Telfer, T. Temim, S. C.

Texter, D. G. Thatte, C. L. Thompson, L. M. Thompson, S. R. Thomson, H. Thronson, C. M. Tierney, T. Tikkanen, L. Tinnin, W. T. Tippet, C. W. Todd, H. D. Tran, J. Trauger, E. G. Trejo, J. H. Vinh Truong, C. L. Tsukamoto, Y. Tufail, J. Tumlinson, S. Tustain, H. Tyra, L. Ubeda, K. Underwood, M. A. Uzzo, S. Vaclavik, F. Valenduc, J. A. Valenti, J. Van Campen, I. van de Wetering, R. P. Van Der Marel, R. van Haarlem, B. Vandenbussche, E. F. van Dishoeck, D. D. Vanterpool, M. R. Vernoy, M. B. Vila Costas, K. Volk, P. Voorzaat, M. F. Voyton, E. Vydra, D. J. Waddy, C. Waelkens, G. M. Wahlgren, F. E. Walker Jr., M. Wander, C. K. Warfield, G. Warner, F. C. Wasiak, M. F. Wasiak, J. Wehner, K. R. Weiler, M. Weilert, S. B. Weiss, M. Wells, A. D. Welty, L. Wheate, T. P. Wheeler, C. L. White, P. Whitehouse, J. M. Whiteleather, W. R. Whitman, C. C. Williams, C. N. A. Willmer, C. J. Willott, S. P. Willoughby, A. Wilson, D. Wilson, D. Wilson V., R. Windhorst, E. C. Wislowski, D. J. Wolfe, M. A. Wolfe, S. Wolff, A. Wondel, C. Woo, R. T. Woods, E. Worden, W. Workman, G. S. Wright, C. Wu, C.-R. Wu, D. D. Wun, K. B. Wymer, T. Yadetie, I. C. Yan, K. C. Yang, K. L. Yates, C. R. Yeager, E. J. Yerger, E. T. Young, G. Young, G. Yu, S. Yu, D. S. Zak, P. Zeidler, R. Zepp, J. Zhou, C. A. Zincke, S. Zonak, E. Zondag, "The James Webb Space Telescope Mission", *Publications of the Astronomical Society of the Pacific,* vol. 135, id 068001, 24 pages, 2023.

[34] N. Epchtein, E. Deul, S. Derriere, J. Borsenberger, D. Egret, G. Simon, C. Alard, L. G. Balázs, B. de Batz, M.-R. Cioni, E. Copet, M. Dennefeld, T. Forveille, P. Fouqué, F. Garzón, H. J. Habing, A. Holl, J. Hron, S. Kimeswenger, F. Lacombe, T. Le Bertre, C. Loup, G. A. Mamon, A. Omont, G. Paturel, P. Persi, A. C. Robin, D. Rouan, D. Tiphène, I. Vauglin, S. J. Wagner, "A preliminary database of DENIS point sources", *Astronomy & Astrophysics*, vol. 349, pp. 236–242, 1999.

[35] R. M. Cutri, M. F. Skrutskie, S. van Dyk, C. A. Beichman, J. M. Carpenter, T. Chester, L. Cambresy, T. Evans, J. Fowler, J. Gizis, E. Howard, J. Huchra, T. Jarrett, E. L. Kopan, J. D. Kirkpatrick, R. M. Light, K. A. Marsh, H. McCallon, S. Schneider, R. Stiening, M. Sykes, M. Weinberg, W. A. Wheaton, S. Wheelock, N. Zacarias, "2MASS All-Sky Catalog of Point Sources", *University of Massachusetts and Infrared Processing and Analysis Center, (IPAC/California Institute of Technology),* 2003.

[36] S. Veilleux, D.-C. Kim, D.B. Sanders, "Optical and Near-Infrared Imaging of the IRAS 1 Jy Sample of Ultraluminous Infrared Galaxies. II. The Analysis", *The Astrophysical Journal Supplement Series*, vol. 143, no. 2, pp. 315–376, 2002.

[37] A. Knyazyan, A. Mickaelian, H. Astsatryan, "Astronomical Catalogs Cross-correlation Objectives and Illustration of a New Correlation Algorithm", *Proc. Conf. Computer Science and Information Technologies*, 2011.

[38] V. Beckmann, C.R. Shrader, "Active Galactic Nuclei", *Wiley-VCH Verlag GmbH, ISBN-13: 978-3527410781,* 350 pages, 2012.

[39] D.B. Sanders, J.M. Mazzarella, D.-C. Kim, J.A. Surace, B.T. Soifer, "The IRAS Revised Bright Galaxy Sample", *The Astronomical Journal*, vol. 126, pp. 1607–1664, 2003.

[40] E. Bertin, M. Dennefeld, M. Moshir, "Galaxy evolution at low redshift? II. Number counts and optical identifications of faint IRAS sources", *Astronomy & Astrophysics*, vol. 323, pp. 685–696, 1997.

[41] M.A. Strauss, J.P. Huchra, "The Distribution of IRAS Galaxies Towards the Boötes Void", *Astronomical Journal*, vol. 95, pp. 1602–1618, 1988.

[42] G. Wang, S.K. Leggett, R.G. Clowes, H.T. MacGillivray, and A. Savage, "Optical identifications of IRAS point sources: the Fornax, Hydra I and Coma clusters", *Monthly Notices of the Royal Astronomical Society,* vol. 248, pp. 112–127, 1991.

[43] K.B. Fisher, J.P. Huchra, M.A. Strauss, M. Davis, A. Yahil, D. Schlegel, "The IRAS 1.2 Jy Survey: Redshift Data", *The Astrophysical Journal Supplement Series*, vol. 100, pp. 69–103, 1995.

[44] Q.R. Yuan, Z.H. Zhu, Z.L. Yang, X.T. He, "Optical identifications of the IRAS faint sources in the Virgo cluster area", *Astronomy & Astrophysics Supplement*, vol. 115, pp. 267–282, 1996.

[45] B. E. Markarian, V. A. Lipovetsky, J. A. Stepanian, L. K. Erastova, A. I. Shapovalova, "The first Byurakan survey. A catalogue of galaxies with UV-continuum", *Communications of the Special Astrophysical Observatory*, vol. 62, pp. 5–117, 1989.

[46] L. Fullmer, C.J. Lonsdale, "Cataloged galaxies and quasars observed in the IRAS survey", *JPL D-1932, Pasadena: Jet Propulsion Laboratory*, 1989.

[47] R.W. Sinnott, "NGC 2000.0: The Complete New General Catalogue and Index Catalogues of Nebulae and Star Clusters by J. L. E. Dreyer", *Cambridge: Sky Publishing Corporation and Cambridge University Press,* 1988.

[48] G. de Vaucouleurs, A. de Vaucouleurs, H.G. Corwin, R.J. Buta, G. Paturel, P. Fouque, "Third Reference Cat. of Bright Galaxies (RC3)", *VizieR On-line Data Catalog: VII/155. Originally published in: Springer-Verlag: New York*, 1995.

[49] M.-P. Véron-Cetty, P. Véron, "A catalogue of quasars and active nuclei: 13th edition", *Astronomy & Astrophysics*, vol. 518, id A10, 8 pp., 2010.

[50] E. Massaro, A. Maselli, C. Leto, P. Marchegiani, M. Perri, P. Giommi, S. Piranomonte, "The 5th edition of the Roma-BZCAT. A short presentation", *Astrophysics and Space Science*, vol. 357, id 75, 4 pp., 2015.

[51] J. Souchay, C. Gattano, A. H. Andrei, D. Souami, B. Coelho, C. Barache, F. Taris, N. Secrest, A. Berthereau, "LQAC-5: The fifth release of the Large Quasar Astrometric Catalogue. A compilation of 592 809 objects with 398 697 Gaia counterparts", *Astronomy & Astrophysics,* vol. 624, id A145, 14 pp., 2019.

[52] B. W. Lyke, A. N. Higley, J. N. McLane, D. P. Schurhammer, A. D. Myers, A. J. Ross, K. Dawson, S. Chabanier, P. Martini, N. G. Busca, H. du Mas des Bourboux, M. Salvato, A. Streblyanska, P. Zarrouk, E. Burtin, S. F. Anderson, J. Bautista, D. Bizyaev, W. N. Brandt, J. Brinkmann, J. R. Brownstein, J. Comparat, P. Green, A. de la Macorra, A. Muñoz Gutiérrez, J. Hou, J. A. Newman, N. Palanque-Delabrouille, I. Pâris, W. J. Percival, P. Petitjean, J. Rich, G. Rossi, D. P. Schneider, A. Smith, M. Vivek, B. A. Weaver, "The Sloan Digital Sky Survey Quasar Catalog: Sixteenth Data Release", *The Astrophysical Journal Supplement Series*, vol. 250, id 8, 24 pp., 2020.

[53] G. T. Richards, A. D. Myers, C. M. Peters, C. M. Krawczyk, G. Chase, N. P. Ross, X. Fan, L. Jiang, M. Lacy, I. D. McGreer, J. R. Trump, R. N. Riegel, "Bayesian High-redshift Quasar Classification from Optical and Mid-IR Photometry", *The Astrophysical Journal Supplement Series,* vol. 219, id 39, 21 pp., 2015.

[54] D. S. Aguado, R. Ahumada, A. Almeida, S. F. Anderson, B. H. Andrews, B. Anguiano, E. Aquino Ortíz, A. Aragón-Salamanca, M. Argudo-Fernández, M. Aubert, V. Avila-Reese, C. Badenes, S. Barboza Rembold, K. Barger, J. Barrera-Ballesteros, D. Bates, J. Bautista, R. L. Beaton, T. C. Beers, F. Belfiore, M. Bernardi, M. Bershady, F. Beutler, J. Bird, D. Bizyaev, G. A. Blanc, M. R. Blanton, M. Blomqvist, A. S. Bolton, M. Boquien, J. Borissova, J. Bovy, W. N. Brandt, J. Brinkmann, J. R. Brownstein, K. Bundy, A. Burgasser, N. Byler, M. Cano Diaz, M. Cappellari, R. Carrera, B. Cervantes Sodi, Y. Chen, B. Cherinka, P. D. Choi, H. Chung, D. Coffey, J. M. Comerford, J. Comparat, K. Covey, G. da Silva Ilha, L. da Costa, Y. S. Dai, G. Damke, J. Darling, R. Davies, K. Dawson, V. de Sainte Agathe, A. Deconto Machado, A. Del Moro, N. De Lee, A. M. Diamond-Stanic, H. Domínguez Sánchez, J. Donor, N. Drory, H. du Mas des Bourboux, C. Duckworth, T. Dwelly, G. Ebelke, E. Emsellem, S. Escoffier, J. G. Fernández-Trincado, D. Feuillet, J.-L. Fischer, S. W. Fleming, A. Fraser-McKelvie, G. Freischlad, P. M. Frinchaboy, H. Fu, L. Galbany, R. Garcia-Dias, D. A. García-Hernández, L. A. Garma Oehmichen, M. A. Geimba Maia, H. Gil-Marín, K. Grabowski, M. Gu, H. Guo, J. Ha, E. Harrington, S. Hasselquist, C. R. Hayes, F. Hearty, H. Hernandez Toledo, H. Hicks, D. W. Hogg, K. Holley-Bockelmann, J. A. Holtzman, B.-C. Hsieh, J. A. S. Hunt, H. S. Hwang, H. J. Ibarra-Medel, C. E. Jimenez Angel, J. Johnson, A. Jones, H. Jönsson, K. Kinemuchi, J. Kollmeier, C. Krawczyk, K. Kreckel, S. Kruk, I. Lacerna, T.-W. Lan, R. R. Lane, D. R. Law, Y.-B. Lee, C. Li, J. Lian, L. Lin, Y.-T. Lin, C. Lintott, D. Long, P. Longa-Peña, J. T. Mackereth, A. de la Macorra, S. R.

Majewski, O. Malanushenko, A. Manchado, C. Maraston, V. Mariappan, M. Marinelli, R. Marques-Chaves, T. Masseron, K. L. Masters, R. M. McDermid, N. Medina Peña, S. Meneses-Goytia, A. Merloni, M. Merrifield, S. Meszaros, D. Minniti, R. Minsley, D. Muna, A. D. Myers, P. Nair, J. Correa do Nascimento, J. A. Newman, C. Nitschelm, M. D. Olmstead, A. Oravetz, D. Oravetz, R. A. Ortega Minakata, Z. Pace, N. Padilla, P. A. Palicio, K. Pan, H.-A. Pan, T. Parikh, J. Parker III, S. Peirani, S. Penny, W. J. Percival, I. Perez-Fournon, T. Peterken, M. H. Pinsonneault, A. Prakash, M. J. Raddick, A. Raichoor, R. A. Riffel, R. Riffel, H.-W. Rix, A. C. Robin, A. Roman-Lopes, B. Rose, A. J. Ross, G. Rossi, K. Rowlands, K. H. R. Rubin, S. F. Sánchez, J. R. Sánchez-Gallego, C. Sayres, A. Schaefer, R. P. Schiavon, J. S. Schimoia, E. Schlafly, D. Schlegel, D. P. Schneider, M. Schultheis, H.-J. Seo, S. J. Shamsi, Z. Shao, S. Shen, S. Shetty, G. Simonian, R. J. Smethurst, J. Sobeck, B. J. Souter, A. Spindler, D. V. Stark, K. G. Stassun, M. Steinmetz, T. Storchi-Bergmann, G. S. Stringfellow, G. Suárez, J. Sun, M. Taghizadeh-Popp, M. S. Talbot, J. Tayar, A. R. Thakar, D. Thomas, P. Tissera, R. Tojeiro, N. W. Troup, E. Unda-Sanzana, O. Valenzuela, M. Vargas-Magaña, J. A. Vázquez-Mata, D. Wake, B. A. Weaver, A.-M. Weijmans, K. B. Westfall, V. Wild, J. Wilson, E. Woods, R. Yan, M. Yang, O. Zamora, G. Zasowski, K. Zhang, Z. Zheng, Z. Zheng, G. Zhu, J. C. Zinn, H. Zou, "The Fifteenth Data Release of the Sloan Digital Sky Surveys: First Release of MaNGA-derived Quantities, Data Visualization Tools, and Stellar Library", *The Astrophysical Journal Supplement Series,* vol. 240, id 23, 25 pp., 2019.

[55] R. Ahumada, C. Allende Prieto, A. Almeida, F. Anders, S. F. Anderson, B. H. Andrews, B. Anguiano, R. Arcodia, E. Armengaud, M. Aubert, S. Avila, V. Avila-Reese, C. Badenes, C. Balland, K. Barger, J. K. Barrera-Ballesteros, S. Basu, J. Bautista, R. L. Beaton, T. C. Beers, B. I. T. Benavides, C. F. Bender, M. Bernardi, M. Bershady, F. Beutler, C. M. Bidin, J. Bird, D. Bizyaev, G. A. Blanc, M. R. Blanton, M. Boquien, J. Borissova, J. Bovy, W. N. Brandt, J. Brinkmann, J. R. Brownstein, K. Bundy, M. Bureau, A. Burgasser, E. Burtin, M. Cano-Díaz, R. Capasso, M. Cappellari, R. Carrera, S. Chabanier, W. Chaplin, M. Chapman, B. Cherinka, C. Chiappini, P. D. Choi, S. D. Chojnowski, H. Chung, N. Clerc, D. Coffey, J. M. Comerford, J. Comparat, L. da Costa, M.-C. Cousinou, K. Covey, J. D. Crane, K. Cunha, G. da Silva Ilha, Y. S. Dai, S. B. Damsted, J. Darling, J. W. Davidson Jr., R. Davies, K. Dawson, N. De, A. de la Macorra, N. De Lee, A. B. de Andrade Queiroz, A. Deconto Machado, S. de la Torre, F. Dell'Agli, H. du Mas des Bourboux, A. M. Diamond-Stanic, S. Dillon, J. Donor, N. Drory, C. Duckworth, T. Dwelly, G. Ebelke, S. Eftekharzadeh, A. D. Eigenbrot, Y. P. Elsworth, M. Eracleous, G. Erfanianfar, S. Escoffier, X. Fan, E. Farr, J. G. Fernández-Trincado, D. Feuillet, A. Finoguenov, P. Fofie, A. Fraser-McKelvie, P. M. Frinchaboy, S. Fromenteau, H. Fu, L. Galbany, R. A. Garcia, D. A. García-Hernández, L. A. Garma Oehmichen, J. Ge, M. A. Geimba Maia, D. Geisler, J. Gelfand, J. Goddy, V. Gonzalez-Perez, K. Grabowski, P. Green, C. J. Grier, H. Guo, J. Guy, P.

Harding, S. Hasselquist, A. J. Hawken, C. R. Hayes, F. Hearty, S. Hekker, D. W. Hogg, J. A. Holtzman, D. Horta, J. Hou, B.-C. Hsieh, D. Huber, J. A. S. Hunt, J. Ider Chitham, J. Imig, M. Jaber, C. E. Jimenez Angel, J. A. Johnson, A. M. Jones, H. Jönsson, E. Jullo, Y. Kim, K. Kinemuchi, C. C. Kirkpatrick IV, G. W. Kite, M. Klaene, J.-P. Kneib, J. A. Kollmeier, H. Kong, M. Kounkel, D. Krishnarao, I. Lacerna, T.-W. Lan, R. R. Lane, D. R. Law, J.-M. Le Goff, H. W. Leung, H. Lewis, C. Li, J. Lian, L. Lin, D. Long, P. Longa-Peña, B. Lundgren, B. W. Lyke, J. T. Mackereth, C. L. MacLeod, S. R. Majewski, A. Manchado, C. Maraston, P. Martini, T. Masseron, K. L. Masters, S. Mathur, R. M. McDermid, A. Merloni, M. Merrifield, S. Mészáros, A. Miglio, D. Minniti, R. Minsley, T. Miyaji, F. G. Mohammad, B. Mosser, E.-M. Mueller, D. Muna, A. Muñoz-Gutiérrez, A. D. Myers, S. Nadathur, P. Nair, K. Nandra, J. Correa do Nascimento, R. J. Nevin, J. A. Newman, D. L. Nidever, C. Nitschelm, P. Noterdaeme, J. E. O'Connell, M. D. Olmstead, D. Oravetz, A. Oravetz, Y. Osorio, Z. J. Pace, N. Padilla, N. Palanque-Delabrouille, P. A. Palicio, H.-A. Pan, K. Pan, J. Parker, R. Paviot, S. Peirani, K. P. Ramŕez, S. Penny, W. J. Percival, I. Perez-Fournon, I. Pérez-Ràfols, P. Petitjean, M. M. Pieri, M. Pinsonneault, V. J. Poovelil, J. T. Povick, A. Prakash, A. M. Price-Whelan, M. J. Raddick, A. Raichoor, A. Ray, S. B. Rembold, M. Rezaie, R. A. Riffel, R. Riffel, H.-W. Rix, A. C. Robin, A. Roman-Lopes, C. Román-Zúñiga, B. Rose, A. J. Ross, G. Rossi, K. Rowlands, K. H. R. Rubin, M. Salvato, A. G. Sánchez, L. Sánchez-Menguiano, J. R. Sánchez-Gallego, C. Sayres, A. Schaefer, R. P. Schiavon, J. S. Schimoia, E. Schlafly, D. Schlegel, D. P. Schneider, M. Schultheis, A. Schwope, H.-J. Seo, A. Serenelli, A. Shafieloo, S. J. Shamsi, Z. Shao, S. Shen, M. Shetrone, R. Shirley, V. Silva Aguirre, J. D. Simon, M. F. Skrutskie, A. Slosar, R. Smethurst, J. Sobeck, B. Cervantes Sodi, D. Souto, D. V. Stark, K. G. Stassun, M. Steinmetz, D. Stello, J. Stermer, T. Storchi-Bergmann, A. Streblyanska, G. S. Stringfellow, A. Stutz, G. Suárez, J. Sun, M. Taghizadeh-Popp, M. S. Talbot, J. Tayar, A. R. Thakar, R. Theriault, D. Thomas, Z. C. Thomas, J. Tinker, R. Tojeiro, H. H. Toledo, C. A. Tremonti, N. W. Troup, S. Tuttle, E. Unda-Sanzana, M. Valentini, J. Vargas-González, M. Vargas-Magaña, J. A. Vázquez-Mata, M. Vivek, D. Wake, Y. Wang, B. A. Weaver, A.-M. Weijmans, V. Wild, J. C. Wilson, R. F. Wilson, N. Wolthuis, W. M. Wood-Vasey, R. Yan, M. Yang, C. Yèche, O. Zamora, P. Zarrouk, G. Zasowski, K. Zhang, C. Zhao, G. Zhao, Z. Zheng, Z. Zheng, G. Zhu, H. Zou, "The 16th Data Release of the Sloan Digital Sky Surveys: First Release from the APOGEE-2 Southern Survey and Full Release of eBOSS Spectra", *The Astrophysical Journal Supplement Series,* vol. 249, id 3, 21 pp., 2020.

[56] G. Paturel, C. Petit, Ph. Prugniel, G. Theureau, J. Rousseau, M. Brouty, P. Dubois, L. Cambrésy, "HYPERLEDA. I. Identification and designation of galaxies", *Astronomy & Astrophysics,* vol. 412, pp. 45–55, 2003.

[57] R.G. McMahon, M.J. Irwin, S.J. Maddox, "The APM-North Catalogue", *VizieR On-line Data Catalog: I/267. Originally published in: Institute of Astronomy, Cambridge, CB3 0HA, UK,* 2000.

[58] Gaia Collaboration, "Gaia DR3 Part 2. Extragalactic", *VizieR On-line Data Catalog: I/356*, 2022.

[59] J. J. Condon, W. D. Cotton, E. W. Greisen, Q. F. Yin, R. A. Perley, G. B. Taylor, J. J. Broderick, "The NRAO VLA Sky Survey", *The Astronomical Journal*, vol. 115, pp. 1693–1716, 1998.

[60] D.J. Helfand, R.L. White, R.H. Becker, "The Last of FIRST: The Final Catalog and Source Identifications", *The Astrophysical Journal,* vol. 801, id 26, 17 pp., 2015.

[61] Gaia Collaboration, "Gaia DR3 Part 1. Main source", *VizieR On-line Data Catalog: I/355*, 2022.

[62] NASA/IPAC Extragalactic Database, *Jet Propulsion Laboratory, California Institute of Technology,* https://ned.ipac.caltech.edu/.

[63] A.M. Mickaelian, H.V. Abrahamyan, G.M. Paronyan, G.A. Mikayelyan, "Fine Classification of QSOs and Seyferts for Activity Types Based on SDSS Spectroscopy", *Frontiers in Astronomy & Space Sciences*, vol. 7, id 505043, 2021

[64] D.E. Osterbrock, "Emission-line spectra of active galactic nuclei and quasars", *Proceedings of the 9th Texas Symposium on Relativistic Astrophysics*, pp. 22–38, 1980.

[65] H. Winkler, "Variability studies of Seyfert galaxies - II. Spectroscopy", *Monthly Notices of the Royal Astronomical Society*, vol. 257, pp. 677–688, 1992.

[66] D.E. Osterbrock, R.W. Pogge, "The spectra of narrow-line Seyfert 1 galaxies", *The Astrophysical Journal*, vol. 297, pp. 166–176, 1985.

[67] S. Veilleux, D.E. Osterbrock, "Spectral classification of emission-line galaxies", *The Astrophysical Journal Supplement Series*, vol. 63, pp. 295–310, 1987.

[68] T.M. Heckman, "An Optical and Radio Survey of the Nuclei of Bright Galaxies - Activity in the Normal Galactic Nuclei", *Astronomy & Astrophysics*, vol. 87, pp. 152–164, 1980.

[69] L.C. Ho, A.V. Filippenko, W.L.W. Sargent, "Low-Luminosity Seyfert Nuclei", *Proceedings of IAU Colloquium No. 159*, pp. 429–433, 1997.

[70] D.W. Weedman, "Emission line galaxies", *Vistas in Astronomy*, vol. 21, pp. 55–70, 1977.

[71] P. Véron, A.C. Gonçalves, M.-P. Véron-Cetty, "AGNs with composite spectra", *Astronomy & Astrophysics*, vol. 319, pp. 52–66, 1997.

[72] H. Astsatryan, A. Knyazyan, A. Mickaelian, L. Sargsyan, "Web portal for the Armenian Virtual Observatory based on Armenian national grid infrastructure", *Parallel Computing and Control Problems, Moscow*, pp. 109–114, 2010.

[73] A. M. Mickaelian, H. V. Astsatryan, A. V. Knyazyan, T. Yu. Magakian, G. A. Mikayelyan, L. K. Erastova, L. R. Hovhannisyan, L. A. Sargsyan, P. K. Sinamyan, "Ten years of the Armenian Virtual Observatory", *Astronomical Surveys and Big Data: Proceedings of an International Symposium dedicated to the 50th anniversary of the Markarian Survey and the 10th anniversary of the Armenian Virtual Observatory held at Byurakan Astrophysical Observatory, Byurakan, Armenia on 5-8 October 2015. Edited by Areg M. Mickaelian, Andy Lawrence and Tigran Yu. Magakian. ASP Conference Series*, vol. 505, pp. 16–23, 2016.

[74] A.M. Mickaelian, G.A. Mikayelyan, H.V. Astsatryan, A.V. Knyazyan, "Armenian Virtual Observatory: Status and activities", *Astronomy & Computing*, vol. 42, id 100689, 2023.

[75] P. J. Quinn, D. G. Barnes, I. Csabai, C. Cui, F. Genova, B. Hanisch, A. Kembhavi, S. C. Kim, A. Lawrence, O. Malkov, M. Ohishi, F. Pasian, D. Schade, W. Voges, "The International Virtual Observatory Alliance: recent technical developments and the road ahead", *In Optimizing Scientific Return for Astronomy Through Information Technologies, Proceedings of the SPIE*, vol. 5493, pp. 137–145, 2004.

[76] R.J. Hanisch, "The Virtual Observatory: I", *Astronomy and Computing*, vol. 7, pp. 1–2, 2014.

[77] R. J. Hanisch, G. B. Berriman, T. J. W. Lazio, S. Emery Bunn, J. Evans, T. A. McGlynn, R. Plante, "The virtual astronomical observatory: re-engineering access to astronomical data", *Astronomy and Computing*, vol. 11, pp. 190–209, 2015.

[78] C.Z. Cui, Y.H. Zhao, "Worldwide R&D of virtual observatory", *Proceedings of the International Astronomical Union No. 248*, pp. 563–564, 2007.

[79] R. Hudec, "Astronomical Photographic Plate Archives", *Proceedings of Astroplate-2014 Conference*, *Prague*, pp. 1, 2014.

[80] L. Kazantseva, "Photographic collections of observatories as scientific heritage", *Proceedings of Astroplate-2014 Conference*, *Prague*, pp. 13, 2014.

[81] R. Nesci, M. Bagaglia, G. Nucciarelli, "Scientific use of digitized plates", *Proceedings of Astroplate-2014 Conference*, *Prague*, pp. 75, 2014.

[82] W. Osborn, "Utilizing Astronomy's Photographic Heritage: Progress, Problems and Challenges", *Proceedings of Astroplate-2014 Conference*, *Prague,* pp. 15, 2014.

[83] J. Stupka, E. Benesova, "Digitization as a Method of Protection and Presentation of Glass Plate Negatives", *Proceedings of Astroplate-2014 Conference*, *Prague,* pp. 31, 2014.

[84] M. Tsvetkov, K. Tsvetkova, "Wide-Field Plate Database: Latest Results", *In Proceedings of IAU Symposium 285: New Horizons in Time-Domain Astronomy, Cambridge University Press*, pp. 417–419, 2012.

[85] J.P. Huchra, "The nature of Markarian galaxies", *The Astrophysical Journal Supplement Series*, vol. 35, pp. 171–195, 1977.

[86] S. Indolia, A.K. Goswami, S. Mishra, P. Asopa, "Conceptual understanding of convolutional neural network—a deep learning approach", *Procedia Computer Science*, vol. 132, pp. 679–688, 2018.

[87] T. Kohonen, "Analysis of a simple self-organizing process", *Biological Cybernetics*, vol. 44(2), pp. 135–140, 1982.

[88] J. MacQueen, "Some methods for classification and analysis of multivariate observations", *In Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, pp. 281–297, 1967.

[89] M. Zhang, "Use density-based spatial clustering of applications with noise (DBSCAN) algorithm to identify galaxy cluster members", *IOP Conference Series: Earth and Environmental Science*, vol. 242 (4), id 042033, 2019.

[90] A. Buades, B. Coll, J.-M. Morel, "A review of image denoising algorithms, with a new one", *Multiscale Modeling & Simulation*, vol. 4 (2), pp. 490–530, 2005.

[91] T. Pavlidis, "Algorithms for shape analysis of contours and waveforms", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 2 (4), pp. 301–312, 1980.

[92] T. Pavlidis, "Algorithms for Graphics and Image Processing", *Springer Science & Business Media,* 448 pages, 2012.

[93] R. J. Hanisch, A. Farris, E. W. Greisen, W. D. Pence, B. M. Schlesinger, P. J. Teuben, R. W. Thompson, A. Warnock, "Definition of the Flexible Image Transport System (FITS)", *Astronomy & Astrophysics*, vol. 376, pp. 359–380, 2001.

[94] R.E. González, R.P. Muñoz, C.A. Hernández, "Galaxy detection and identification using deep learning and data augmentation", *Astronomy and Computing*, vol. 25, pp. 103–109, 2018.

[95] P. Arbelaez, M. Maire, C. Fowlkes, J. Malik, "Contour detection and hierarchical image segmentation", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33 (5), pp. 898–916, 2011.

[96] K. He, X. Zhang, S. Ren, J. Sun, "Deep residual learning for image recognition", *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778, 2016.

[97] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, L.-C. Chen, "MobileNetV2: Inverted residuals and linear bottlenecks", *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4510–4520, 2018.

[98] G.S. Tran, T.P. Nghiem, V.T. Nguyen, C.M. Luong, J.-C. Burie, "Improving accuracy of lung nodule classification using deep learning with focal loss", *Journal of Healthcare Engineering*, vol. 2019, id 5156416, 9 pages, 2019.

[99] D.P. Kingma, J. Ba, "Adam: A method for stochastic optimization", *arXiv:1412.6980*, 2014.

[100] S. J. Pan, Q. Yang, "A survey on transfer learning", *IEEE Transactions on Knowledge and Data Engineering*, vol. 22 (10), pp. 1345–1359, 2010.

[101] T. Carneiro, R.V. Medeiros Da NóBrega, T. Nepomuceno, G.-B. Bian, V.H.C. De Albuquerque and P.P.R. Filho, "Performance Analysis of Google Colaboratory as a Tool for Accelerating Deep Learning Applications", *in IEEE Access*, vol. 6, pp. 61677–61685, 2018.

[102] Cloud-ML, "Computational Astrophysics Library", *Available at: https://github.com/ArmHPC/Computational-Astrophysics,* 2021.

LIST OF ACRONYMS

Table A.1. Acronyms and abbreviations used in the Thesis

Acronym / label	Meaning / definition
2MASS	Two Micron All Sky Survey
AGN	Active Galactic Nucleus / Active Galactic Nuclei
AKARI	Meaning "light" in Japanese
APM	Automated Plate Measurement Machine
ArVO	Armenian Virtual Observatory
BAO	Byurakan Astrophysical Observatory
BIG	Byurakan-IRAS Galaxies
BLL	BL Lacertae object
BLS	Broad-Line Seyfert
BZCAT	Roma-BZCAT, Multi-Frequency Catalogue of Blazars
CNN	Convolutional Neural Network
Comp.	Composite spectrum galaxy
DENIS	Deep Near-Infrared Survey of the Southern Sky
DFBS	Digitized First Byurakan Survey
DR	Data Release
DR16Q	Sloan Digital Sky Survey Data Release 16 quasar catalogue
FBS	First Byurakan Survey
FIR	Far-InfraRed
FIRST	Faint Images of the Radio Sky at Twenty Centimeters survey
FIS	Far-Infrared Surveyor, an AKARI instrument
FITS	Flexible Image Transport System, the standard format used for astronomical data
FSC	Faint Source Catalogue (IRAS)
FWHM	Full Width at Half Maximum
HII	Ionized hydrogen region / HII galaxy
HLIRG	Hyper-Luminous Infrared Galaxy
HYPERLEDA	Extragalactic galaxy database used for catalogue matching and galaxy identification
IPAC	Infrared Processing and Analysis Center
IR	InfraRed
IRAS	InfraRed Astronomical Satellite
IRC	InfraRed Camera, an AKARI instrument
ISO	InfraRed Space Observatory
IVOA	International Virtual Observatory Alliance
JWST	James Webb Space Telescope
LAC	Lacertae
LINER	Low-Ionization Nuclear Emission-line Region
LIRG	Luminous InfraRed Galaxy
LQAC	Large Quasar Astrometric Catalogue
MIR	Mid-InfraRed
ML	Machine Learning
Mrk	Markarian designation
MSE	Mean Squared Error
NED	NASA/IPAC Extragalactic Database
NGC	New General Catalogue
NIR	Near-InfraRed
NLAGN	Narrow-Line Active Galactic Nucleus

Acronym / label	Meaning / definition
NLQSO	Narrow-Line Quasar
NLS	Narrow-Line Seyfert
NRAO	National Radio Astronomy Observatory
NVSS	NRAO VLA Sky Survey
OHP	Observatoire de Haute-Provence
PN	Planetary Nebula
PSC	Point Source Catalogue (IRAS)
QSO	Quasi-Stellar Object (quasar)
RC3	Third Reference Catalogue of Bright Galaxies
ResNet	Residual Network, a neural-network architecture
SAO	Special Astrophysical Observatory
SB	Starburst galaxy
sd	Subdwarf
sdA	Subdwarf A star
sdB	Subdwarf B star
sdO	Subdwarf O star
SDSS	Sloan Digital Sky Survey
SFR	Star-Formation Rate
Sy	Seyfert
ULIRG	Ultra-Luminous InfraRed Galaxy
VLA	Very Large Array
VO	Virtual Observatory
WISE	Wide-field Infrared Survey Explorer

ACKNOWLEDGEMENTS

First and foremost, I would like to express my sincere gratitude to my scientific supervisor, Dr. Areg M. Mickaelian, for his guidance, encouragement, and continued support throughout this research. His experience in infrared astronomy, optical identification, and the study of extragalactic objects has been invaluable in shaping this work. I am grateful for his advice, thoughtful discussions, and support at every stage of preparing this Thesis.

I would also like to express my sincere appreciation to Dr. Ashot A. Akopian for his valuable comments, constructive suggestions, and insightful observations, which contributed to improving the quality of this Thesis. His feedback and thoughtful remarks were greatly appreciated throughout the preparation of this work.

I would also like to thank the colleagues and staff of the V. Ambartsumian Byurakan Astrophysical Observatory for providing a supportive scientific environment and for their collegiality, discussions, and assistance. I appreciate the contributions of the researchers and collaborators whose observations, expertise, and published studies have informed the investigations presented here.

Finally, I am deeply grateful to my family and friends for their patience, understanding, encouragement, and support throughout the years of study and research. Their support has helped me meet the challenges involved in completing this Thesis.