Galaxy Morphologies in the Early Universe

A thesis
Submitted towards the partial fullfilment of
BS-MS dual degree programme
by

GADEKAR VEDANT BHAGESHWAR

DATE: 15 May 2024

under the guidance of

PROF. YOGESH WADADEKAR

NATIONAL CENTRE FOR RADIO ASTROPHYSICS

from May 2023 to Mar 2024

INDIAN INSTITUTE OF SCIENCE EDUCATION AND RESEARCH PUNE

Certificate

This is to certify that this dissertation entitled **Galaxy Morphologies in Early Universe** submitted towards the partial fulfilment of the BS-MS degree at the Indian Institute of Science Education and Research, Pune, represents original research carried out by Gadekar Vedant Bhageshwar at the National Centre of Radio Astrophysics, under the supervision of Prof. Yogesh Wadadekar during academic year May 2023 to March 2024.

Prof. Yogesh Wadadekar

Committee:

Prof. Yogesh Wadadekar

Prof. Susmita Adhikari

Declaration

I, hereby, declare that the matter embodied in the report titled Galaxy Morphologies in the Early Universe is the result of the investigations carried out by me at the 'National Centre for Radio Astrophysics, TIFR-Pune' from the period 15 May 2023 to 15 March 2024 under the supervision of Yogesh Wadadekar and the same has not been submitted elsewhere for any other degree.

GADEKAR VEDANT
BHAGESHWAR
20191083
BS-MS
IISER PUNE

DATE: 15/03/2024

Acknowledgments

First and foremost, I would like to express my sincere gratitude to my project supervisor, Prof. Yogesh Wadadekar, for his invaluable guidance, unwavering support, and constant encouragement throughout this endeavor.

I am deeply grateful to Prof. Susmita Adhikari, the project expert, for her invaluable contributions and insightful suggestions, which have significantly enriched the quality of this work.

I would also like to extend my heartfelt thanks to Dr. Ryan Hausen for his generous assistance and guidance during the project. His expertise and willingness to share his expertise have been invaluable assets in overcoming the challenges encountered along the way.

I am indebted to Rashi Jain, a PhD scholar at NCRA, for providing a sample of galaxies and visual classifications that I could use to test my model.

My sincere appreciation goes to my family and friends for their unconditional love, encouragement, and unwavering belief in me. Their constant support has been a source of strength and motivation throughout this journey.

Contents

List of Figures

List of Tables

Abbreviations

Abbreviations used in the Table

1. **NS**: Number of Sources
2. **CS**: Number of Correct Sources
3. **IS** : Number of Incorrect Sources
4. **PC**: Percentage of Correct Sources
5. **PI**: Percentage of Incorrect Sources

Abstract

This study investigates the morphological classifications of galaxies in the early universe by employing the capabilities of Morpheus, a deep learning framework, in conjunction with the UNCOVER and CANDELS datasets from the James Webb Space Telescope (JWST) and the Hubble Space Telescope (HST). The primary objective was to automatically classify galactic sources into distinct morphological classes, which would aid future spectroscopic studies and theoretical advancements in understanding galaxy evolution. The initial phase involved replicating existing Morpheus findings on GOODS-S data, thereby validating the model's reliability. Subsequent testing on UNCOVER data, particularly utilizing the JWST F444W band, demonstrated the model's effectiveness in segmenting and classifying high-redshift galaxies with an accuracy of 89.21%, although challenges arose in deblending closely situated sources. To enhance efficiency, a novel model, termed BigMorpheus, was developed, incorporating an attention mechanism that significantly improved the processing speed with an accuracy of 85%. This research highlights the potential of machine learning models, such as BigMorpheus, which can not only classify galaxies but also segment astronomical images and overcome the scalability and efficiency challenges posed by the impending influx of astronomical data from advanced telescopes.

Chapter 1

Introduction

1.0.1 Galaxies and a primer on their formation

Galaxies are fundamental building blocks of the large-scale structure of the universe. After the Big Bang, small density fluctuations in the primordial dark matter distribution led to the gravitational collapse of these overdense regions. Dark matter particles coalesced into halos that formed deep gravitational potential wells. These wells attracted baryonic gas, primarily hydrogen and helium, left over from the early universe. In regions where this gas could sufficiently cool, condense and ignite, stars were formed.

Large galaxies arose when substantial amounts of gas were converted into stars, with these stars remaining gravitationally bound in a dark matter halo. A wide array of physical processes - including gas accretion, mergers, and feedback mechanisms from supernovae and active galactic nuclei - proceeded to shape and evolve galaxies.[1] Galaxies thereby serve as unique laboratories, set up in the distant heavens, where astronomers empirically test hypotheses about the physical laws governing the structure formation of the universe. [2]

1.0.2 High Redshift Galaxies and their Importance to Astronomers

Galaxy formation started a few hundred million years after the Big Bang. Since the light from the early universe is redshifted due to cosmic expansion (the wavelength of the detected light increases and becomes ‘redder’), the earliest galaxies in the universe are also called high redshift galaxies. Studying such galaxies is of immense importance as they are one of the earliest formed gravitationally bound structures in the universe and can be used to test cosmological theories. They are mostly ‘undisturbed’ from external factors such as mergers and internal factors such as AGNs. They are progenitors of galaxies present in the local universe making them objects of interest.

1.0.3 Galaxy Morphology and its connection to Formation and Evolution

Galaxy morphology is the study of the structure, shapes, and arrangements of material within galaxies. Astronomers classify galaxies into various types by categorizing features such as the distributions of stars, interstellar gas and dust, and gravitational patterns of both luminous and dark matter. Galaxy morphology is a crucial lens through which we can understand the complex processes governing galaxy formation and evolution. Tracing its importance back to the work of Hubble in 1926, morphology reveals key insights into many factors— from initial formation conditions, dissipation, and cosmic environment to the internal dynamics and influence of supermassive black holes. The galaxies’ structural end state (morphology) encapsulates their history of mergers, accretions, and interactions with large-scale tidal fields, among other physical influences.

Understanding morphology is not just an academic exercise; it has practical implications for the broader field of astronomy. Much like how the classification of stars via colour-magnitude diagrams revolutionized our understanding of stellar evolution, a detailed taxonomy of galaxy morphologies could be a game-changer. Such classification is essential for identifying distinct galaxy populations over a range of redshifts based on their fundamental features [3]. By making these correct distinctions, we can isolate the key factors driving galaxy evolution, thereby gaining a comprehensive understanding of how galaxies have formed and are continuing to evolve. [4]

1.0.4 Motivation

The upcoming deluge of data from advanced telescopes like the James Webb Space Telescope (JWST) and the Large Synoptic Survey Telescope (LSST) presents both an opportunity and a challenge for studying galaxy morphologies. JWST's capabilities in sensitive infrared imaging and multi-object spectroscopy will unearth various galaxy characteristics, necessitating rapid morphological classifications for efficient spectroscopic target selection. Meanwhile, LSST's unprecedented scale—covering billions of galaxies across thousands of square degrees—demands robust analysis methods capable of scaling to this enormity. Traditional methods, even those that leverage community engagement like Galaxy Zoo, struggle to meet the scalability and efficiency required for these new astronomical frontiers. This is where machine learning, and particularly deep learning, can revolutionize our approach. We can automate the complex task of galaxy morphology classification by utilizing a deep learning framework like Morpheus for semantic segmentation. This enhances the speed and reliability of the classification process and enables us to sift through vast datasets to quickly identify and study key galactic sources in the high redshift universe. Hence, machine learning offers an indispensable toolset for navigating and extracting meaningful insights from the upcoming astronomical big data era. [5]

1.0.5 Problem Statement and its Implications

This project primarily aimed to leverage Morpheus's capabilities, a deep learning framework, for the comprehensive study of galaxy populations in the high redshift universe. Utilizing Morpheus, the project aims to automatically classify galactic sources into various classes that could serve as candidates for future studies. These classifications would provide the basis for in-depth spectroscopic analyses, supplementing theoretical advancements in our understanding of galaxy evolution. While the current version of Morpheus operates on input from four different bands, this project aimed to explore its adaptability to single-band JWST images, enhancing its utility for reliable classification. The model's adaptability to single-band images also paves the way for its application in real-time analysis during sky surveys, enabling immediate identification and flagging of objects of interest for follow-up studies. In summary, the scalability and efficiency of this machine-learning model offer a powerful tool for targeted research and large-scale astronomical surveys.

Chapter 2

Theory and Background Work

2.0.1 History and Overview of Galaxy Morphologies

Galaxy morphology has a long-standing history that predates the recognition of galaxies as extra-galactic entities. When objects now classified as galaxies were initially observed, their resolved structure clearly distinguished them from stars. Since then, structure and morphology have remained the primary means of describing and studying galaxies. Early studies of galaxy morphologies, which were based on visual inspections (qualitative), have now been extended using quantitative methods and the recent usage of AI ML techniques. [4]

2.0.2 Visual Classification Schemes

Descriptions of galaxy structure and morphology are found even before the telescope era. Astronomical investigations of galactic structures and morphologies became more feasible through the utilization of photographic techniques. Based on the work of many others, Hubble published his famous Hubble sequence in 1926. The scheme can be summarized in a diagram famously known as the Hubble Tuning Fork (2.1) that classifies galaxies into 4 broad classes - ellipticals, lenticulars, spirals and irregular galaxies. This methodology effectively classified nearby galaxies, ensuring its success[6].

In their work, de Vaucouleurs in 1959[7] and van den Bergh in 1976[8] expanded the Hubble classification system to include internal structures like bars and rings, as well as the various configurations of spiral arms. Visual morphological classification has had great success in inferring the physical properties of galaxies, as Holmberg (1958)[9] showed.

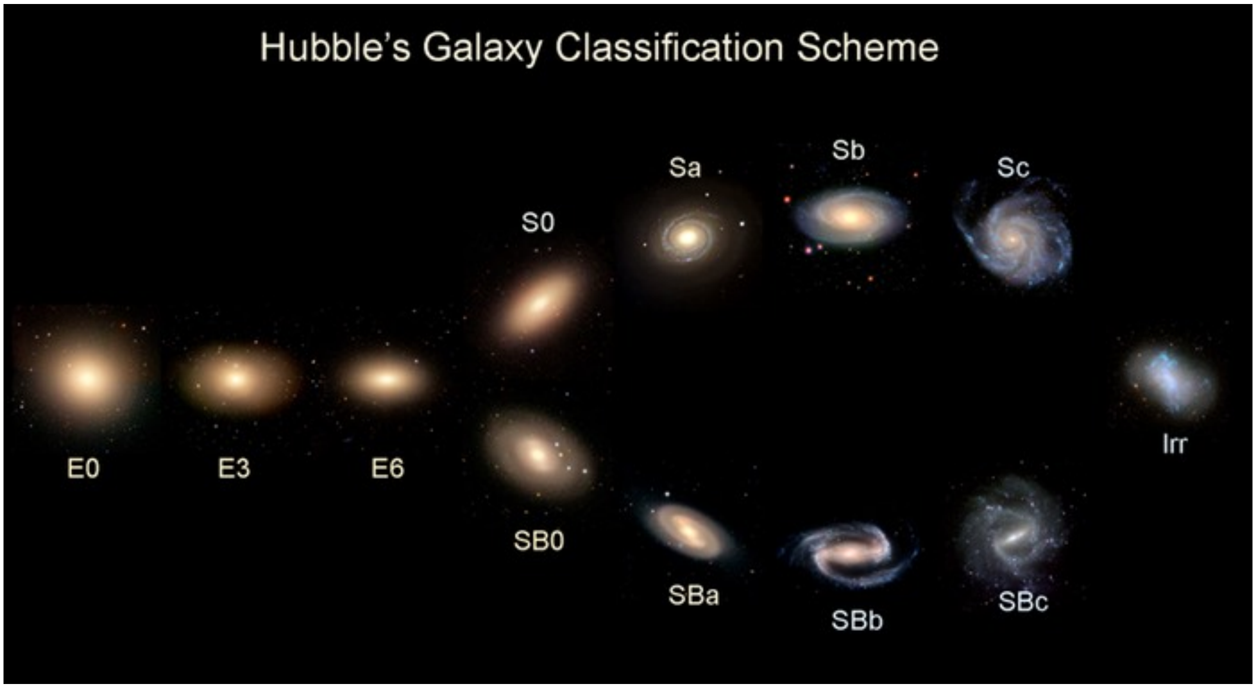


Figure 2.1: Hubble Classification Scheme. Source.

2.0.3 Parametric Classification Methods

Quantitative measurements of how light is spread across various regions in a galaxy became possible because of the adoption of CCDs (Charged Coupled Devices). De Vaucouleurs[10], in 1948, showed that ellipticals follow a universal light profile called the de Vaucouleurs profile. Sersic (1963)[11] later was able to unify the light distributions of ellipticals and spirals using a more general light profile.

$$I(R) = I_0 \times exp(-b(n)) \times [\left(\frac{R}{R_e}\right)^{1/n} - 1]$$

The Sersic index, denoted as n, characterizes the profile's shape, while the parameter $b(n)$ ensures that R_e represents the effective radius, encompassing half the galaxy's luminosity and varies as a function of n. It is established that a Sersic index of $n = 4$ corresponds to the elliptical galaxies, whereas $n = 1$ aligns with disk galaxies. In this sense, measuring the Sersic index can give an idea about the morphologies of galaxies.

During the late twentieth century, decomposition methods that distributed light within galaxies into various components such as bulge, disk, and rings became widely used. Galaxies having disks or spiral arms have most of their light from blue young stars in the disk and present lower values of B/T ratio, which measures what fraction of a galaxy's total light comes from the bulge. Ellipticals, because of the presence of larger bulges, have higher B/T ratios.

Based on assumed profiles like de Vaucouleurs or Sérsic, parametric models for galaxy light profile decomposition are ill-suited for high-redshift galaxies ($z > 1$). These models are designed for well-behaved bulge+disk systems akin to nearby galaxies. Still, high-z galaxies exhibit irregular, clumpy morphologies shaped by active star formation and mergers, distorting their light distributions beyond the capabilities of azimuthally averaged models. Compounding issues are the low spatial resolution and signal-to-noise of high-redshift observations, hampering reliable model fitting. While useful locally, parametric methods fail to capture high-redshift galaxy morphologies meaningfully, necessitating non-parametric, model-independent approaches.

2.0.4 Non-Parametric Classification Methods

Non-parametric methods try to capture the most essential features of the structure of a galaxy's light distribution, but they do not assume or model the light distribution. Since galaxy morphologies start becoming clumpy and irregular with increasing redshift, as shown by HST observations [12], these parameters are excellent in capturing the features across all redshifts.

CAS and Gini/M_{20} parameters are the most commonly used methods for measuring a galaxy's structure. Below, I describe the CAS and Gini/M_{20} parameters.

CAS parameters

CAS Parameters were introduced in various studies by Conselice[3] and Lotz[13] in 2000s. The parameters use Petrosian radius as a characteristic size of galaxies as it is less affected by the noise and surface brightness dimming.

1. **Concentration Index (C) :** The concentration of light is employed as a methodology to quantify the distribution of light within a galaxy, distinguishing the central region from the outer regions. Various approaches exist for determining the concentration of light in galaxies. These approaches include computing the ratios of radii that contain specific portions of the galaxy's luminosity and calculating the ratio between the intensities of light at two separate radii.(e.g., Bershady et al[14] ; Graham et al. 2005[15]). Luminosity within a specified Petrosian radius determines the radii used in the calculations. The most popular way sets these radii as two circular boundaries that contain internal and peripheral portions. (typically 20% and 80% or 30% and 70%) (r_{inner}, r_{outer}) of the total galaxy flux.

   $$C = 5 \times (r_{inner}/r_{outer})$$

   A higher value of C indicates that a larger proportion of the galaxy's luminosity is concentrated within a central region. In poor observational conditions, calculating concentration with different radii and areas can produce different results. (e.g., Graham et al[16]; Graham et al[15])

2. **Asymmetry Index (A) :** Asymmetry index (A), quantifies the degree of asymmetry in a galaxy after rotating it by 180 degrees along its line-of-sight center axis (2.2). This metric serves as a measure of the proportion of a galaxy's light that originates from asymmetrical structures. Asymmetry Index (A) is calculated as :

$$A = \min\left(\frac{\sum|I_o - I_{180}|}{\sum|I_o|}\right) - \min\left(\frac{\sum|B_o - B_{180}|}{\sum|I_o|}\right)$$

Where I_0 represents the initial image of the galaxy, and I_{180} is the image rotated by 180 degrees from its centre. Since the image of a galaxy will also contain background noise, it is treated in a similar manner as the original flux. Region B_0 used in the equation can be considered as an empty region of the sky surrounding the galaxy. The measurement process also involves finding a global minimum of the asymmetry value as the centre of the galaxy is unknown. Hence, the process is carried out iteratively until a minimum is found. Ellipticals have an extremely small asymmetry value between $A \sim 0.02 \pm 0.02$, while spiral galaxies have a larger value from $A \sim 0.07 - 0.2$. The value for irregular galaxies is obviously higher and is shown by measuring the index for Ultra-Luminous Infrared Galaxies (ULIRGs), which are often mergers (Conselice 2003[3]).

3. **Clumpiness (S) :** The clumpiness parameter (S) quantifies the amount of light originating from clumpy components with higher spatial frequencies within a galaxy. Galaxies in the high-redshift regime, characterized by clumpy star formation, tend to exhibit higher clumpiness values. The clumpiness parameter is calculated as:

$$S = 10 \times \left[\left(\frac{\sum(I_{x,y} - I^{\sigma}_{x,y})}{\sum I_{x,y}}\right) - \left(\frac{\sum(B_{x,y} - B^{\sigma}_{x,y})}{\sum I_x}\right)\right]$$

Where $I_{x,y}$ represents the original galaxy image, and $I^{\sigma}_{x,y}$ is a secondary image obtained by smoothing the original image with a kernel dependent on the galaxy's radius. To isolate the component of light present in high-frequency structures, a secondary image is created using a smoothing kernel that depends on the radius of the galaxy. Leftover flux image containing high-frequency structures are obtained from subtracting the observed galaxy image from the secondary smoothed image. Notably, the central regions of galaxies are excluded from this analysis, as they often harbour unresolved bright sources. Smoother sources, such as elliptical galaxies, exhibit lower clumpiness values, whereas spiral and star-forming galaxies tend to have higher clumpiness values.

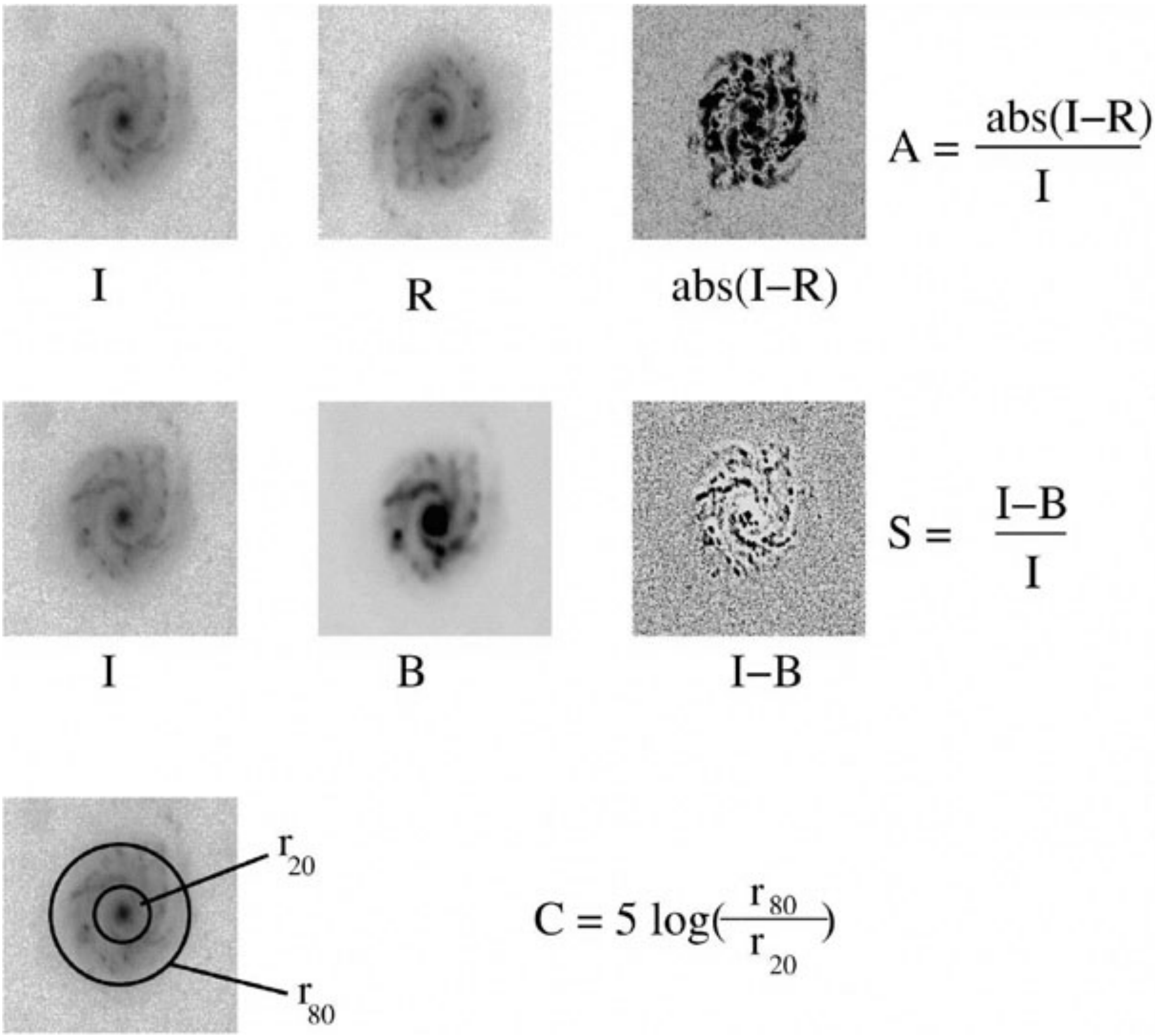


Figure 2.2: Description of Non-Parametric Methods. Source : [4].

Gini/M_{20} Parameters

These parameters focus on the relative distribution of light across pixels and have extensively been used in the literature to identify galaxies that have undergone mergers. They are not sensitive to background noise as they dont involve subtraction.

1. **Gini Coefficient (G) :** The Gini coefficient, a parameter initially introduced by Lotz (2004)[13],

is employed to quantify the degree of equity in the distribution of light among the pixels of a galaxy image. This coefficient assumes a value of 0 when the distribution is perfectly uniform, with each pixel contributing equally to the galaxy's total light. Conversely, a value of 1 signifies a scenario where the entirety of the galaxy's light is concentrated within a single pixel.

To calculate the Gini coefficient, the pixels comprising the galaxy image, denoted by 'n,' are ordered according to their brightness, and a cumulative distribution is subsequently generated. Representing the flux from each pixel as f_i, the formula for the Gini coefficient is expressed as:

$$G = \frac{1}{|\bar{f}|n(n-1)} \times \sum_i (2i - n - 1) \times f_i$$

Where the pixels are ranked in increasing order of flux, such that $f_1 \geq f_2 \ldots \geq f_n$ and $\bar{f}$ is the average pixel flux value.

2. M_{20} **parameter (**M_{20}**)** parameter, like the concentration parameter, offers a quantitative estimate indicating the degree of light concentration in an image. A highly negative value reflects an accumulation of light in a galaxy; however, this does not mean that the accumulation is at the galaxy's centre, since the light could be densely packed in any region within it. The M_{20} parameter is computed from the moment of the most luminous 20% of the galaxy's pixels, subsequently normalized by the overall moment of light spanning the entire galaxy (Lotz et al. [13], [17]). The M_{20} index is calculated as

$$M_{20} = \log 10(\frac{\sum_i M_i}{M_{tot}}),\ while \sum_i f_i < 0.2 f_{tot}$$

where M_{tot} is defined as

$$M_{tot} = \sum_{i}^{n} M_i = \sum_{i}^{n} f_i \left[(x_i - x_c)^2 + (y_i - y_c)^2 \right]$$

here, x_c and y_c denote the centre of the galaxy, and in the case of M_{20}, this centre is defined as the location where the value of M_{tot} is minimized (Lotz et al. 2004[13]). The separation between nearby ellipticals, spirals, and ULIRGs obtained using M_{20} is comparable to that found by the CAS parameters (see Lotz et al. [13], [17]).

2.0.5 Machine Learning Techniques for studies in Galaxy Morphology Classifications

Machine learning techniques, particularly artificial neural networks, have gained significant traction in astronomical research over the past two decades. The initial applications of neural networks in this domain were focused on star-galaxy classification (Odewahn et al. [18]; Bertin [19]) and the classification of galaxy spectra (Folkes et al.[20]), followed by their utilization for photometric redshift estimation (Firth et al.[21]; Collister and Lahav [22]). The appeal of artificial neural networks in astronomy lies in their ability to learn and distinguish between features in an automated manner.

Much of the earlier work in this field was limited by the availability of small datasets. However, the advent of the Galaxy Zoo project and the Sloan Digital Sky Survey (SDSS) (Lintott et al.[23]; York et al.[24]) revolutionized the domain by providing large-scale datasets necessary for conducting supervised learning algorithms. Support vector machines (SVMs) have also been extensively employed by researchers in this area (Huertas-Company et al. [25]).

Mairal et al.[26] pioneered the use of convolutional neural networks (CNNs), where the values in the convolutional kernels encode the features and are learned during the training process. CNNs have demonstrated remarkable success in galaxy morphological classification studies. Initial work by Dieleman [27] showcased the application of CNNs for classifying the Galaxy Zoo catalogue with an astonishing accuracy of 99%. Shortly thereafter, Huertas-Company et al. [28] employed CNNs to classify galaxies in five CANDELS fields with 99% accuracy. Numerous subsequent studies have leveraged CNNs for galaxy classification tasks, including radio galaxies (Aniyan and Thorat [29]; Alhassan et al.[30]), the work of Domínguez Sánchez et al.[31], and deblending and classifying galaxies (Burke et al.[32]).

Katebi et al.[33] introduced the use of capsule networks, which are adept at understanding higher-level features like CNNs but also capable of comprehending the lower-level features that constitute these higher-level features. In 2020, Hausen and Robertson[5] developed a machine learning model named 'Morpheus' that employs a UNet architecture to perform semantic segmentation on astronomical images and subsequently aggregates the classification to generate object-level classifications, deblended segmentation maps, and value-added catalogues. More recently, Robertson in 2023[34] introduced an attention-based fast semantic segmentation model, termed 'Big Morpheus,' which was applied to James Webb Space Telescope (JWST) images, maintaining the accuracy of the previous model while accelerating the computational process.

Chapter 3

Methods

3.0.1 General Overview

The initial phase of the project involved replicating the findings of Hausen [5] on a smaller dataset to gain a comprehensive understanding of the operational mechanics of Morpheus and develop the necessary code for training a semantic segmentation model. Subsequently, the currently publicly accessible version of Morpheus was evaluated using a sample obtained from Data Release 1 of the JWST UNCOVER program. Following this evaluation, the focus shifted towards BigMorpheus, a model demonstrated to work for JWST images by Robertson[34] but has not yet been released to the public. Although minor modifications were made, the same architecture was utilized to train a model and test it against a smaller yet more robust sample obtained from Data Release 2 of the JWST UNCOVER program.

3.0.2 Replication of GOODS-S results

Hausen [5] describes the development of Morpheus and the results of testing the UNet architecture on mosaics of 5 fields from HST data. To assess the efficacy and reliability of Morpheus, an independent run of the model was conducted on images from the GOODS South field, which is part of the broader Hubble Legacy Fields data release. The primary objectives were twofold: first, to compare the performance of Morpheus against the results obtained by the model's authors, and second, to familiarize oneself with the intricate details and nuances of how the model performs under different conditions.

The initial step involved downloading the GOODS South field images in $F125W$, $F160W$, $F606W$, and $F850LP$ bands. As these images covered various sky regions and were captured in different bands, selecting areas that overlapped across these bands was essential. This was particularly important because the original classification done by the authors was based on sources present in all four bands under consideration. Once the appropriate image areas were identified and cropped from the mosaics, the next step was to mask these images. Masking was necessary to ensure that the classification would focus solely on the sources present in all four image bands, thereby aligning the run with the conditions under which the authors' classification was performed.

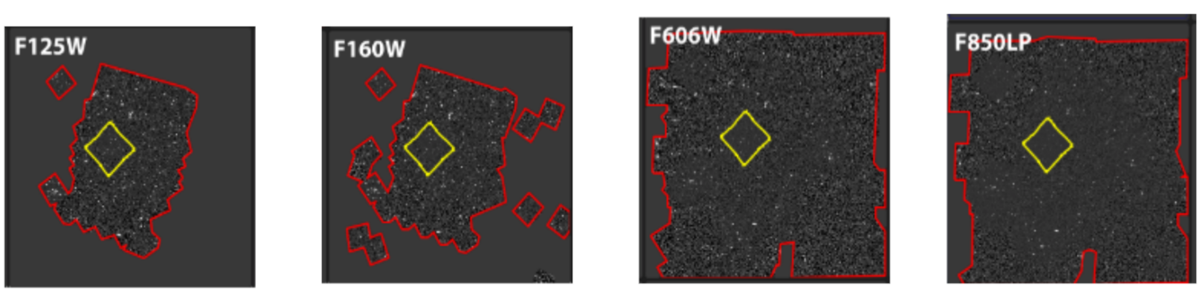
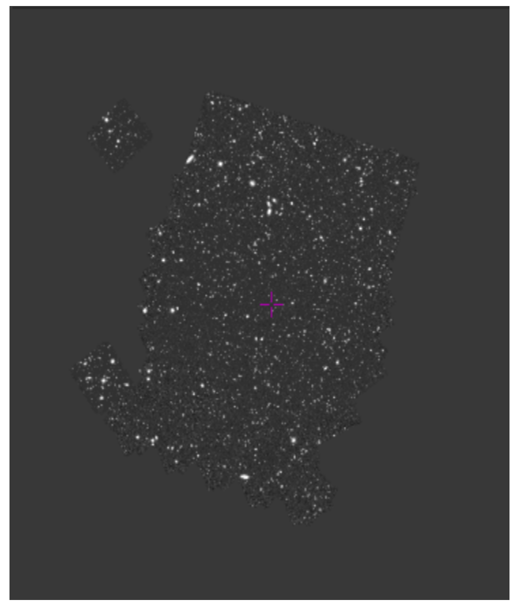


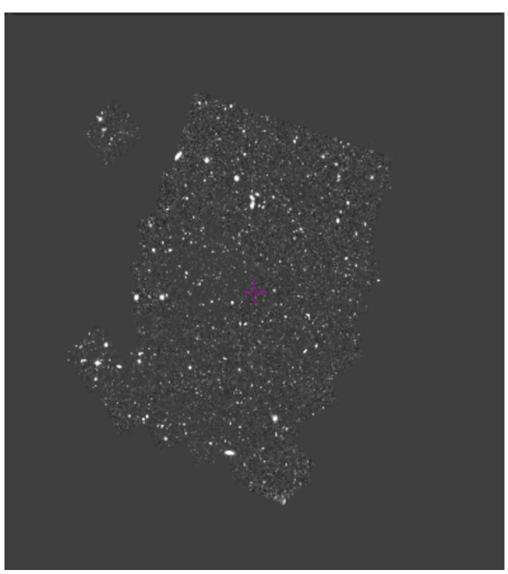
Figure 3.1: Footprints of different bands showcasing the difference in their coverage.

For masking, a simple operation was performed where all pixels from F125W that had no sources or had a pixel value equal to zero were set to zero in all other band images. This ensured that only those sources would be present in all four bands that were originally present in $F125W$. A publicly available version of Morpheus was used to classify the now cropped and masked GOODS-S images. The results from this run were then checked against those from the authors' run.

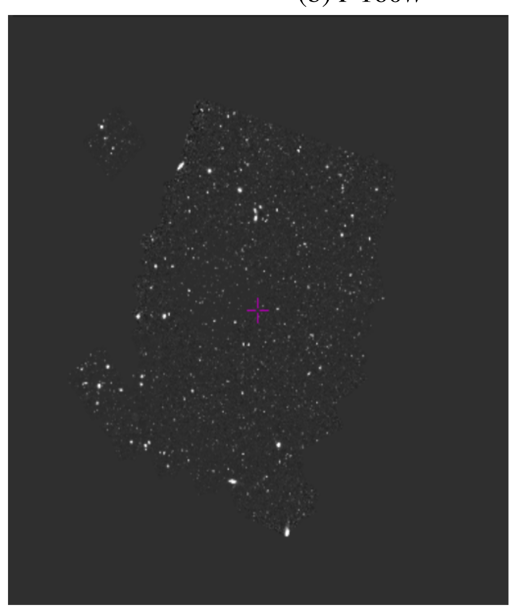
(a) $F125W$



(b) $F160W$



(c) $F606W$



(d) $F850LP$

Figure 3.2: Mosaics from all the bands after masking.

3.0.3 Testing publicly available version of Morpheus on UNCOVER Data

To evaluate the applicability of the Morpheus model beyond its original training set, which consists of HST CANDELS data, its performance was tested on images from the JWST UNCOVER Program. A sample of galaxies was selected from the Photometric Catalogue that was released as part of Data Release 1 of the UNCOVER program.

1. Galaxies with redshift between 3 and 6. The redshifts were obtained by using EAZY software.

2. Remove sources that were flagged as having unreliable photometry.

3. Remove sources flagged as having SNR (Signal to Noise Ratio) less than 3 in $F444W$ band.

This provided a robust and large sample of 1020 galaxies to test the model. Three-arcsecond cutouts from the F444W band were generated around these coordinates to ensure they were larger than 40 pixels in height and width, as required by Morpheus. Using these cutouts, they were passed on to Morpheus. It is noteworthy that since Morpheus requires four images for a source, the single cutout was passed on 'four' times to Morpheus. The selected sample was subjected to visual classification to benchmark the model's performance.

3.0.4 Training BigMorpheus

Since BigMorpheus has not yet been publicly released, it was necessary to construct a model that performs semantic segmentation. Here, we first explain the machine learning concepts used in Hausen[5] & Robertson[34] and then later outline how our version of BigMorpheus was trained.

Supervised Machine Learning Algorithms

Supervised learning involves training a machine to learn the relationship between inputs and the desired outputs using a set of examples. Subsequently, the machine can predict the output for unseen inputs. This approach differs from model-fitting algorithms, where the model is predefined. In supervised learning, the machine constructs the model based on the provided examples, enabling it to learn and model complex relationships between inputs and outputs, thereby offering superior flexibility compared to predefined model approaches.

The supervised learning algorithm comprises three main stages. During the training process, a loss function motivated by the final task is defined, and the model aims to minimize this loss function, thereby finding a relationship between inputs and outputs. The machine receives feedback on its predicted relationship after every batch of training examples, and it adjusts the model parameters to minimize the loss.

The validation phase involves optimizing the hyperparameters, which are parameters that modify the configurations and structure of the model, and evaluating the model's performance. During the test phase, the trained model is applied to previously unseen data, and the results are compared to other relevant benchmarks specific to the domain.

Evaluation metrics play a crucial role in supervised learning, as interpreting a model's performance and assessing whether the model has indeed understood the task at hand can be challenging. Evaluation metrics track a model's performance and provide insights into the model's understanding of the task.[35]

Artifical Neural Networks

Neural networks are composed of neurons that are analogous to neurons in the human brain. Neural network models typically consist of three types of layers: the input layer, hidden layers, and the output layer. The input layer serves as the entry point where input data is fed in the form of vectors. The hidden layers, which can be numerous, are selected and designed based on the specific task at hand. The output layer is the final layer, where the processed inputs are operated upon to produce the desired outputs.

The input data flows through the input layer and reaches the hidden layers. Here, the input is processed by a weight matrix W and bias terms, which represent a linear combination of the input values. These linear combinations are then passed through a non-linear activation function, introducing non-linearity into the model and enabling the learning of more complex patterns and features. The equation gives the mathematical representation of this process:

$$\vec{x}_1 = f_1(W_1, \vec{x}_0)$$

where $\vec{x}_1$ is the result of applying W, the weight matrix to the input $\vec{x}_0$ and then applying f as a non-linear activation function. After passing through multiple hidden layers, let's assume three layers, the neurons in the output layer receive the processed values :

$$\vec{x}_3 = f_3(W_3, \vec{x}_2) = f_3(W_3 f_2(W_2 f_1(W_1, \vec{x}_0)))$$

In this context, the model parameters refer to the values of the weights (coefficients of the linear combinations) in the weight matrices. The model hyperparameters, on the other hand, correspond to the choices of non-linear activation functions, the number of hidden layers, and the structure of these hidden layers (e.g., the number of neurons in each hidden layer). Additional techniques, such as dropout, are applied to prevent the model from overfitting the relationship between the training inputs and outputs.[35]

Convolutional Neural Networks

Convolutional neural networks (CNNs) are a specialized type of neural network primarily employed when the inputs are in audio or visual form. CNNs are composed of two distinct layer types: convolutional layers and pooling layers.

Convolutional layers are hidden layers where convolution operations replace the traditional matrix multiplication operation used in neural networks. These layers take a stack of inputs (e.g., the colour channels of a colour image) and pass them through stacks of weight matrices, also known as filters.

$$X_n^{(l)} = f\left(\sum_{k=1}^{K} W_n^{(k,l)} * X_{n-1}^{(k)} + b_n^{(l)}\right)$$

where, $X_n^{(l)}$ is the output from the convolutional layer, $X_{n-1}^{(k)}$ is the input to the convolutional layer, $W_n^{(k,l)}$ are the weight matrices also known as filters, $*$ represents the convolution operation and f is the non linear activation function.

The depth of the first convolutional layer filter is typically set to match the input's depth. In subsequent layers, the number of filters can be tuned according to the problem at hand, becoming a hyperparameter. Generally, each filter in the model can be interpreted as being trained to recognize a particular feature from the input image. Convolutional filters excel in visual tasks due to their ability to learn local features and capture correlations between local regions in the input.

A pooling layer often follows a convolutional layer and has no trainable parameters. These layers serve to reduce the resolution of the inputs, thereby reducing the dimensionality of the input and the model. This allows the entire model to identify features at various resolutions and levels, enhancing its representational capacity. [27]

Semantic Segmentation

Having understood the basic building blocks of Morpheus, we now understand why Morpheus performs excellently in morphology classification.

Semantic segmentation is a term popularly used in computer vision applications where every pixel in a given image is assigned to a particular category.[36] Because of this pixel-level granularity, semantic segmentation is highly effective in image analysis. In Morpheus, the authors exploited this by creating pixel-level annotated datasets and using supervised learning, specifically a neural network architecture, to learn and generate classifications. The classifications for pixels associated with a source are then aggregated to create object-level classifications. Because of semantic segmentation, Morpheus can distinguish between source and non-source(background) pixels and perform source detection within an image. Morpheus can also provide segmentation maps in astronomy, which can be extremely useful and automate the process of generating value-added catalogues.

UNet Architecture

The currently public version of Morpheus uses the UNet Architecture. The following section describes the UNet Architecture.

The U-Net architecture is a powerful model specifically designed for semantic segmentation tasks, where the goal is to classify each pixel in an image into a particular category. The architecture stands out due to its unique "U" shape, which is constructed from two main parts: the contraction path (downsampling) and the expansion path (upsampling).

In the contraction path, the network captures the context of the input image through a series of convolutional and max pooling layers. This downsampling process reduces the spatial dimensions of the image while increasing the depth of the feature maps, allowing the model to learn more complex features at each layer.

The expansion path then performs the opposite operation. It progressively upsamples the feature maps, increasing their spatial resolution. During this process, the network uses skip connections, a critical component of the U-Net architecture, to concatenate feature maps from the contraction path with those of the corresponding upsampling stage. These connections enable precise localization by combining high-resolution features from the contraction path with the upsampled output, facilitating the accurate delineation of object boundaries.

This architecture is particularly suited for semantic segmentation because it efficiently balances the capture of context and the localization of features within an image. The downsampling path ensures that the network understands the overall context and structure of the image, while the upsampling path, reinforced by skip connections, ensures detailed localization of features. This combination enables the U-Net to perform exceptionally well in tasks requiring precise segmentation of objects within images, making it a popular choice for medical image analysis and other applications where accuracy in pixel-level classification is crucial. [37]

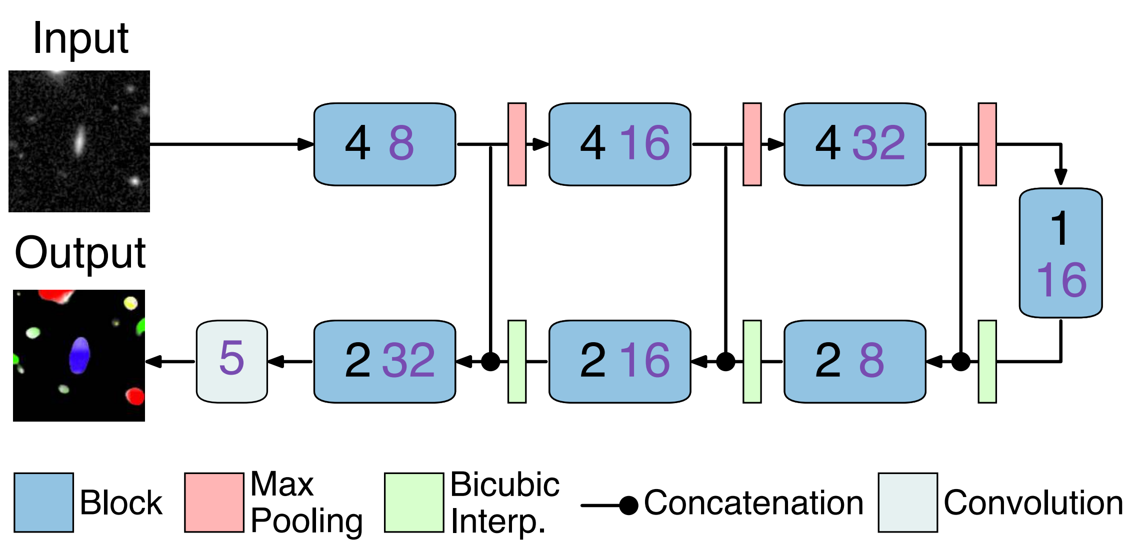


Figure 3.3: Morpheus's UNet Architecture. Source : [5].

Fast Attention Network (FANet)

Big Morpheus is based on an architecture known as Fast Attention Network (FANet). Since the architecture involves the concept of attention, I will first introduce attention and then will proceed to explain the FANet architecture.

Self Attention : Self-attention is a powerful mechanism that has revolutionized many areas of machine learning, particularly in natural language processing and computer vision tasks. At its core, self-attention allows a model to weigh different parts of the input data when computing a representation for a specific part of that data. This is achieved by calculating the relevance, or attention, between different positions in the input sequence or image.

The self-attention mechanism works by first mapping the input data (e.g., a sequence of words or an image) into three different representations: query (Q), key (K), and value (V). These representations are typically obtained by linearly projecting the input data through separate weight matrices.

1. Query (Q): The query represents the current position or element in the input data for which the model needs to compute a representation. It encodes the information about what the model should attend to or focus on.

2. Key (K): The key represents the other positions or elements in the input data that the model can potentially attend to. It encodes the information that will be used to determine the relevance or importance of each position concerning the query.

3. Value (V): The value represents the actual data or information that the model will incorporate into the representation of the query position, weighted by the attention scores.

The self-attention mechanism computes the dot product between the query and the keys and these are called as attention scores, which indicate the similarity or relevance between the query and each key position. These attention scores are then normalized using a softmax function, resulting in a probability distribution over the input positions.

The normalized attention scores are used to compute a weighted sum of the values, where each value is weighted by its corresponding attention score. This weighted sum represents the attended representation for the query position, incorporating information from all other positions in the input data, weighted by their relevance to the query.

Mathematically, the self-attention operation can be expressed as:

$$Attention(Q,K,V) = softmax((Q \times K^T)/\sqrt{d_k}) \times V$$

Where: Q, K, and V are the query, key, and value representations, d_k is the dimension of the key vectors used for scaling the dot product, and softmax normalizes the attention scores to sum to 1.

This self-attention mechanism allows the model to selectively focus on relevant parts of the input data when computing representations, capturing long-range dependencies and contextual information more effectively than traditional models like convolutional neural networks or recurrent neural networks.

In the context of semantic segmentation, self-attention can be particularly beneficial. Objects in an image can have complex shapes and interactions, requiring the model to consider contextual information beyond the local neighbourhood. Self-attention enables the model to gather relevant information from distant regions, helping it better understand the overall scene and make more accurate pixel-level predictions.

By incorporating self-attention mechanisms into neural network architectures for semantic segmentation, the model can effectively capture long-range dependencies and leverage global context, potentially leading to improved segmentation performance, especially for complex scenes and objects with intricate shapes. This capturing of correlations between parts of text or images speeds up the process of semantic segmentation. [38]

Network Architecture Let's first understand the key component of FANet, which is the Fast Attention (FA) module. Unlike standard self-attention, where the attention is computed between all pairs of spatial locations (leading to quadratic complexity), the FA module employs a more efficient approach. It encodes the input feature maps into three separate representations: a Query map (Q), a Key map (K), and a Value map (V).

The Query map represents the current position or element in the feature map for which the model needs to compute a representation. The Key map encodes the relevance or importance of each position concerning the query, and the Value map contains the actual information that the model will incorporate into the query's representation, weighted by the attention scores.

The attention scores are computed as the dot product between the Query and Key maps, which measures the similarity or relevance between the query and each key position. These attention scores are then normalized using a cosine similarity function, resulting in a probability distribution over the input positions. The normalized attention scores are used to compute a weighted sum of the Value map, where each value is weighted by its corresponding attention score. This weighted sum represents the attended representation for the query position, incorporating information from all other positions in the feature map, weighted by their relevance to the query.

Mathematically, the Fast Attention operation can be expressed as:

$$X = 1/n \times (Q \times (K^T \times V))$$

Where n is the spatial size and $K^T * V$ is computed first, reducing the computational complexity to $O(nc^2)$, where c is the channel dimension. This is significantly lower than the quadratic complexity of traditional self-attention mechanisms, making FANet more efficient and suitable for real-time applications.

The overall FANet architecture is shown in the figure. It consists of an encoder, a decoder and a Fast Attention module.

Features at different levels of complexity are identified by the encoder. These features are then processed by the Fast Attention modules, which aggregate contextual information over the spatial domain, enhancing the feature representations.

Unlike traditional architectures like UNet, which rely solely on convolutional operations and skip connections, FANet incorporates the Fast Attention modules at multiple stages of the encoder-decoder pipeline. This allows for more effective capture of long-range dependencies and spatial context, which is crucial for accurate semantic segmentation, especially in complex scenes with intricate object shapes and interactions.

Moreover, FANet employs an additional strategy of extra spatial reduction at intermediate feature stages. This involves applying additional down-sampling operations to the intermediate feature maps, which helps retain high-resolution details while reducing computational cost. This strategy is particularly beneficial for real-time applications, where efficiency is crucial.

Compared to older architectures like UNet, FANet offers several advantages:

1. Improved capture of long-range dependencies: By incorporating the Fast Attention modules, FANet can effectively model global relationships within the input data, leading to better segmentation performance, especially for complex scenes and objects with intricate shapes.

2. Computational efficiency: The Fast Attention module has a lower computational complexity than traditional self-attention, making FANet more suitable for real-time applications and deployments on resource-constrained devices.

3. Retention of high-resolution details: FANet reduces the spatial dimensions and hence is able to preserve details about how different pixels are related from the inputs and hence maintains computational efficiency, potentially improving segmentation accuracy.

4. Flexible feature reuse: FANet enables efficient feature reuse in video scenarios, which is particularly advantageous for applications involving video data, such as autonomous driving or video surveillance. [39]

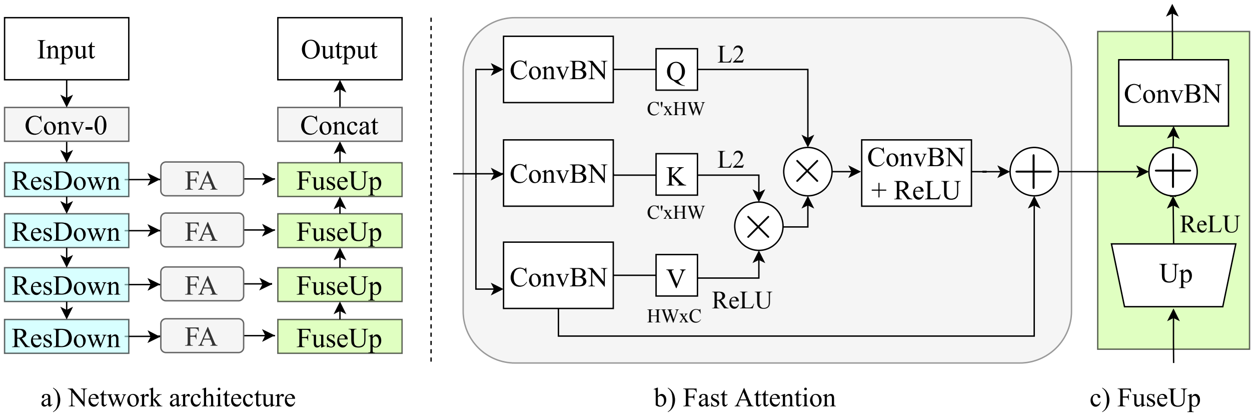


Figure 3.4: Morpheus's FANet Architecture.Source : [39]

Training our model

The input images were downloaded from the CANDELS data release at the STScI archive for the F160W band, following Robertson[34] . The annotated dataset from the Morpheus Data Release has semantic labels for all five CANDELS fields. This annotated dataset, where every pixel has a probability distribution associated with it that assigns the probability of each pixel belonging to a particular class, can be used to train a newer semantic segmentation model.

Since not all pixels in the input image have a valid annotation in the Morpheus Data Release, the images were cropped to obtain areas that have Morpheus annotations.

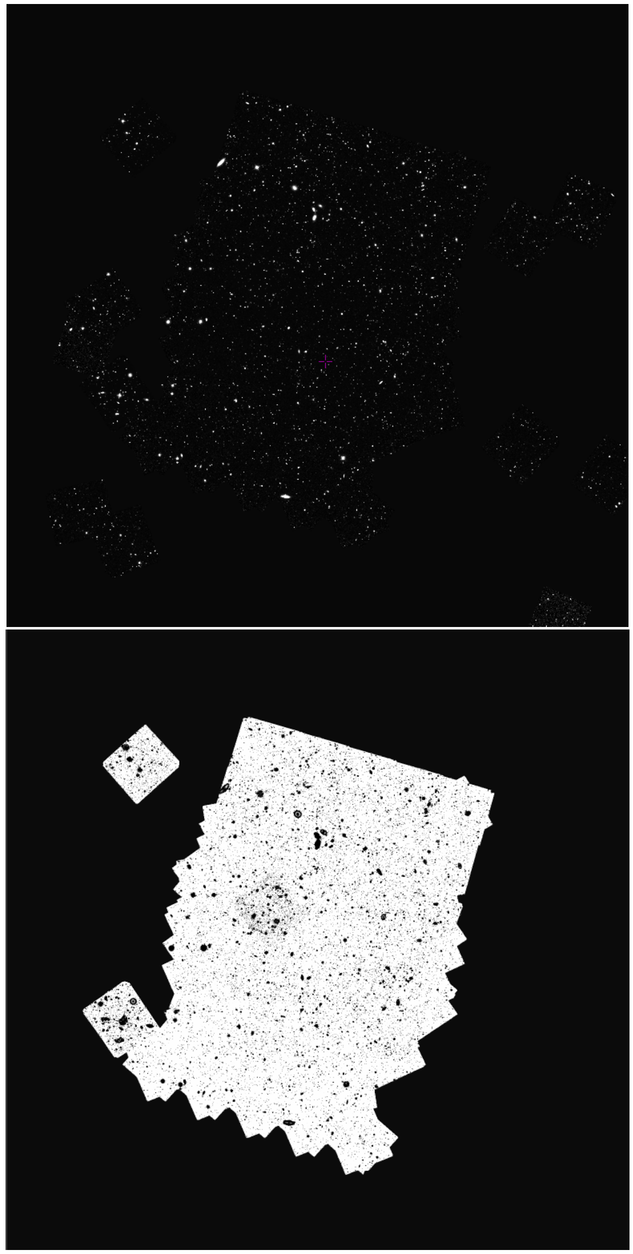


Figure 3.5: The left image is the input image, and the right image is the image output from Morpheus for the background class. As it is seen from the background image, there are areas in the input images that are not classified, and hence, we need to crop the images.

128 * 128-sized cutouts were generated from the cropped input images and the subsequent Morpheus outputs to create input-output pairs for the model. These images were then standardized using the formula provided to ensure that every source in every training image is treated on an equal footing and there is no bias towards brighter sources. During training, this standardization is done so that the mean is 0 and the variance is 1.

For each position x in the image, x becomes

$$x_{standardized} = \frac{x - mean}{max(stddev, 1.0/\sqrt{(N)})}$$

where $stddev$ is the standard deviation of all values in X and N are the number of pixels in the image.

After standardizing, 80% of all the images are then divided to form the training dataset and the rest are the part of validation dataset. The training images are then randomly shuffled and grouped into batches, and every batch is passed on through the model architecture. The loss function is calculated for a batch, and then the backpropagation proceeds to minimise it. In the FANet architecture, the convolutional filters learn features present in the image. The attention modules encode the relationship between different pixels and help in segmenting the source. The next section describes the loss function we have used in our task. The training metrics and the model performance are outlined in the Results chapter.

Loss Function A common method for training deep learning networks involves establishing a loss function that yields a statistic derived from the output classifications. This model parameters are subsequently refined through stochastic gradient descent, a process during which gradients are determined by employing the backpropagation.[40]). We use the same loss function that is mentioned in Hausen 2020 [5] except for a single modification. For our work we have used $\lambda_D = 100$ and $\lambda_w = 1$, where the λ_D and $\lambda_w = 1$ are the coefficients of the Dice and Cross Weighted Entropy Loss.

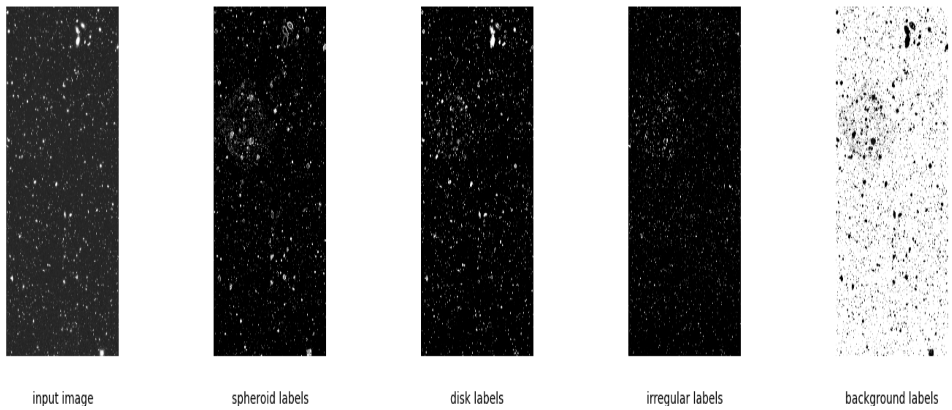


Figure 3.6: Cropped Images of GOODS South Field from which cutouts of size 128 * 128 were used for training the model.

Chapter 4

RESULTS

4.0.1 General Overview

Since we achieved 3 goals: Replicating the results of Morpheus on GOODS-S field data, Using a publicly available version of Morpheus on UNCOVER Data and retraining a newer model and again testing it on UNCOVER data, the following section will describe the results obtained in each in detail.

4.0.2 Replication of GOODS-S Results

The output of my run, as described in the Methods section, was a catalogue containing information on the dominant morphological class prediction for each source, as well as the locations of these sources within the images. To compare the results with those of the authors, a specific set of criteria was employed to match sources between the two runs.

Specifically, sources closest to those in the authors' catalogue were identified, and it was examined whether the probability of being classified into any class exceeded a predetermined threshold (70%) in both catalogs. The threshold value of 70% was arbitrarily determined, and any higher value would have been suitable.

Of all the sources identified in the images, 31.90% were deemed invalid because the class assigned with the highest probability did not exceed the 70% threshold.

A complete match—100% agreement—was found between the dominant class predictions in my run and those in the authors' run. This validated the reliability and reproducibility of the Morpheus model and enriched the understanding of its performance characteristics under different conditions.

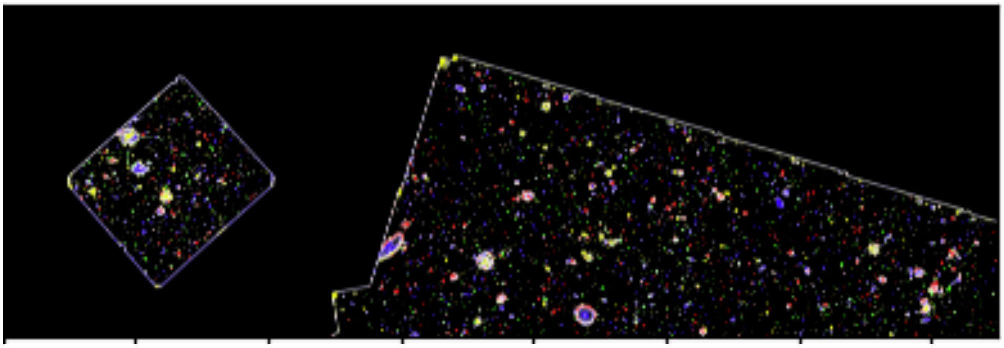


Figure 4.1: Colour classified image from my run of Morpheus .

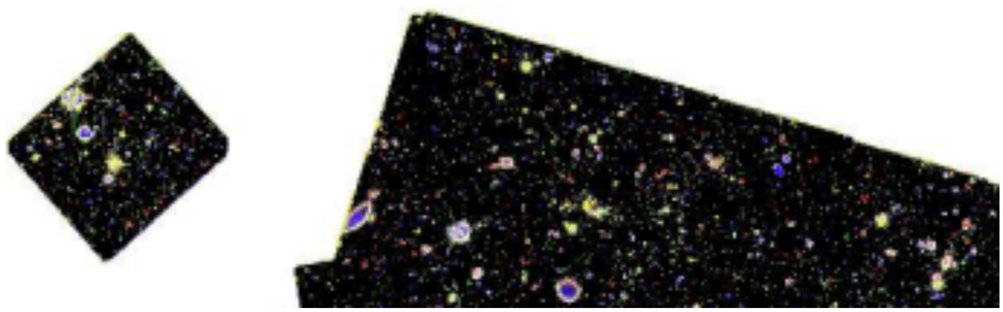


Figure 4.2: Colour classified image from author's run of Morpheus .

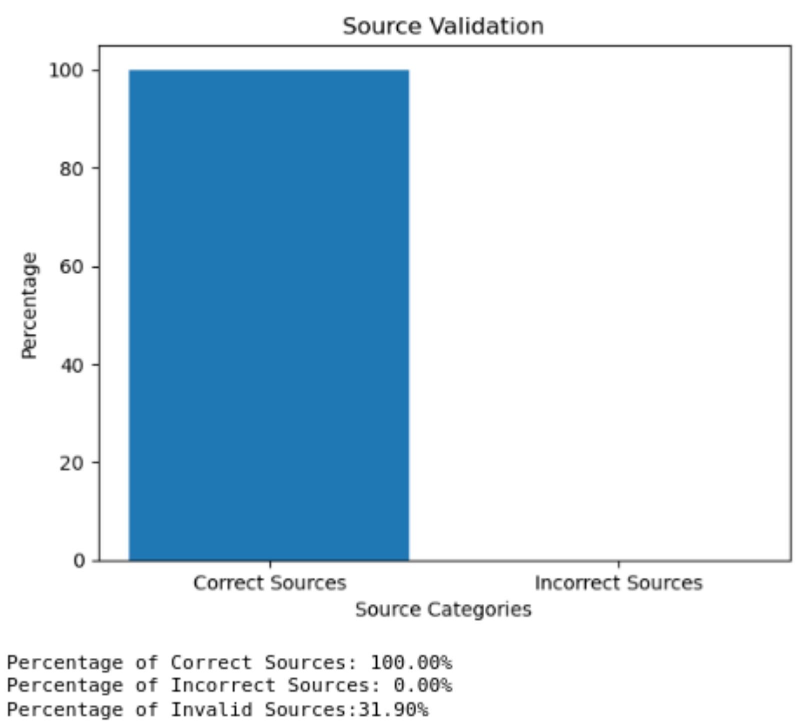


Figure 4.3: Plot outlining the comparison of results from my run and author's run.

4.0.3 Testing Morpheus on UNCOVER Data Release 1

Initially, an attempt was made to classify the JWST images by feeding Morpheus with data from four different bands, specifically $F150W$, $F277W$, $F356W$, and $F444W$.

However, the results obtained did not meet the expected level of accuracy. Additional complications arose when deciding which four bands to include in the classification process. Even after multiple iterations and combinations, the results remained below the desired level of accuracy.

This might have been due to the difference in the coverage of the filters between JWST and HST, resulting in galaxies appearing differently in various parts of the spectrum.

In a departure from this multi-band approach, Morpheus was run using the F444W band images. This yielded significantly better results in terms of both segmentation and classification. The F444W band, being the reddest and largest band used in the UNCOVER survey, effectively captured and resolved high-redshift galaxies. Therefore, it was instrumental in the successful operation of Morpheus on the JWST images.

The model's performance in classifying galaxies in the high-redshift universe was evaluated using two categories:

1. correct classification where the model's output matches visual classification.

2. incorrect classification where the model's output does not match visual classification.

The tables and figures showed that the model demonstrated high accuracy and realistically described a source. However, it faced challenges with deblending when multiple sources were close to the target, occasionally resulting in incorrect classifications.

Tables(4.1) and figures(4.4 , 4.5 , 4.6) are provided to outline the model's overall and class-wise performance.

Classes	NS	CS	IS	PC	PI
Disk	277	235	42	84.8	15.2
Spheroid	186	141	45	75.8	24.2
Irregular	308	294	14	95.5	4.5
Disk + Sph	34	32	2	94	6
Disk + Irr	120	117	3	97	3
Sph + Irr	91	87	4	95.6	4.4
Disk + Sph + Irr	4	4	0	100	0

Table 4.1: Classification Data

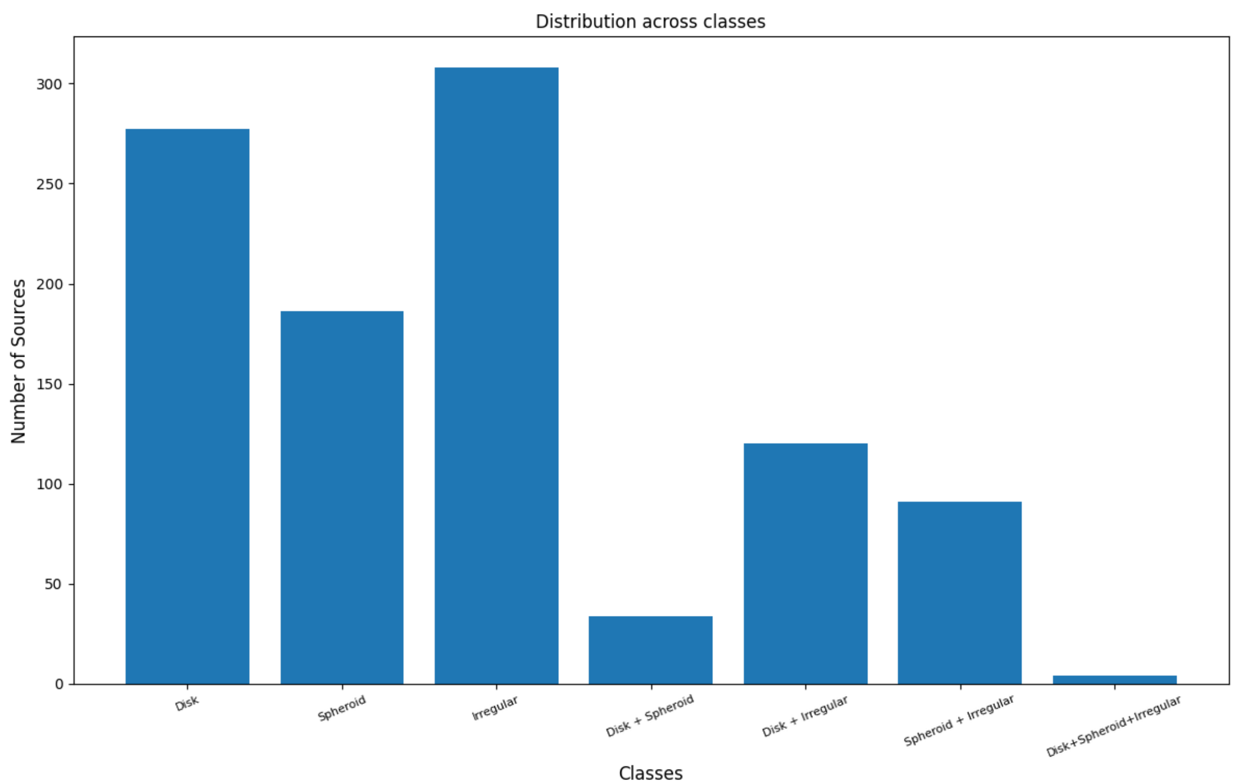


Figure 4.4: Distribution of Sources across various classes.

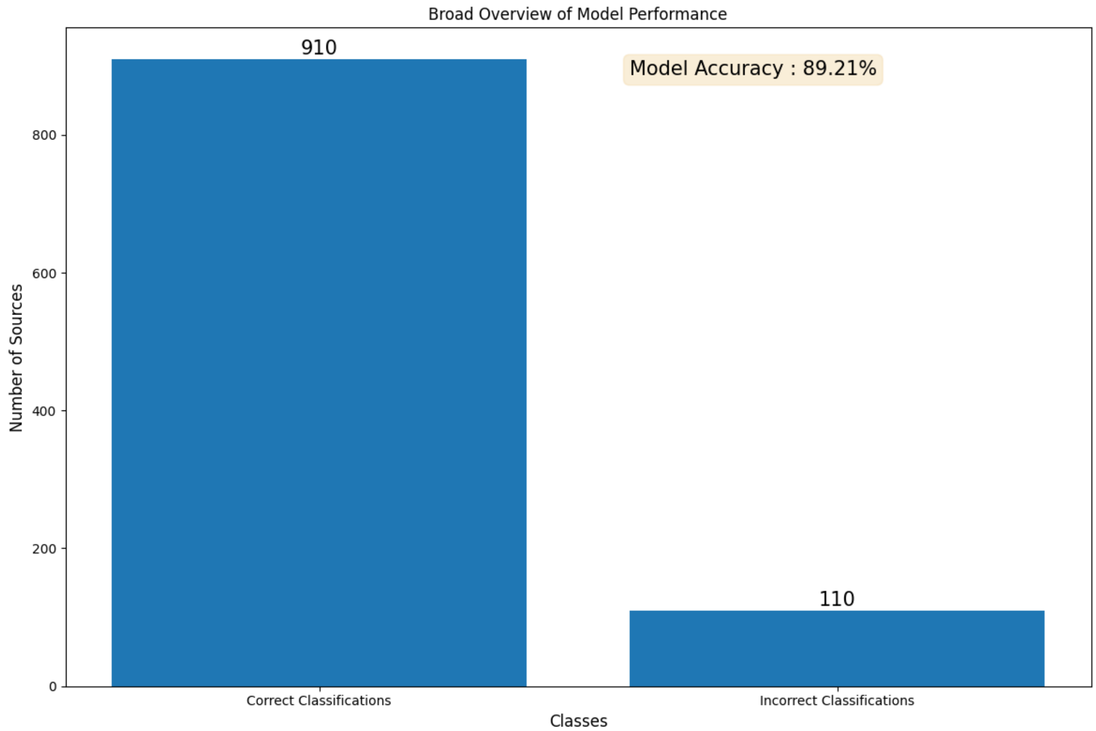


Figure 4.5: Broad overview of Model's Performance.

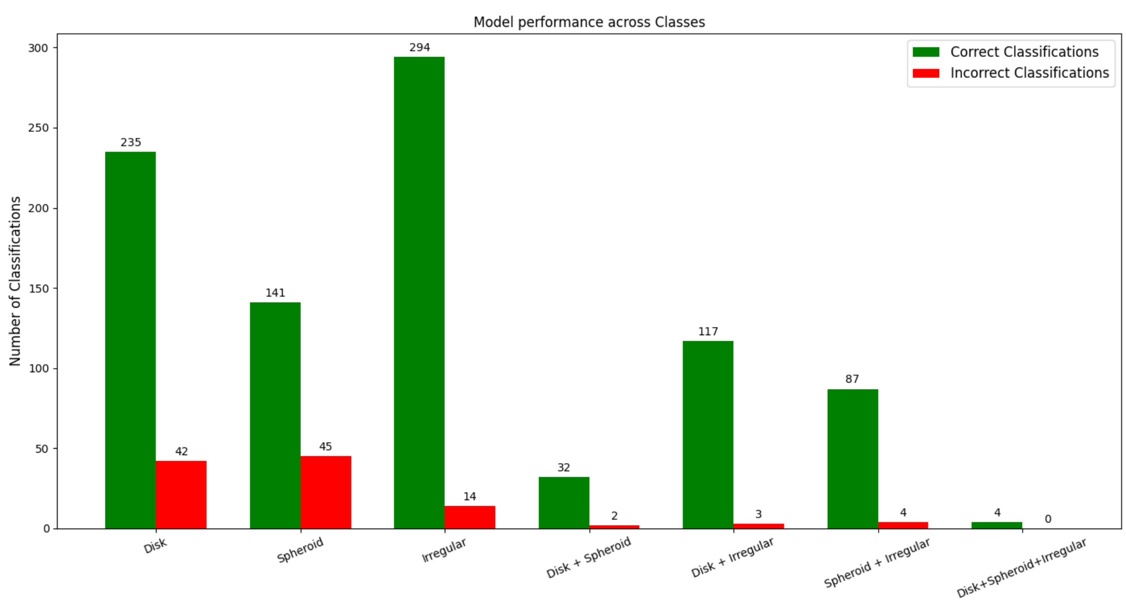


Figure 4.6: Class wise Model Performance.

4.0.4 Training and Testing BigMorpheus

Robertson [34] briefly mentions BigMorpheus, a version of Morpheus that has now included an attention mechanism and is faster and more robust in the task of Semantic segmentation.

For training a model similar to the one mentioned in Robertson [34] , which is not publicly available, data from the Morpheus Initially, an attempt was made to train the model using 512 × 512 input images, but during testing, issues arose because high-redshift galaxies appeared small in these cutouts.

Consequently, the trained model began classifying them as point sources or stars. Data Release was used, which provides semantic segmented labels (probability values for every pixel belonging to a particular class).

To solve this problem, the size of the input images was reduced to 128 × 128, and the probabilities associated with point sources were added to the spheroidal class.

This approach was reasonable because JWST cannot resolve point sources or stars in high-redshift galaxies. Since the testing sample consisted of galaxies between redshifts 3 and 6, it would not contain stars or point sources at these redshifts. Additionally, any high-redshift source that appears point-like in JWST data is most likely a galaxy whose morphology should be spheroidal. Based on this assumption, a FANet architecture, as described in the paper, was trained.

The model was trained on 24,620 and tested on 6155 128 × 128-sized cutouts for 20 epochs. A comprehensive set of performance metrics was systematically recorded throughout the model training process, portraying the progressive evolution of the model's capability to classify and segment astronomical images with increasing accuracy.

The accuracy plot exhibited an immediate ascent to high values, indicating that the model rapidly acquired the fundamental patterns necessary for effective segmentation. This initial rapid learning phase was crucial, as it helped the model separate source pixels from background pixels.

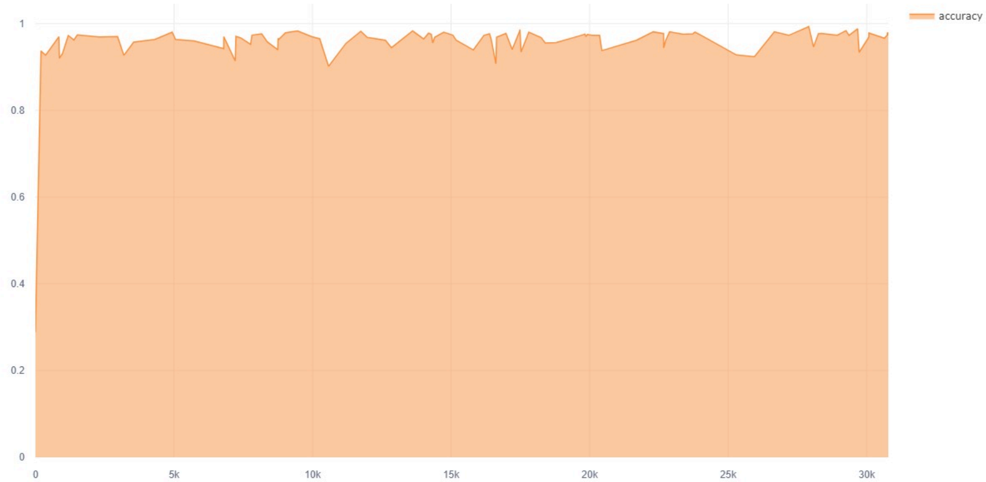


Figure 4.7: Progression of Accuracy.

As training progressed, accuracy plots for specific classes, such as background, disk, irregular, and spheroid, displayed fluctuations, reflecting the model's ongoing adjustment to distinguish between features of various morphologies.

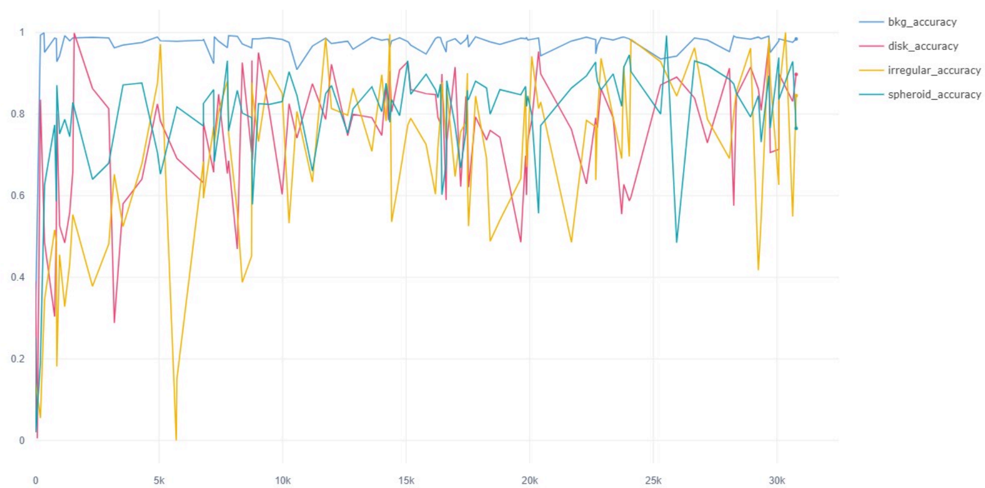


Figure 4.8: Progression of Class Wise Accuracy.

A similar trend was observed in the loss plot, which exhibited an initial steep decline, followed by a plateau, indicating that the model had begun to converge and that further improvements would be incremental. Furthermore, the cross-entropy plots, measuring the model's predictive confidence, indicated a stabilization at low levels, signifying a high level of confidence in its predictions.

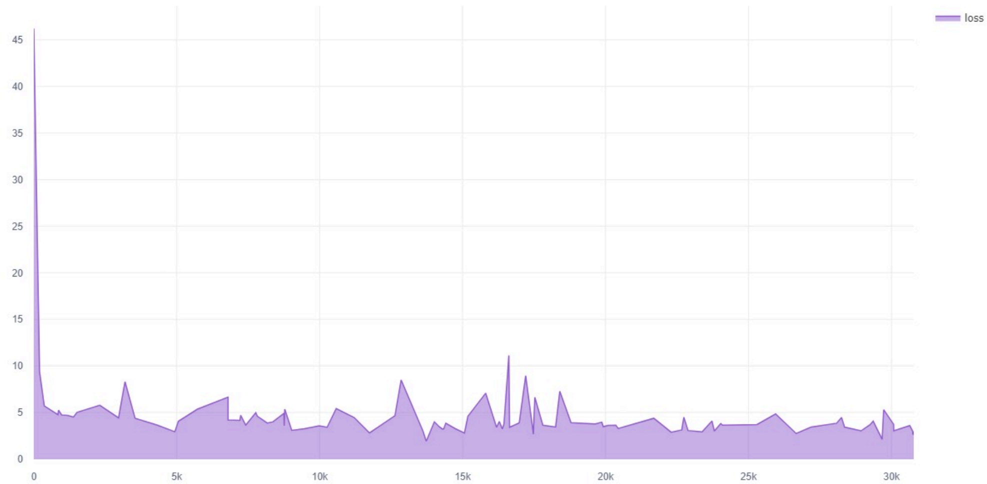


Figure 4.9: Progression of Loss.

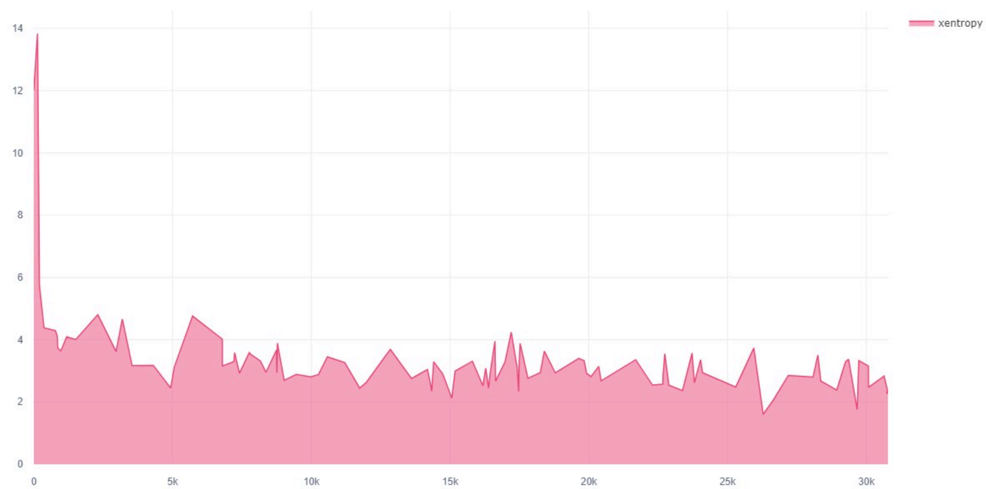


Figure 4.10: Progression of Cross Entropy Loss.

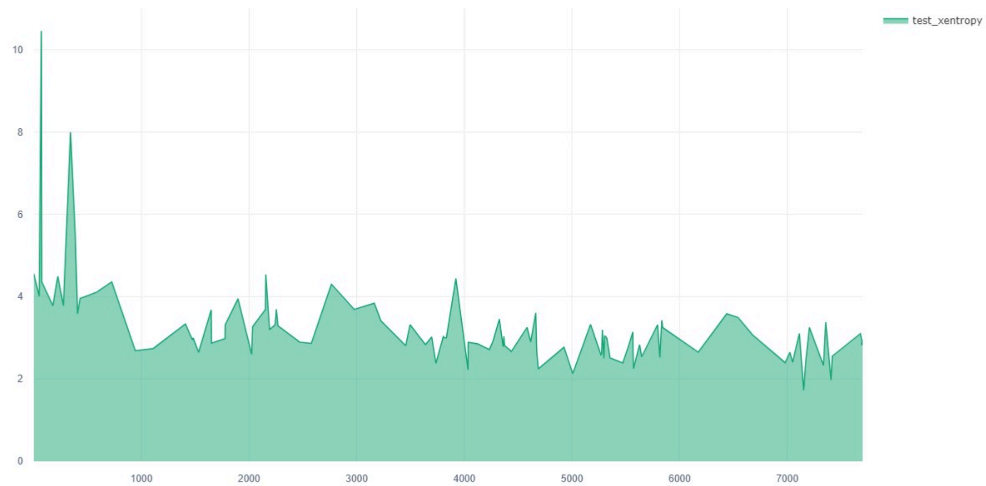


Figure 4.11: Progression of Validation Cross Entropy Loss.

Evaluations on the validation set yielded similar patterns, affirming the model's generalization ability. High validation accuracies reinforced the model's proficiency in classification, while the validation loss values aligned with the training loss, suggesting that the model did not overfit the training data and maintained its efficacy on unseen data.

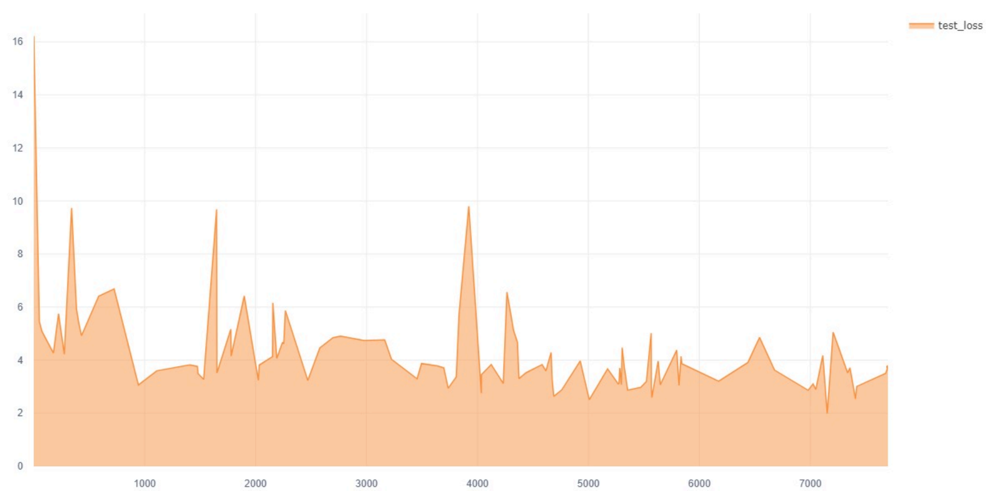


Figure 4.12: Progression of validation Loss.

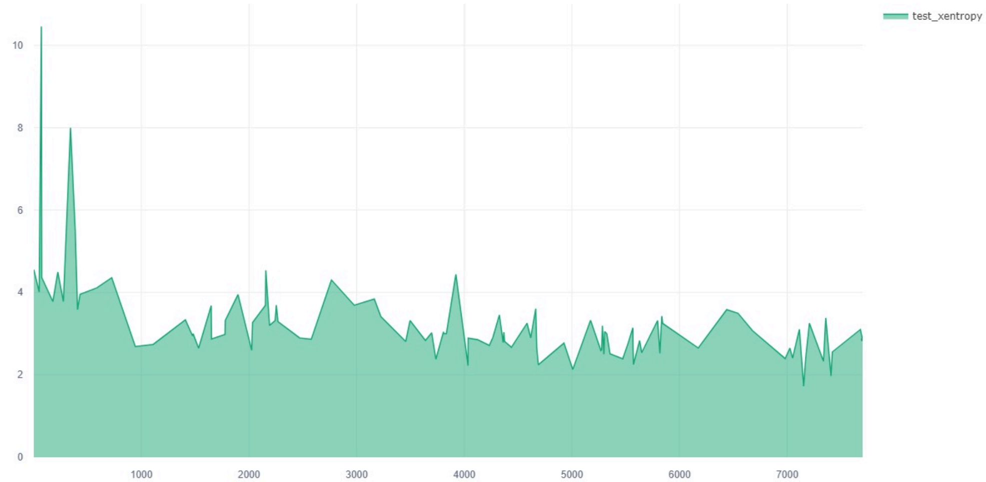


Figure 4.13: Progression of validation Cross Entropy Loss.

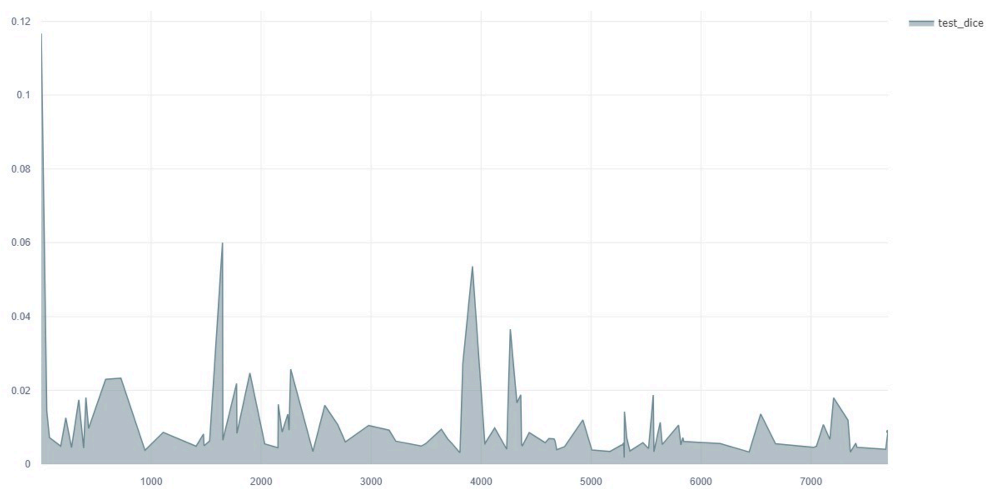


Figure 4.14: Progression of Validation Dice Loss.

After training the model, it was tested on a sample of galaxies from the UNCOVER Data Release 2. The model was tested on 674 galaxies, which were more robustly determined than the sample of 1,020 galaxies from Data Release 1. The model's performance was satisfactory, and it performed correct classifications with an accuracy of 85%. It was also faster than the UNet Architecture due to its attention mechanism.

Although in high-redshift JWST observations, some sources appeared extremely close to each other, confusing the model, as it perceives images as arrays of numbers. Due to this, the model was unable to distinguish between the local maxima of these sources, leading to issues with deblending.

Even in scenarios where deblending was an issue, the model was able to correctly classify the sources and did not flag them as irregular in nature. If two or more sources were misunderstood as one source, the model classified the morphologies of both sources correctly but was unable to segment the boundaries of both sources accurately.

Since the morphologies in high-redshift galaxies may not be regular and distinct, mixed morphologies (e.g., disk + spheroid, disk + irregular) were included in the classification scheme according to the cited reference. The model could predict one of the classes in most scenarios where a source had mixed morphologies. However, this was not an issue, as assigning a distinct morphology is also a reasonable classification. Since the model's performance was benchmarked against visual classification by a human, classifications may become subjective in some scenarios.

The tables and figures presented illustrated that the model exhibited high accuracy and provided a realistic characterization of sources. However, challenges arose when deblending multiple sources in close proximity to the target, occasionally leading to incorrect classifications.

Tables (4.2) and figures (4.15, 4.16, 4.17) were provided to outline the model's overall performance as well as its performance across different classes.

Classes	**NS**	**CS**	**IS**	**PC**	**PI**
Disk	110	70	40	63.6	36.4
Spheroid	118	102	16	86.4	13.6
Irregular	79	68	11	86.1	13.9
Disk + Sph	32	23	9	71.9	28.1
Disk + Irr	176	158	18	89.8	10.2
Sph + Irr	109	104	5	95.4	4.6
Disk + Sph + Irr	50	48	2	96	4

Table 4.2: Classification Data

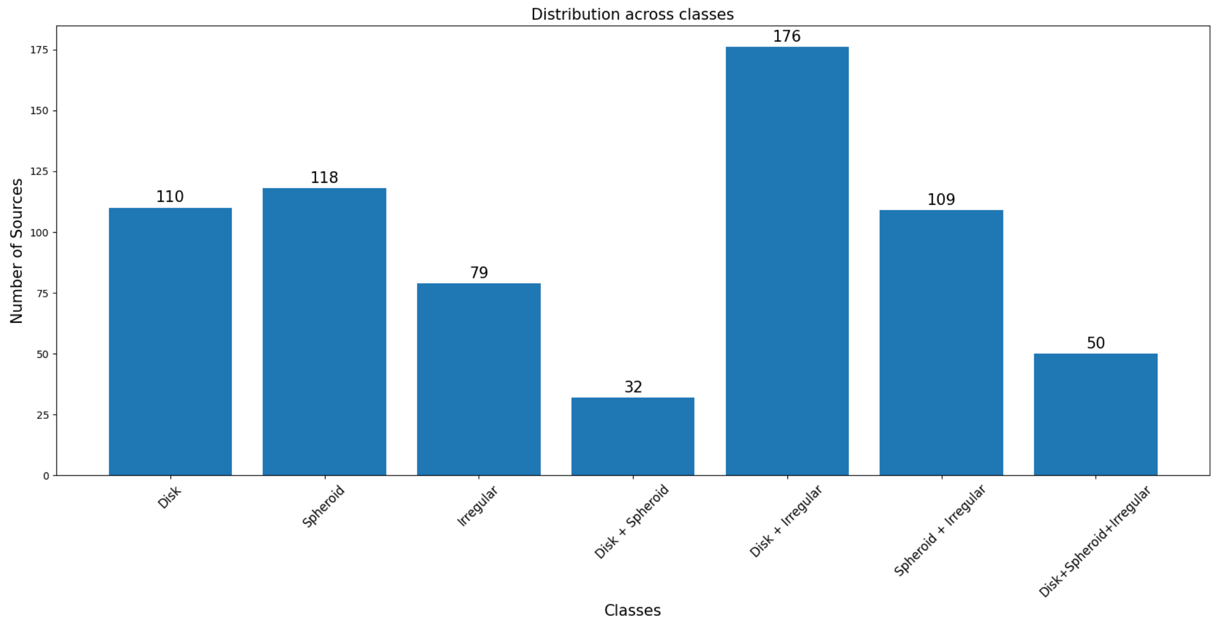


Figure 4.15: Distribution of Sources across various classes.

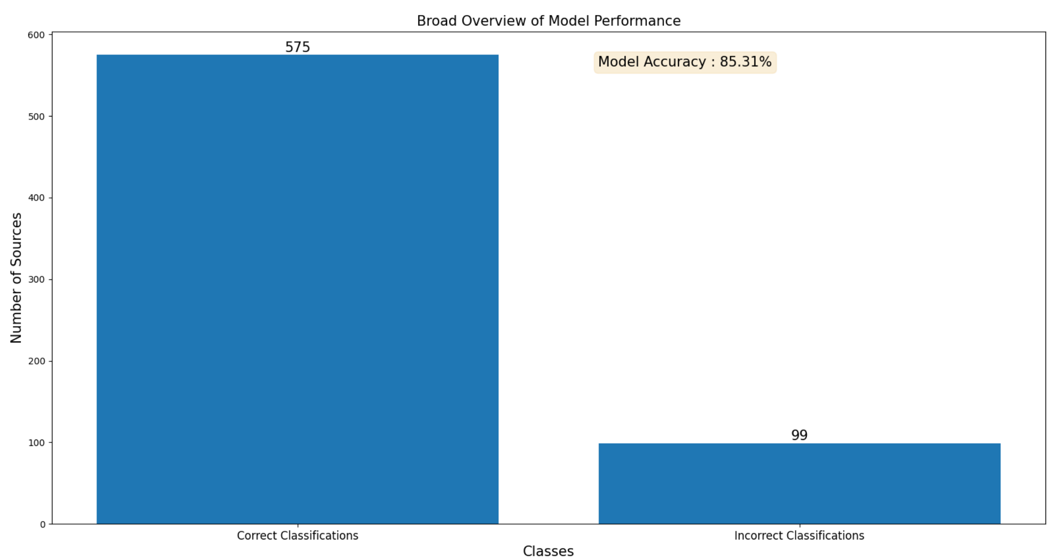


Figure 4.16: Broad overview of Model's Performance.

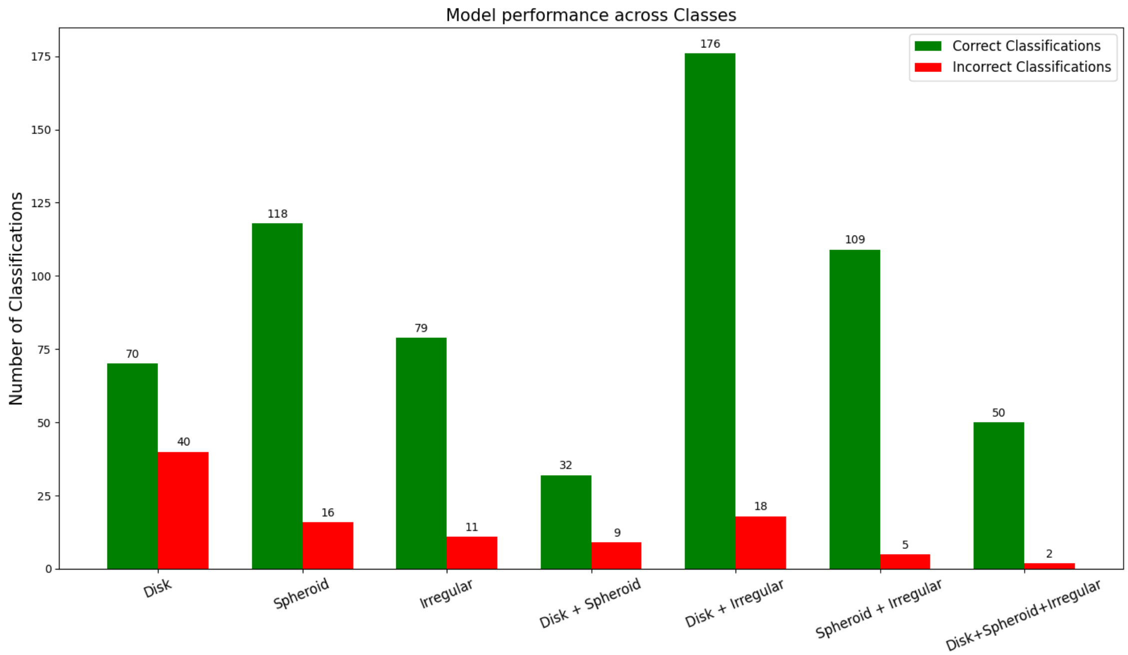


Figure 4.17: Class wise Model Performance.

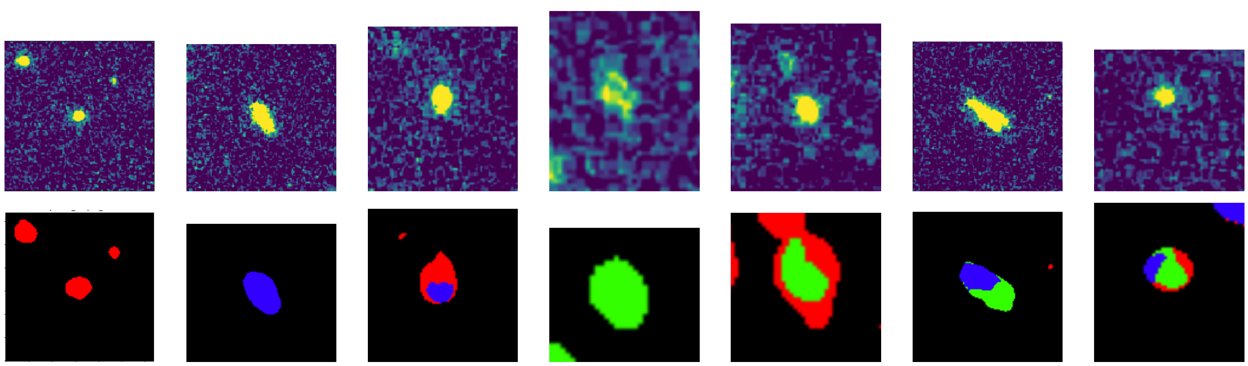


Table 4.3: Classification and Segmentation Results from Our Model. Here, the Red pixels are spheroidal, the Blue pixels are disk and the Green Pixels are Irregular

Conclusion and Outlook

This thesis successfully demonstrates the application and development of advanced machine learning techniques for the classification and semantic segmentation of astronomical images from the high redshift universe, leveraging data from both the JWST and HST. Through replicating Morpheus results, applying it to UNCOVER data, and developing BigMorpheus, this work showcases the potential of deep learning in astronomical studies. BigMorpheus, in particular, with its fast attention mechanism, presents a significant advancement in processing speed which will be extremely helpful in the big data era of Astronomy. Future directions include refining these models to improve deblending capabilities and extend applicability to broader datasets, thereby enhancing our understanding of galaxy formation and evolution.

References

[1] A. J. Benson, "Galaxy formation theory," *Physics Reports*, vol. 495, no. 2–3, pp. 33–86, Oct. 2010, ISSN: 0370-1573. DOI: 10.1016/j.physrep.2010.06.001. URL: http://dx.doi.org/10.1016/j.physrep.2010.06.001.

[2] Wikipedia contributors, *Galaxy formation and evolution — Wikipedia, the free encyclopedia*, [Online; accessed 14-March-2024], 2024. URL: https://en.wikipedia.org/w/index.php?title=Galaxy_formation_and_evolution&oldid=1198200364.

[3] C. J. Conselice, "The Relationship between Stellar Light Distributions of Galaxies and Their Formation Histories,", vol. 147, no. 1, pp. 1–28, Jul. 2003. DOI: 10.1086/375001. arXiv: astro-ph/0303065 [astro-ph].

[4] C. J. Conselice, "The evolution of galaxy structure over cosmic time," *Annual Review of Astronomy and Astrophysics*, vol. 52, no. 1, pp. 291–337, Aug. 2014, ISSN: 1545-4282. DOI: 10.1146/annurev-astro-081913-040037. URL: http://dx.doi.org/10.1146/annurev-astro-081913-040037.

[5] R. Hausen and B. E. Robertson, "Morpheus: A deep learning framework for the pixel-level analysis of astronomical image data," *The Astrophysical Journal Supplement Series*, vol. 248, no. 1, p. 20, May 2020, ISSN: 1538-4365. DOI: 10.3847/1538-4365/ab8868. URL: http://dx.doi.org/10.3847/1538-4365/ab8868.

[6] E. P. Hubble, "Extragalactic nebulae.,", vol. 64, pp. 321–369, Dec. 1926. DOI: 10.1086/143018.

[7] G. de Vaucouleurs, "Classification and Morphology of External Galaxies.," *Handbuch der Physik*, vol. 53, p. 275, Jan. 1959. DOI: 10.1007/978-3-642-45932-0_7.

[8] S. van den Bergh, "A new classification system for galaxies.,", vol. 206, pp. 883–887, Jun. 1976. DOI: 10.1086/154452.

[9] E. Holmberg, "A photographic photometry of extragalactic nebulae.," *Meddelanden fran Lunds Astronomiska Observatorium Serie II*, vol. 136, p. 1, Jan. 1958.

[10] G. de Vaucouleurs, "Recherches sur les Nebuleuses Extragalactiques," *Annales d'Astrophysique*, vol. 11, p. 247, Jan. 1948.

[11] J. L. Sérsic, "Influence of the atmospheric and instrumental dispersion on the brightness distribution in a galaxy," *Boletin de la Asociacion Argentina de Astronomia La Plata Argentina*, vol. 6, pp. 41–43, Feb. 1963.

[12] J. S. Kartaltepe, C. Rose, B. N. Vanderhoof, *et al.*, "Ceers key paper. iii. the diversity of galaxy structure and morphology at z = 3–9 with jwst," *The Astrophysical Journal Letters*, vol. 946, no. 1, p. L15, Mar. 2023, ISSN: 2041-8213. DOI: 10.3847/2041-8213/acad01. URL: http://dx.doi.org/10.3847/2041-8213/acad01.

[13] J. M. Lotz, J. Primack, and P. Madau, "A New Nonparametric Approach to Galaxy Morphological Classification,", vol. 128, no. 1, pp. 163–182, Jul. 2004. DOI: 10.1086/421849. arXiv: astro-ph/0311352 [astro-ph].

[14] M. A. Bershady, A. Jangren, and C. J. Conselice, "Structural and Photometric Classification of Galaxies. I. Calibration Based on a Nearby Galaxy Sample,", vol. 119, no. 6, pp. 2645–2663, Jun. 2000. DOI: 10.1086/301386. arXiv: astro-ph/0002262 [astro-ph].

[15] A. W. Graham, S. P. Driver, V. Petrosian, C. J. Conselice, M. A. Bershady, S. M. Crawford, and T. Goto, "Total Galaxy Magnitudes and Effective Radii from Petrosian Magnitudes and Radii,", vol. 130, no. 4, pp. 1535–1544, Oct. 2005. DOI: 10.1086/444475. arXiv: astro-ph/0504287 [astro-ph].

[16] A. W. Graham, P. Erwin, N. Caon, and I. Trujillo, "A Correlation between Galaxy Light Concentration and Supermassive Black Hole Mass,", vol. 563, no. 1, pp. L11–L14, Dec. 2001. DOI: 10.1086/338500. arXiv: astro-ph/0111152 [astro-ph].

[17] J. M. Lotz, M. Davis, S. M. Faber, *et al.*, "The Evolution of Galaxy Mergers and Morphology at z < 1.2 in the Extended Groth Strip,", vol. 672, no. 1, pp. 177–197, Jan. 2008. DOI: 10.1086/523659. arXiv: astro-ph/0602088 [astro-ph].

[18] S. C. Odewahn, E. B. Stockwell, R. L. Pennington, R. M. Humphreys, and W. A. Zumach, "Automated Star/Galaxy Discrimination With Neural Networks,", vol. 103, p. 318, Jan. 1992. DOI: 10.1086/116063.

[19] E. Bertin, "Classification of Astronomical Images with a Neural Network,", vol. 217, no. 1-2, pp. 49–51, Jul. 1994. DOI: 10.1007/BF00990023.

[20] S. R. Folkes, O. Lahav, and S. J. Maddox, "An artificial neural network approach to the classification of galaxy spectra.,", vol. 283, no. 2, pp. 651–665, Dec. 1996. DOI: 10.1093/mnras/283.2.651. arXiv: astro-ph/9608073 [astro-ph].

[21] A. E. Firth, O. Lahav, and R. S. Somerville, "Estimating photometric redshifts with artificial neural networks,", vol. 339, no. 4, pp. 1195–1202, Mar. 2003. DOI: 10.1046/j.1365-8711.2003.06271.x. arXiv: astro-ph/0203250 [astro-ph].

[22] A. A. Collister and O. Lahav, "ANNz: Estimating Photometric Redshifts Using Artificial Neural Networks,", vol. 116, no. 818, pp. 345–351, Apr. 2004. DOI: 10.1086/383254. arXiv: astro-ph/0311058 [astro-ph].

[23] C. J. Lintott, K. Schawinski, A. Slosar, *et al.*, "Galaxy Zoo: morphologies derived from visual inspection of galaxies from the Sloan Digital Sky Survey,", vol. 389, no. 3, pp. 1179–1189, Sep. 2008. DOI: 10.1111/j.1365-2966.2008.13689.x. arXiv: 0804.4483 [astro-ph].

[24] D. G. York, J. Adelman, J. Anderson John E., *et al.*, "The Sloan Digital Sky Survey: Technical Summary,", vol. 120, no. 3, pp. 1579–1587, Sep. 2000. DOI: 10.1086/301513. arXiv: astro-ph/0006396 [astro-ph].

[25] M. Huertas-Company, D. Rouan, L. Tasca, G. Soucail, and O. Le Fèvre, "A robust morphological classification of high-redshift galaxies using support vector machines on seeing limited images. I. Method description,", vol. 478, no. 3, pp. 971–980, Feb. 2008. DOI: 10.1051/0004-6361:20078625. arXiv: 0709.1359 [astro-ph].

[26] J. Mairal, P. Koniusz, Z. Harchaoui, and C. Schmid, *Convolutional kernel networks*, 2014. arXiv: 1406.3332 [cs.CV].

[27] S. Dieleman, K. W. Willett, and J. Dambre, "Rotation-invariant convolutional neural networks for galaxy morphology prediction,", vol. 450, no. 2, pp. 1441–1459, Jun. 2015. DOI: 10.1093/mnras/stv632. arXiv: 1503.07077 [astro-ph.IM].

[28] M. Huertas-Company, R. Gravet, G. Cabrera-Vives, *et al.*, "A catalog of visual-like morphologies in the 5 candels fields using deep learning," *The Astrophysical Journal Supplement Series*, vol. 221, no. 1, p. 8, Oct. 2015, ISSN: 1538-4365. DOI: 10.1088/0067-0049/221/1/8. URL: http://dx.doi.org/10.1088/0067-0049/221/1/8.

[29] A. K. Aniyan and K. Thorat, "Classifying radio galaxies with the convolutional neural network," *The Astrophysical Journal Supplement Series*, vol. 230, no. 2, p. 20, Jun. 2017, ISSN: 1538-4365. DOI: 10.3847/1538-4365/aa7333. URL: http://dx.doi.org/10.3847/1538-4365/aa7333.

[30] W. Alhassan, A. R. Taylor, and M. Vaccari, "The first classifier: Compact and extended radio galaxy classification using deep convolutional neural networks," *Monthly Notices of the Royal Astronomical Society*, vol. 480, no. 2, pp. 2085–2093, Jul. 2018, ISSN: 1365-2966. DOI: 10.1093/mnras/sty2038. URL: http://dx.doi.org/10.1093/mnras/sty2038.

[31] H. Domínguez Sánchez, M. Huertas-Company, M. Bernardi, D. Tuccillo, and J. L. Fischer, "Improving galaxy morphologies for SDSS with Deep Learning,", vol. 476, no. 3, pp. 3661–3676, Feb. 2018. DOI: 10.1093/mnras/sty338. arXiv: 1711.05744 [astro-ph.GA].

[32] C. J. Burke, P. D. Aleo, Y.-C. Chen, X. Liu, J. R. Peterson, G. H. Sembroski, and J. Y.-Y. Lin, "Deblending and classifying astronomical sources with mask r-cnn deep learning," *Monthly Notices of the Royal Astronomical Society*, vol. 490, no. 3, pp. 3952–3965, Oct. 2019, ISSN: 1365-2966. DOI: 10.1093/mnras/stz2845. URL: http://dx.doi.org/10.1093/mnras/stz2845.

[33] R. Katebi, Y. Zhou, R. Chornock, and R. Bunescu, "Galaxy morphology prediction using Capsule Networks,", vol. 486, no. 2, pp. 1539–1547, Jun. 2019. DOI: 10.1093/mnras/stz915. arXiv: 1809.08377 [astro-ph.IM].

[34] B. E. Robertson, S. Tacchella, B. D. Johnson, *et al.*, "Morpheus Reveals Distant Disk Galaxy Morphologies with JWST: The First AI/ML Analysis of JWST Images,", vol. 942, no. 2, L42, p. L42, Jan. 2023. DOI: 10.3847/2041-8213/aca086. arXiv: 2208.11456 [astro-ph.GA].

[35] D. Baron, *Machine learning in astronomy: A practical overview*, 2019. arXiv: 1904.07248 [astro-ph.IM].

[36] Y. Mo, Y. Wu, X. Yang, F. Liu, and Y. Liao, "Review the state-of-the-art technologies of semantic segmentation based on deep learning," *Neurocomputing*, vol. 493, pp. 626–646, 2022, ISSN: 0925-2312. DOI: https://doi.org/10.1016/j.neucom.2022.01.005. URL: https://www.sciencedirect.com/science/article/pii/S0925231222000054.

[37] O. Ronneberger, P. Fischer, and T. Brox, *U-net: Convolutional networks for biomedical image segmentation*, 2015. arXiv: 1505.04597 [cs.CV].

[38] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, *Attention is all you need*, 2023. arXiv: 1706.03762 [cs.CL].

[39] P. Hu, F. Perazzi, F. C. Heilbron, O. Wang, Z. Lin, K. Saenko, and S. Sclaroff, *Real-time semantic segmentation with fast attention*, 2020. arXiv: 2007.03815 [cs.CV].

[40] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, no. 6088, pp. 533–536, Oct. 1986. DOI: 10.1038/323533a0. URL: https://doi.org/10.1038/323533a0.