

---

Electronic Thesis and Dissertation Repository

---

8-22-2022 1:00 PM

# Classification of Stellar Age and Galaxy Components within Spiral Galaxies by use of Hubble Space Telescope Imagery and Machine Learning

Robin Kwik, *The University of Western Ontario*

Supervisor: Wang, Jinfei, *The University of Western Ontario*

Co-Supervisor: Barmby, Pauline, *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Master of Science degree in Geography and Environment

© Robin Kwik 2022

Follow this and additional works at: <https://ir.lib.uwo.ca/etd>

---

## Recommended Citation

Kwik, Robin, "Classification of Stellar Age and Galaxy Components within Spiral Galaxies by use of Hubble Space Telescope Imagery and Machine Learning" (2022). *Electronic Thesis and Dissertation Repository*. 8831.

<https://ir.lib.uwo.ca/etd/8831>

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact [wlsadmin@uwo.ca](mailto:wlsadmin@uwo.ca).

# Classification of Stellar Age and Galaxy Components within Spiral Galaxies by use of Hubble Space Telescope Imagery and Machine Learning

## Abstract

Galaxies have complex formations of components such as stars, dust, and gas, whose spatial and temporal relationships can help us to better understand the formation and evolution of galaxies, and ultimately the Universe. The main objective of this study is to test how machine learning can be used to classify galaxy components and stellar ages within spiral galaxies based on values of pixels in Hubble Space Telescope imagery, Euclidean distance calculations, textural features, and band ratios. We develop two machine learning models using maximum likelihood, random forest, and support vector machine algorithms. We find the models are successful for classification of galaxy components and stellar age, with Euclidean distance and textural features being the most important parameters. These methods can contribute to the rapid processing of high resolution astronomical imagery of galaxies and other celestial phenomena.

## Keywords

Spiral galaxies, Hubble Space Telescope, supervised machine learning, galaxy components, stellar age, texture analysis, random forest, support vector machine, maximum likelihood.

## Summary for Lay Audience

The Universe is thought to have formed around 14 billion years ago, with our Milky Way galaxy forming soon after. The Milky Way and all other galaxies are made of components such as stars, dust, and gas. Different types of galaxies exhibit different patterns of components: elliptical galaxies are round in shape and host a large population of older stars, spiral galaxies are characterized by arms extending from them and have a higher population of young stars, and irregular galaxies lack patterns, exhibiting random distributions of young and old stars. Temperature and brightness determine the colour of stars we observe. Younger stars are hotter and brighter and appear bluer in colour, while older stars are colder and dimmer, appearing redder in colour.

In this thesis, we use remote sensing techniques to observe galaxy components and the ages of stars within two spiral galaxies. Remote sensing can be defined as the gathering of information about different types of phenomena by distant observation; for example, by using satellites or telescopes. We take information gathered from light emitted by components within the spiral galaxies in the form of Hubble Space Telescope imagery. The Hubble Space Telescope is able to take images using filters that filter out different types of light (e.g., blue light) and focus on specific colours of light emitted from phenomena.

We train a computer to automatically classify stellar age and galaxy component membership of each pixel in the Hubble Space Telescope images; this process is called machine learning. We use information stored within images to train the computer: Hubble Space Telescope images in several colours of light, distance of pixels from the spiral arms and galaxy center within the galaxies, patterns of spatial distribution of the galaxy components and stellar ages, and band ratios that compare the amounts of different colours of light emitted from the galaxies (e.g., blue light divided by green light). By observing the different ages of stars and the spatial relationships between the components with galaxies, we can better understand the formation and evolution of galaxies, the Universe, and ultimately how matter formed.

## Co-Authorship Statement

The work presented in this thesis follows the integrated-article layout. The majority of the work, including analysis and writing of the manuscripts, was done by me. Supervisors Dr. Wang and Dr. Barmby provided guidance, expertise, proofreading, and comments on the work contained in this thesis.

### **Chapter 2:**

The co-authors of the following published article are Dr. Wang, Dr. Barmby, and Dr. Benne W. Holwerda:

**Kwik, R.J.**, Wang, J., Barmby, P., & Holwerda, B.W. (2022). Galactic component mapping of galaxy UGC 2885 by machine learning classification. *Advances in Space Research*, 70, 229–247. <https://doi.org/10.1016/j.asr.2022.04.032>

### **Chapter 3:**

The following article has been submitted for publication and is co-authored by Dr. Wang, Dr. Barmby, and Dr. M. Carmen Sánchez-Gil:

**Kwik, R.J.**, Wang, J., Barmby, P., & Sánchez-Gil, M.C. Age classification of spiral galaxy M83 by supervised machine learning.

## Acknowledgments

Firstly, I thank my supervisors, Dr. Jinfei Wang and Dr. Pauline Barmby, for all their patience, guidance, and encouragement they provided me with through the research process. They gave me the opportunity to merge my two interests and study galaxies using remote sensing. It has been an experience that I will always remember. Their kindness towards has made me so grateful to have such remarkable supervisors. I have known them both since my Bachelors, and am so thankful for everything they have done for me.

I would also like to thank Dr. Philip Stooke for taking the time to provide constructive feedback and advice throughout the research process. I greatly appreciate it.

To my mom and dad, Penny and Richard Kwik, my sister Jessica, and my best friend and companion Wade, I thank you for all of your encouragement and support. I could not have done this without you all. Our trips to the campsite, canoeing, kayaking, fishing, and all the time we spend together at home and going antiquing, being 'nerds', and searching for uranium glass is so cherished. I could not have asked for a better family.

# Table of Contents

Classification of Stellar Age and Galaxy Components within Spiral Galaxies by use of Hubble Space Telescope Imagery and Machine Learning.....	ii
Abstract.....	ii
Summary for Lay Audience.....	iii
Co-Authorship Statement.....	iv
Acknowledgments.....	v
Table of Contents.....	vi
List of Tables.....	x
List of Figures.....	xii
List of Appendices.....	xvi
Chapter 1.....	1
1 Introduction.....	1
1.1 Background.....	1
1.1.1 Galaxies and Galactic Components.....	1
1.1.2 Observation Methods.....	4
1.1.3 Literature Review: Galaxy Component and Stellar Age Estimation Methods.....	6
1.1.4 Research Objectives.....	9
1.2 Study Areas.....	10
1.3 References.....	13
Chapter 2.....	19
2 Galactic Component Mapping of Galaxy UGC 2885 by Machine Learning Classification.....	19
2.1 Introduction.....	19
2.2 Methods.....	22
2.2.1 Study Area.....	22

2.2.2	Data.....	23
2.3	HST Image Processing .....	26
2.3.1	Coordinate Transformation.....	26
2.3.2	Band Ratio .....	28
2.3.3	Texture Features.....	30
2.3.4	Distance from Spiral Arms and Galaxy Center.....	33
2.4	Classification Scheme .....	37
2.5	Machine Learning Algorithms.....	39
2.5.1	Maximum Likelihood Classifier .....	39
2.5.2	Support Vector Machine.....	40
2.5.3	Random Forest .....	40
2.5.4	Classification Groups .....	41
2.6	Accuracy Assessment.....	42
2.7	Results .....	43
2.7.1	Parameter Importance.....	43
2.7.2	Galaxy Component Classification Performance .....	49
2.7.3	Model Performance Summary .....	54
2.8	Discussion.....	54
2.9	Conclusions .....	57
2.10	References .....	59
Chapter 3	.....	69
3	Age Classification of Spiral Galaxy M83 by Supervised Machine Learning .....	69
3.1	Introduction .....	69
3.2	Methods.....	72
3.2.1	Study Area .....	72
3.2.2	Data.....	73

3.3	Parameter Creation.....	78
3.3.1	Euclidean Distance .....	80
3.3.2	Texture.....	82
3.4	Reference Data.....	86
3.4.1	Classification Scheme.....	87
3.5	Machine Learning Classification .....	90
3.5.1	Support Vector Machine .....	90
3.5.2	Random Forest .....	91
3.6	Accuracy Assessment.....	92
3.6.1	Classification Groups .....	95
3.7	Results .....	95
3.7.1	Parameter Importance.....	96
3.7.2	Age Classification Performance.....	103
3.7.3	Model Performance .....	104
3.8	Discussion.....	105
3.9	Conclusions .....	108
3.10	References .....	110
Chapter 4	.....	116
4	Conclusion.....	116
4.1	Summary .....	116
4.2	Conclusions .....	117
4.3	Limitations and Future Work .....	118
4.4	References .....	121
5	Appendices .....	122
	Appendix A: Textures .....	122
	Appendix B: Band Ratios.....	132

Appendix C: Classified Rasters .....	134
Appendix D: Copyright Releases from Publications .....	152
Curriculum Vitae .....	153

## List of Tables

Table 2-1: Band ratios used for classification and their respective equations. .... 28

Table 2-2: Calibration factors for the BVI imagery as found in FITS Liberator 3 Image Headers tab under “PHOTFNU”. .... 29

Table 2-3: Number of polygons and pixels within 70% and 30% of the polygon dataset. .... 42

Table 2-4: Confusion matrix of SVM classification of the top seven layer using 70% of the polygon set for training and 30% for testing. The green shading shows the number of correctly classified pixels for each class, orange shading shows the PA and UA row and column, and the blue shading shows the OA statistic. Bolded numbers identify the areas of the most confusion between classes. .... 49

Table 2-5: Confusion matrix of SVM classification of the top seven layer using 30% of the polygon set for training and 70% for testing. The green shading shows the number of correctly classified pixels for each class, orange shading shows the PA and UA row and column, and the blue shading shows the OA statistic. Bolded numbers identify the areas of the most confusion between classes. .... 49

Table 2-6: Average of user’s accuracy for each class in the MLC, RF, and SVM models. The bolded numbers represent the model with the highest user’s accuracy for each class. SVM model outperforms the MLC and RF models. .... 51

Table 2-7: Average of producer’s accuracy for each class in the MLC, RF, and SVM models.<sup>3</sup> The bolded numbers represent the model with the highest producer’s accuracy for each class. SVM and RF models outperform the MLC. .... 51

Table 2-8: Average of F1 Score for each class in the MLC, RF, and SVM models.<sup>3</sup> The bolded numbers represent the model with the highest F1 score accuracy for each class. SVM model outperforms the MLC and RF models. .... 51

Table 2-9: Top performing classifications from Figure 2-11 and their corresponding accuracy statistics. .... 54

Table 2-10: Average accuracies for each model. ....	54
Table 3-1: HST wavelength bands for M83 and their respective wavelength ranges. ....	74
Table 3-2. Chosen classification scheme of age map pixels. ....	88
Table 3-3. Number of M83 pixels for each class within the 80% and 20% splits of the reference set. ....	93
Table 3-4. Chosen classification groups. ....	95
Table 3-5. Chosen classification groups. ....	98
Table 3-6. Confusion Matrix of best performing classification: RF classification model with mean textures (F438W, F555W-F547M, F657N, and F814W) and distance from spiral arm and galaxy center. Top: classification using 80% of training data for model training and 20% for model testing; bottom: classification using 20% of training data for model training and 80% for model testing. The italicized numbers on the diagonal show the number of correctly classified pixels in each age class. The bolded numbers represent the overall map accuracy (OA). Columns named UA are the UA for each age class and rows named PA are the PA for each age class. ....	99
Table 3-7. Age class UA, PA, and F1 score statistics for RF and SVM models. Accuracy statistics are averaged over all eight classifications for each model. ....	103
Table 3-8. All classifications and their overall accuracy (OA) statistics for RF and SVM classification models. The OA reported is average of the classifications trained on 80% and 20% of the total reference set. The top 5 highest accuracies are bolded. ....	105

# List of Figures

Figure 1-1. Examples of the three main types of galaxies: a) spiral galaxy (image credit: Hubble Heritage Team, ESA, NASA); b) elliptical galaxy (image credit: NASA and ESA; Acknowledgment: A. Levan (U. Warwick), N. Tanvir (U. Leicester), and A. Fruchter and O. Fox (STScI)); c) irregular galaxy (image credit: NASA and ESA; Acknowledgment: T. Roberts (Durham University, UK), D. Calzetti (University of Massachusetts) and the LEGUS Team, R. Tully (University of Hawaii), and R. Chandar (University of Toledo)). ..... 2

Figure 1-2. Hubble Space Telescope image of galaxy UGC 2885. Image credit: NASA, ESA, and B. Holwerda (University of Louisville). ..... 11

Figure 1-3. Hubble Space Telescope image of galaxy M83. Image credit: NASA, ESA, and the Hubble Heritage Team (STScI/AURA); Acknowledgement: W. Blair (STScI/Johns Hopkins University) and R. O'Connell (University of Virginia). ..... 12

Figure 2-1: HST colour composite map of UGC 2885 where the F475W band is displayed in blue, the F606W band is displayed in green, and the F814W band is displayed in red. This image is in celestial orientation; therefore, east is towards the left rather than to the right of the north direction. Parsecs refers to a measure of distance equaling approximately 3.26 light years. Image credit: NASA, ESA, and B.W. Holwerda (University of Louisville). ..... 23

Figure 2-2: HST wide band filter wavelength ranges. (Dressel et al., 2022) ..... 24

Figure 2-3: UGC 2885 blue-green band imagery (F475W): a) original HST imagery showing the projected galaxy with an inclination of 74°; b) deprojected image of UGC 2885 – stretched vertically by 363%. ..... 25

Figure 2-4: Flowchart showing methods used in this study. .... 27

Figure 2-5: Band F606W Haralick textures with 20 brightest foreground stars masked. .... 33

Figure 2-6: Features used to calculate Euclidean distance and the corresponding Euclidean distance rasters: a) spiral arm and galaxy center features; b) galaxy center distance raster; c) spiral arm distance raster. Spiral arms and galaxy center features shown in red and blue

respectively. The background image shows the deprojected UGC 2885 F475W blue-green band (Holwerda et al., 2020). Both distance rasters are reprojected to the galaxy's 74° inclination..... 36

Figure 2-7: Examples of training site selection using the classification scheme with six classes: a) young stellar population; b) old stellar population; c) dust lanes; d) galaxy center; e) outer disc; f) celestial background. Background image is a RGB colour composite of HST band data of UGC 2885: F475W band is displayed in blue, the F606W band is displayed in green, and the F814W band is displayed in red. .... 38

Figure 2-8: Scatterplot of classification scheme pixels and their respective digital number (DN) values within the blue-green and near-infrared HST bands. .... 39

Figure 2-9: MDG plot of the top 30 important input parameters. .... 44

Figure 2-10: SVM classifications using 70% of the polygon set as training and 30% of the polygons as testing: a) HST bands (F475W, F606W, and F814W); b) Mean texture; c) HST bands and distance; d) Top 7 MDG parameters. We include inset maps, represented by red squares, to show the small details within the inner disc of the galaxy. .... 47

Figure 2-11: Top performing classifications: a) HST bands and distance RF; b) HST bands and distance SVM; c) Top 7 MDG parameters RF; d) Top 7 MDG parameters SVM. All maps shown are the product of classifications using 70% of the polygon set as training and 30% as testing. We include inset maps, represented by red squares, to show the fine details within the inner disc of the galaxy. .... 53

Figure 3-1. Hubble Space Telescope map of galaxy M83: Purple: ultraviolet band (F336W), Blue: blue band (F438W), Green: visual band (F555W) + Strömgren y (F547M), Orange: H-alpha (F657N), Red: near-infrared (F814W). Image credit: NASA, ESA, and the Hubble Heritage Team (STScI/AURA) ..... 73

Figure 3-2: HST WFC3 narrow, medium, and wide band filter wavelength ranges. (Dressel et al., 2022) ..... 75

Figure 3-3. HST imagery of M83 geometrically corrected to Earth orientation (East to the right): a) HST blue band F438W; b) HST visual band F555W-547M; c) HST H-alpha band F657N; d) HST near infrared band F814W. .... 77

Figure 3-4. Flowchart of methodology used in this study ..... 79

Figure 3-5. Euclidean distance features drawn over the HST imagery and their respective distance calculations: a) Euclidean distance from spiral arm features; b) Euclidean distance from galaxy center feature; c) Spiral arm line feature traced over the HST blue band image; d) Galaxy center polygon feature drawn over galaxy center. Background image is the HST blue band (F438W) shown in grayscale ..... 81

Figure 3-6. Haralick textures calculated from HST blue (F438W) imagery of M83: a) HST blue band image (F438W); b) angular second moment texture; c) contrast texture; d) correlation texture; e) dissimilarity texture; f) entropy texture; g) homogeneity texture; h) mean texture; i) standard deviation texture. White color represents higher values and black represents lower values. .... 85

Figure 3-7. Original age map of young stellar populations within M83 (Sánchez-Gil et al. 2019). Age is in millions of years (Myr). Blue pixels represent the youngest regions while red pixels represent the oldest. The black box indicates the location of the inset shown in Figure 3-7..... 87

Figure 3-8. Pixel resolution comparison for the same region of the HST imagery and the M83 age map as shown on the black box in Figure 3-6: a) HST band imagery; b) reference age map. The HST imagery is an RGB color composite where R – F657N, G – F555W-547M, and B – F438W. The images represent an area of approximately 317.24 x 482.24 parsecs as measured on the geometrically corrected imagery. The red “bubbles” within the HST imagery is H-alpha emission showing the young star forming regions. .... 89

Figure 3-9. Training sites drawn over color composite of HST imagery of M83: a) age class A1 (1 and 2 Myr); b) age class A2 (3 Myr); c) age class A3 (4 Myr); d) age class A4 (5 Myr); e) age class A5 (6, 7, and 8 Myr). Colour composite RGB assignment: R – F657N, G – F555W-547M, B – F438W. .... 90

Figure 3-10. Scatterplot showing the original digital number (DN) values of the HST imagery (0 – 65535) for the age classes A1 to A5. We show the DN values for the 20% reference set with 18,825 pixel samples..... 93

Figure 3-11. MDG results for the top 30 parameters. .... 97

Figure 3-12. RF classification results are compared to show the differences in classification when using different parameters: a) HST imagery classification b) HST imagery and distance classification; c) mean textures classification; d) mean textures and distance classification; e) top 3 MDG parameters classification; f) top 16 MDG parameters classification; g) reference age map. All maps are classification results using 80% of the reference set for training and 20% of the reference set for testing. ....102

# List of Appendices

## Appendix A: Textures

Figure A 1. Haralick textures calculated on the blue-green HST image (F475W) of UGC 2885.....122

Figure A 2. Haralick textures calculated on the near-infrared HST image (F814W) of UGC 2885.....124

Figure A 3. Haralick textures calculated on the visual HST image (F555W-547M) of M83.....126

Figure A 4. Haralick textures calculated on the Hydrogen-alpha HST image (F657N) of M83.....128

Figure A 5. Haralick textures calculated on the near-infrared HST image (F814W) of M83.....130

## Appendix B: Band Ratios

Figure B 1. Band ratios calculated from HST imagery of galaxy UGC 2885.....132

## Appendix C: Classified Rasters

Figure C 1. Maximum likelihood classifier classifications of components within galaxy UGC 2885 not shown in Chapter 2. Classifications use 70% of the reference set for training and 30% of the reference set for testing.....134

Figure C 2. Random forest classifications of components within galaxy UGC 2885 not shown in Chapter 2. Classifications use 70% of the reference set for training and 30% of the reference set for testing.....136

Figure C 3. Support vector machine classifications of components within galaxy UGC 2885 not shown in Chapter 2. Classifications use 70% of the reference set for training and 30% of the reference set for testing.....138

Figure C 4. Maximum likelihood classifier classifications of components within galaxy UGC 2885 not shown in Chapter 2. Classifications use 30% of the reference set for training and 70% of the reference set for testing.....139

Figure C 5. Random forest classifications of components within galaxy UGC 2885 not shown in Chapter 2. Classifications use 30% of the reference set for training and 70% of the reference set for testing.....141

Figure C 6. Support vector machine classifications of components within galaxy UGC 2885 not shown in Chapter 2. Classifications use 30% of the reference set for training and 70% of the reference set for testing.....143

Figure C 7. Random forest classifications of stellar age within galaxy M83 not shown in Chapter 3. Classifications use 80% of the reference set for training and 20% of the reference set for testing.....145

Figure C 8. Support vector machine classifications of stellar age within galaxy M83 not shown in Chapter 3. Classifications use 80% of the reference set for training and 20% of the reference set for testing.....146

Figure C 9. Random forest classifications of stellar age within galaxy M83 not shown in Chapter 3. Classifications use 20% of the reference set for training and 80% of the reference set for testing.....148

Figure C 10. Support vector machine classifications of stellar age within galaxy M83 not shown in Chapter 3. Classifications use 20% of the reference set for training and 80% of the reference set for testing.....150

Appendix D: Copyright Releases from Publications

# Chapter 1

## 1 Introduction

### 1.1 Background

#### 1.1.1 Galaxies and Galactic Components

Astronomy is defined as the study of celestial phenomena such as stars and planets. Researchers study celestial phenomena to gain a better understanding of how our universe and matter was formed. Everything we can see in the universe can give us hints to these processes. An excellent study area is the galaxy we exist within, the Milky Way. The Milky Way is thought to be made up of over 100 billion stars (Bennett et al., 2015), where each star is estimated to have at least one planet orbiting it (Cassan et al., 2012). By studying light from other stars and galaxy components we can better understand the formation of our Milky Way galaxy and matter within our Solar System. One limitation in astronomy is that we cannot observe the full structure of the Milky Way because we exist within it, so we can only look through it. To address this issue we can observe other galaxies outside of our own.

Galaxies are massive aggregations of matter such as stars, dust, and gas. The three main types of galaxies are spiral, elliptical, and irregular (Figure 1-1). Spiral galaxies have arms extending from them and have larger populations of young stars that are bluer in appearance. Elliptical galaxies are round in shape and have a larger population of old and red stars. Irregular galaxies cannot be classified as spiral or elliptical, as they lack similar structural characteristics (Bennett et al., 2015). This thesis focuses on spiral galaxies that have a diversity of galaxy components and stellar ages. Galaxies are important to study because they can help us better understand how the Universe, Milky Way galaxy, and all matter formed. All matter we find on Earth is a product of our galaxy's formation, making us deeply connected with the components within it.

(a) Spiral galaxy NGC 1300



(b) Elliptical galaxy NGC 4993



(c) Irregular galaxy NGC 4485



**Figure 1-1. Examples of the three main types of galaxies: a) spiral galaxy (image credit: Hubble Heritage Team, ESA, NASA); b) elliptical galaxy (image credit: NASA and ESA; Acknowledgment: A. Levan (U. Warwick), N. Tanvir (U. Leicester), and A. Fruchter and O. Fox (STScI)); c) irregular galaxy (image credit: NASA and ESA; Acknowledgment: T. Roberts (Durham University, UK), D. Calzetti (University of Massachusetts) and the LEGUS Team, R. Tully (University of Hawaii), and R. Chandar (University of Toledo)).**

The main substructures within spiral galaxies include a disc, bulge, bar, and spiral arms. All spiral galaxies have discs which are the flat portion of the galaxy that is filled with the galaxy components. Bulges of spiral galaxies are in the center of the galaxy and have a large concentration of older stars. A bar is defined as a bar-like structure filled with stars that extends through the galaxy center; bars are not characteristic of all spiral galaxies.

Spiral arms also vary between galaxies with some having stronger spiral arm structures than others. There are several types of spiral galaxies including grand design, interacting, and flocculent, with each of these types having different characteristics of substructures. Grand design spiral galaxies have two well-defined spiral arms. Interacting spiral galaxies are those that have a companion galaxy that they are merging with through gravitational influence. Flocculent galaxies do not exhibit clear spiral structures like those seen in grand design type galaxies. Some spiral galaxies exhibit age gradients where young stars are found mainly along the spiral arms and older stars are found in between the spiral arms and in the galaxy center. Spiral density wave theory helps to explain these age gradients; as the density wave moves throughout the galaxy, it compresses gas and ignites star formation, creating star clusters that are concentrations of many stars in a small area (Lin & Shu, 1964). As the density wave continues moving throughout the galaxy, it may leave some of the newly formed star clusters behind. These young stars slowly age and disperse, creating a stellar field of many old stars. In nearby galaxies we are able to observe stellar populations in great detail.

Many studies have been undertaken to find evidence of these density waves within nearby spiral galaxies through the identification of age gradients (e.g., Abdeen et al., 2022; Sánchez-Gil et al., 2019; Shabani et al., 2018). Researchers have disagreed upon the existence of age gradients. One reason for this is the use of different methods of age gradient identification. One method of age gradient identification involves analysis of the mid infrared emission peaks within the galaxy. Mid infrared emission is a good identifier of stellar mass, so is useful for identifying the high mass within the spiral arms (Scheepmaker et al., 2009). Tracing dust lanes along the spiral arms has also been tested (Shabani et al., 2018; Bialopetravičius & Narbutis, 2020). Dust lanes along the spiral arm are approximations of the galactic shock front; the shock front is defined as the region characterized by the compression of gas by the density wave that creates dust lanes and ignites star formation (Shu, 2016). The characteristics of spiral galaxies affect their ability to produce age gradients. Several studies have identified grand design spiral galaxies as exhibiting clear age gradients (Sánchez-Gil et al., 2011) while interacting and flocculent galaxies are found to not exhibit clear age gradients (Dobbs & Pringle, 2010).

Galaxy components exhibit different spatial and temporal patterns: younger stars are much more compact, clumpier, and brighter than older stars (Whitmore et al. 2011; Jensen et al. 1981). Typically older stars can be found spread out within regions called stellar fields. Older stars are also not surrounded by as much ionized gas or dust as young stars due to their location in the inter-arm regions where there is less density of galaxy components. Texture features or image descriptors, a method of identifying patterns within digital imagery, have been tested for galaxy morphology (Diaz-Hernandez et al., 2016; Ntwaetsile & Geach, 2021) and outlier identification (Shamir, 2021). However, there is a lack of texture feature testing for galaxy components and smaller scale structures within galaxies.

### 1.1.2 Observation Methods

We study celestial phenomena such as galaxies through the use of remote sensing. Remote sensing can be defined as the observation of phenomena and acquiring of data without direct contact (Lillesand et al., 2015). Remote sensing focuses on electromagnetic radiation emittance and reflectance. Electromagnetic radiation is defined on a spectra ranging from longer wavelengths to shorter ones where a wavelength is denoted by  $\lambda$ . Frequency describes the number of wavelengths occurring within a certain time and is denoted by  $\nu$ . On the electromagnetic spectrum, the shorter the wavelength of radiation the higher its frequency. Remote sensing for Earth-based observation is typically done through satellites orbiting Earth or by unmanned aerial vehicles (UAV) flying over Earth's surface.

For celestial observation, astronomers perform remote sensing using telescopes that can capture images of astronomical phenomena in specific wavelength ranges on the electromagnetic spectrum. Observing celestial phenomena is done through passive remote sensing where the telescopes capture the incoming electromagnetic radiation. For Earth observation, passive sensors rely on illumination from the Sun as a source of electromagnetic radiation. The second type of observation, active remote sensing, is where a remote sensor sends electromagnetic radiation to an object, the electromagnetic radiation then bounces back, and the remote sensor gathers information such as distance to the object or the height of the object off of the ground. An example of this is the use of radar, where a remote sensing device is capable of sending pulses of radiation to an object and making measurements based on the return time and the strength of the returning signal (Lillesand

et al., 2015). For celestial observation, active remote sensing is limited to objects within our solar system due to the vast distances of extrasolar phenomena.

Astronomers observe celestial phenomena in two ways, through imaging and spectroscopy. Imaging involves creating images of celestial phenomena by capturing the light emitted from them. Telescopes take images using filters designed to filter out light in specific wavelengths on the electromagnetic spectrum. For instance, younger and bluer stars emit strongly in the ultraviolet or blue wavelengths and less so in the near-infrared wavelengths, so filtering out the near-infrared light can help to observe young stars better. Older and redder stars on the other hand have stronger emission in the red and near-infrared wavelengths and less emission in ultraviolet or blue wavelengths (van den Bergh, 1975). Dust lanes that run through the galaxy absorb some of the light from the stars behind them, preferentially the bluer light, making the dust lanes appear darker in bluer bands and brighter in the longer wavelength bands. We are only able to see visible wavelengths (colours of the rainbow), so imaging helps astronomers observe galaxy components and processes that we would not normally be able to identify with our eyes alone. The equivalent to imaging in Earth-based remote sensing is multispectral sensing. This method takes several images in broad and narrow bands. Broad bands sense a larger range of electromagnetic radiation whereas narrow bands sense a smaller range.

Spectroscopy is the study of the spectra of celestial phenomena. Spectra shows the electromagnetic radiation emission of phenomena by splitting the incoming light into a spectrum; this is similar to when we see rainbows in the sky (Pössel, 2020). Spectroscopy helps astronomers identify properties of stars such as temperature, composition, and size of the star. Therefore, analyzing spectra of celestial phenomena is important for understanding the populations of stars within galaxies. Further, by studying the spectral emission from galaxies, astronomers can determine the redshift – the stretching of electromagnetic radiation wavelengths due to the vast distances that the electromagnetic radiation has to travel (Bennett et al., 2015) – that tells us how far away the galaxy is from us (Long & de Souza, 2017). The remote sensing equivalent to spectroscopy is hyperspectral sensing which is useful for analyzing phenomena using spectral information in many narrow bands that capture very specific wavelength ranges (Lillesand et al., 2015).

Astronomers use narrow bands to identify specific emission lines in both spectroscopy and imaging; for instance, the high ultraviolet output resulting from OB star formation ionizes hydrogen gas which can be seen through the Hydrogen-alpha emission line at 657 nm on a spectra. Narrow band images are denoted by an N in Hubble Space Telescope imagery – for instance, the Hydrogen-alpha band F657N – whereas broad bands are denoted using a W (e.g., blue band F438W).

In this thesis, we focus our methods on processing imaging data from the Hubble Space Telescope (HST), a famous space telescope that is capable of producing high-resolution imagery. High-resolution telescopes allow us to see great detail and structure of galaxies, but have limited wavelength sensing capabilities. For instance, HST is only able to take images in far ultraviolet to near-infrared wavelengths. Some examples of other high-resolution telescopes include James Webb Space Telescope (Gardner et al., 2006), Euclid Mission (Laureijs et al., 2011), Nancy Roman Space Telescope (Spergel et al., 2015), and The Cosmological Advanced Survey Telescope for Optical and UV Research (CASTOR; Coté et al., 2019). To capture images in longer wavelengths such as radio emission, a larger telescope is needed (Barmby, 2019). Ground-based telescopes are often used for this purpose because radio waves are capable of penetrating Earth's atmosphere and astronomers are able to use an array of many telescopes on the Earth surface (interferometry) to obtain a higher resolution (Bennett et al., 2015). The smaller the wavelength being sensed, the smaller the telescope mirrors can be, so space telescopes are often used to sense smaller wavelengths. Putting large telescopes in space is costly and poses many risks.

### 1.1.3 Literature Review: Galaxy Component and Stellar Age Estimation Methods

This Thesis focuses on galaxy components and stellar populations within spiral galaxies. Galaxy components are organized in spatial and temporal relationships. All components are directly related to each other. For instance, the density waves within spiral galaxies compress gas, creating dust lanes and stars. The stars age and eventually die, returning to dust. Within the shock front where gas is compressed into dust lanes, O and B stars are produced and their locations are used by researchers as identifiers of star formation within

galaxies. The OB stars are young, hot, and bright, emitting strongly in ultraviolet wavelengths which is capable of ionizing Hydrogen gas surrounding the star forming complex. The direct relation between the ultraviolet emission from the OB stars and the Hydrogen gas is helpful for identification of stellar age (Sánchez-Gil et al., 2011; 2019).

One method of age estimation of stellar populations within galaxies is using a diagram describing brightness (luminosity) and temperature of stars called the Hertzsprung-Russel (H-R) diagram (Bennett et al., 2015). An example from the H-R diagram are the OB stars previously mentioned that are extremely hot and bright, using up their fuel (Hydrogen gas) in a shorter time span. At the opposite end of the diagram are M stars which have the lowest temperature and luminosity, making them red and dim in appearance. We can produce different version of the H-R diagram called colour magnitude diagrams (CMDs) which compares magnitude in one wavelength band such as visual (V) to a ratio of two wavelength bands (referred to as colours in astronomy), for example, the blue (B) band divided by the visual (V) band (B-V). Colour-colour diagrams are also used, where colours are plotted on both axes; for example, B-V on the horizontal axis and U-B (ultraviolet divided by blue) on the vertical. CMDs can provide information on the star formation history of stellar populations within galaxies, making them useful for understanding formation and evolution of galaxies and their components (Brown, 2005).

As shown by the H-R diagram and CMDs, stars emit unique ratios of wavelengths of electromagnetic radiation. Many methods have been developed to estimate stellar properties of galaxy components such as age and metallicity. Simple stellar population (SSP) models are used for understanding star cluster properties; SSPs describe stellar populations that were created simultaneously from the same gas cloud within the same region of space, and therefore all have the same composition (Bruzual A., 2010). Models of SSPs are used to compute spectral energy distributions (SEDs) that describe the complete set of stellar populations within a galaxy. As an example, Wei et al. (2021) use SED model fitting to produce maps of stellar properties in a spiral galaxy M51 and its companion galaxy NGC 5195. The authors make use of 28 images ranging from UV to mid-infrared that are taken by several telescopes. They find that spiral galaxy M51 has a large population of younger stars while elliptical galaxy NGC 5195 has a large population

of old stars. From SED model fitting, Wei et al. (2021) are able to understand the recent star formation history of these two galaxies. For the same galaxy system, Lee et al. (2011) use a method of fitting called a pixel colour magnitude diagram (pCMD) that estimates properties of stellar populations within pixels based on CMDs produced. They use high resolution BVI imagery from the HST. The authors identify several age groups ranging from millions to billions of years. Their findings agree with those of Wei et al. (2021) where they concluded that the spiral arms of M51 display active star formation. Both papers used different methods and came to similar conclusions, however, Lee et al. (2011) made use of higher resolution imagery.

SED modelling is common within the astronomical research community, however, it does have its own limitations. Galaxies are much more complex than depicted by the SED models (Walcher et al., 2011). SED models often require images in many wavelength bands from multiple telescopes, many with different pixel resolutions. Due to the inconsistency, resampling of the pixel resolution is performed; resampling involves reducing the images containing pixels of higher resolution down to the same resolution as the images with the lowest pixel resolution. This averages the pixels, creating a loss of important information. SED model fitting is also time consuming, so the development of a model requiring less data while reducing the time to complete SED fitting would be beneficial for astronomical research.

One of the overarching goals of astronomers is to understand the formation and evolution of our universe. Studying galaxies is an excellent way to achieve this goal (King, 1971). Studying the spatial and temporal relationships of stars and galaxy components can help describe the processes that take place within galaxies and govern how galaxy components are organized. Galaxy formation and evolution occurs over a timespan of billions of years; we cannot observe these processes in real time, so astronomers need to study galaxies in many stages of life and make inferences about how processes occur over celestial time. High resolution telescopes such as the Hubble Space Telescope are able to see great detail within galaxies, and will also be able to see further into space as telescopes improve. Observing nearby galaxies can help us to analyze galaxy component distribution and

attempt to recreate how those galaxies might have formed through simulations (Bovy et al., 2016).

Identification and study of structures within spiral galaxies such as the bar, disc, and bulge have been performed in order to better understand the formation and evolution of galaxies (Lingard et al., 2020). Morphological identification of galaxies has also been widely studied (Lintott et al., 2008). Pixel-based analysis and mapping of galaxies is becoming more common as high resolution imagery is more readily available, and astronomers often use complex models and many images or spectra to analyze the pixel-level information of imagery. Pixel-level analysis of galaxies can contribute to a better understanding of the processes taking place within the substructures of galaxies. Further, galaxies of the same morphological type and of the same life cycle stage show differences in their galaxy component populations (Sánchez-Gil et al., 2011), so pixel analysis can help to better understand how these differences occur.

#### 1.1.4 Research Objectives

Higher resolution telescopes being launched and the massive amounts of data we receive from those telescopes is an advancement for the astronomical field, but presents a problem. Processing all of the data requires a significant amount of time to complete. SED models are common for stellar property estimation based on known spectra of galaxy components. However, with the limitations that SED models present, the development of a more rapid model that estimates stellar properties using less data would contribute to astronomical research. In an attempt to speed up the data processing, analysis by artificial intelligence such as machine learning and deep learning has become more prominent (Fluke & Jacobs, 2020). Machine learning and deep learning models can be trained using the same data required for SED models, such as telescope imagery. After training, the model can then automatically classify many other data such as images. In this Thesis, we train machine learning models using manually selected training sites from Hubble Space Telescope imagery in order to determine how machine learning can contribute to improvement of the galaxy component and stellar age classification methods. Because SED models are not able to account for the diverse populations of components within galaxies, manual selection may account for this, thus improving stellar property estimation. Texture features used for

morphological classification and outlier identification in astronomical research can contribute to a more efficient method of stellar age classification due to the different patterns observed in stellar populations of different ages. Here we also attempt to fill this knowledge gap and test texture features for galaxy component and stellar age classification.

This Thesis includes the research done in two separate studies: one focusing on classification of all galaxy components within UGC 2885, and the other focusing on classification of stellar age within M83. The overall objective of this thesis is to determine how machine learning can be used to automate classification of components and their properties within spiral galaxies. We identify the following objectives:

1. Evaluate how machine learning methods can be used to classify galaxy components and stellar age within spiral galaxies and determine how reference data affects classification accuracy of stellar age;
2. Determine the most useful parameters for galaxy component and stellar age classification;
3. Determine which galaxy components and stellar ages are best classified.

## 1.2 Study Areas

In this thesis we focus on two spiral galaxies. The first is UGC 2885, about 79.1 megaparsecs (Mpc) or about 258 million light years away from us, which is one of the most massive spiral galaxies in the nearby Universe (Figure 1-2; Hunter et al., 2013). We study the galaxy component populations within UGC 2885 as it is an excellent study area due to the diversity of galaxy components within it. The second is M83, sitting much closer at 4.61 Mpc or 15 million light years, and exhibiting high amounts of star formation (Figure 1-3; Blair et al., 2014). We study the age of stellar populations within M83.

UGC 2885 has a diameter of 5.5 arcminutes as projected on the sky (Rubin et al., 1980). Although UGC 2885 is a larger galaxy than M83, M83 appears larger on the sky with a diameter of 11.5 arcminutes due to its nearer proximity (Sánchez-Gil et al., 2019).

Arcminutes are the celestial coordinate system equivalent to minute measurements used in Earth-based coordinate systems.



**Figure 1-2. Hubble Space Telescope image of galaxy UGC 2885. Image credit: NASA, ESA, and B. Holwerda (University of Louisville).**



**Figure 1-3. Hubble Space Telescope image of galaxy M83. Image credit: NASA, ESA, and the Hubble Heritage Team (STScI/AURA); Acknowledgement: W. Blair (STScI/Johns Hopkins University) and R. O'Connell (University of Virginia).**

## 1.3 References

- Abdeen, S., Davis, B. L., Eufrazio, R., Kennefick, D., Kennefick, J., Miller, R., Shields, D., Monson, E. B., Bassett, C., & O'Mara, H. (2022). Evidence in favour of density wave theory through age gradients observed in star formation history maps and spatially resolved stellar clusters. *Monthly Notices of the Royal Astronomical Society*, 512(1), 366–377. <https://doi.org/10.1093/mnras/stac459>
- Barmby, P. (2019). Astronomical observations: A guide for allied researchers. *The Open Journal of Astrophysics*, 2(1), 10.21105/astro.1812.07963. <https://doi.org/10.21105/astro.1812.07963>
- Bennett, J. O., Donahue, M., Schneider, N., & Voit, M. (2015). *The essential cosmic perspective (Seventh edition)*. Pearson.
- Bialopetravičius, J., & Narbutis, D. (2020). Study of star clusters in the M83 galaxy with a convolutional neural network. *The Astronomical Journal*, 160(6), 264. <https://doi.org/10.3847/1538-3881/abbf53>
- Blair, W. P., Chandar, R., Dopita, M. A., Ghavamian, P., Hammer, D., Kuntz, K. D., Long, K. S., Soria, R., Whitmore, B. C., & Frank Winkler, P. (2014). An expanded HST/WFC3 survey of M83: project overview and targeted supernova remnant search. *The Astrophysical Journal*, 788(1), 55. <https://doi.org/10.1088/0004-637X/788/1/55>
- Bovy, J., Rix, H.-W., Schlafly, E. F., Nidever, D. L., Holtzman, J. A., Shetrone, M., & Beers, T. C. (2016). The stellar population structure of the galactic disk. *The Astrophysical Journal*, 823(1), 30. <https://doi.org/10.3847/0004-637X/823/1/30>
- Brown, T. M. (2005). Star formation histories in the Local Group. *New Astronomy Reviews*, 49(7–9), 474–477. <https://doi.org/10.1016/j.newar.2005.08.025>
- Bruzual A., G. (2010). Star clusters as simple stellar populations. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 368(1913), 783–799. <https://doi.org/10.1098/rsta.2009.0258>

- Cassan, A., Kubas, D., Beaulieu, J.-P., Dominik, M., Horne, K., Greenhill, J., Wambsganss, J., Menzies, J., Williams, A., Jørgensen, U. G., Udalski, A., Bennett, D. P., Albrow, M. D., Batista, V., Brilliant, S., Caldwell, J. A. R., Cole, A., Coutures, Ch., Cook, K. H., ... Wyrzykowski, Ł. (2012). One or more bound planets per Milky Way star from microlensing observations. *Nature*, *481*(7380), 167–169. <https://doi.org/10.1038/nature10684>
- Côté, P., Abraham, B., Balogh, M., Capak, P., Carlberg, R., Cowan, N., Djazovski, O., Drissen, L., Drout, M., Dupuis, J., Evans, C., Fantin, N., Ferrarese, L., Fraser, W., Gallagher, S., Girard, T., Gleisinger, R., Grandmont, F., Hall, P., ... Venn, K.. (2019). *CASTOR: A Flagship Canadian Space Telescope*. Zenodo. <https://doi.org/10.5281/ZENODO.3758463>
- Diaz-Hernandez, R., Ortiz-Esquivel, A., Peregrina-Barreto, H., Altamirano-Robles, L., & Gonzalez-Bernal, J. (2016). Automatic approach to solve the morphological galaxy classification problem using the sparse representation technique and dictionary learning. *Experimental Astronomy*, *41*(3), 409–426. <https://doi.org/10.1007/s10686-016-9495-0>
- Dobbs, C. L., & Pringle, J. E. (2010). Age distributions of star clusters in spiral and barred galaxies as a test for theories of spiral structure. *Monthly Notices of the Royal Astronomical Society*, *409*(1), 396–404. <https://doi.org/10.1111/j.1365-2966.2010.17323.x>
- Fluke, C. J., & Jacobs, C. (2020). Surveying the reach and maturity of machine learning and artificial intelligence in astronomy. *WIREs Data Mining and Knowledge Discovery*, *10*(2). <https://doi.org/10.1002/widm.1349>
- Gardner, J. P., Mather, J. C., Clampin, M., Doyon, R., Greenhouse, M. A., Hammel, H. B., Hutchings, J. B., Jakobsen, P., Lilly, S. J., Long, K. S., Lunine, J. I., Mccaughrean, M. J., Mountain, M., Nella, J., Rieke, G. H., Rieke, M. J., Rix, H.-W., Smith, E. P., Sonneborn, G., ... Wright, G. S. (2006). The James Webb Space

Telescope. *Space Science Reviews*, 123(4), 485–606.

<https://doi.org/10.1007/s11214-006-8315-7>

Hunter, D. A., Elmegreen, B. G., Rubin, V. C., Ashburn, A., Wright, T., Józsa, G. I. G., & Struve, C. (2013). Star formation in two luminous spiral galaxies. *The Astronomical Journal*, 146(4), 92. <https://doi.org/10.1088/0004-6256/146/4/92>

Jensen, E. B., Talbot, R. J., Jr., & Dufour, R. J. (1981). M83. III - Age and brightness of young and old stellar populations. *The Astrophysical Journal*, 243, 716.

<https://doi.org/10.1086/158640>

King, I. R. (1971). Stellar populations in galaxies. *Publications of the Astronomical Society of the Pacific*, 83, 377.

Laureijs, R., Amiaux, J., Arduini, S., Auguères, J.-L., Brinchmann, J., Cole, R., Cropper, M., Dabin, C., Duvet, L., Ealet, A., Garilli, B., Gondoin, P., Guzzo, L., Hoar, J., Hoekstra, H., Holmes, R., Kitching, T., Maciaszek, T., Mellier, Y., ... Zucca, E. (2011). *Euclid Definition Study Report*. ArXiv:1110.3193 [Astro-Ph].

<http://arxiv.org/abs/1110.3193>

Lee, J. H., Kim, S. C., Park, H. S., Ree, C. H., Kyeong, J., & Chung, J. (2011). Hubble Space Telescope pixel analysis of the interacting face-on spiral galaxy NGC 5194 (M51A). *The Astrophysical Journal*, 740(1), 42.

<https://doi.org/10.1088/0004-637X/740/1/42>

Lillesand, T. M., Kiefer, R. W., & Chipman, J. W. (2015). *Remote sensing and image interpretation (Seventh edition)*. John Wiley & Sons, Inc.

Lin, C. C., & Shu, F. H. (1964). On the spiral structure of disk galaxies. *The Astrophysical Journal*, 140, 646. <https://doi.org/10.1086/147955>

Lingard, T. K., Masters, K. L., Krawczyk, C., Lintott, C., Kruk, S., Simmons, B., Simpson, R., Bamford, S., Nichol, R. C., & Baeten, E. (2020). Galaxy Zoo Builder: Four-component photometric decomposition of spiral galaxies guided

by citizen science. *The Astrophysical Journal*, 900(2), 178.

<https://doi.org/10.3847/1538-4357/ab9d83>

Lintott, C. J., Schawinski, K., Slosar, A., Land, K., Bamford, S., Thomas, D., Raddick, M. J., Nichol, R. C., Szalay, A., Andreescu, D., Murray, P., & Vandenberg, J. (2008). Galaxy Zoo: Morphologies derived from visual inspection of galaxies from the Sloan Digital Sky Survey. *Monthly Notices of the Royal Astronomical Society*, 389(3), 1179–1189. <https://doi.org/10.1111/j.1365-2966.2008.13689.x>

Long, J. P., & de Souza, R. S. (2017). *Statistical methods in astronomy*.

<https://doi.org/10.48550/ARXIV.1707.05834>

Ntwaetsile, K., & Geach, J. E. (2021). Rapid sorting of radio galaxy morphology using Haralick features. *Monthly Notices of the Royal Astronomical Society*, 502(3), 3417–3425. <https://doi.org/10.1093/mnras/stab271>

Pössel, M. (2020). A beginner's guide to working with astronomical data. *The Open Journal of Astrophysics*, 3(1), 10.21105/astro.1905.13189.

<https://doi.org/10.21105/astro.1905.13189>

Rubin, V. C., Thonnard, N., & Ford, W. K., Jr. (1980). Rotational properties of 21 SC galaxies with a large range of luminosities and radii, from NGC 4605 / $R = 4\text{kpc}$ / to UGC 2885 / $R = 122\text{ kpc}$ /. *The Astrophysical Journal*, 238, 471.

<https://doi.org/10.1086/158003>

Sánchez-Gil, M. C., Alfaro, E. J., Cerviño, M., Pérez, E., Bland-Hawthorn, J., & Jones, D. H. (2019). Hierarchical Bayesian approach for estimating physical properties in nearby galaxies: Age Maps (Paper II). *Monthly Notices of the Royal Astronomical Society*, 483(2), 2641–2670. <https://doi.org/10.1093/mnras/sty3106>

Sánchez-Gil, M. C., Jones, D. H., Pérez, E., Bland-Hawthorn, J., Alfaro, E. J., & O'Byrne, J. (2011). Age patterns in a sample of spiral galaxies: Age patterns in a sample of spiral galaxies. *Monthly Notices of the Royal Astronomical Society*, 415(1), 753–772. <https://doi.org/10.1111/j.1365-2966.2011.18759.x>

- Scheepmaker, R. A., Lamers, H. J. G. L. M., Anders, P., & Larsen, S. S. (2009). The spatial distribution of star and cluster formation in M 51. *Astronomy & Astrophysics*, *494*(1), 81–93. <https://doi.org/10.1051/0004-6361:200811068>
- Shabani, F., Grebel, E. K., Pasquali, A., D’Onghia, E., Gallagher, J. S., Adamo, A., Messa, M., Elmegreen, B. G., Dobbs, C., Gouliermis, D. A., Calzetti, D., Grasha, K., Elmegreen, D. M., Cignoni, M., Dale, D. A., Aloisi, A., Smith, L. J., Tosi, M., Thilker, D. A., ... Pellerin, A. (2018). Search for star cluster age gradients across spiral arms of three LEGUS disc galaxies. *Monthly Notices of the Royal Astronomical Society*, *478*(3), 3590–3604. <https://doi.org/10.1093/mnras/sty1277>
- Shamir, L. (2021). Automatic identification of outliers in Hubble Space Telescope galaxy images. *Monthly Notices of the Royal Astronomical Society*, *501*(4), 5229–5238. <https://doi.org/10.1093/mnras/staa4036>
- Shu, F. H. (2016). Six decades of spiral density wave theory. *Annual Review of Astronomy and Astrophysics*, *54*(1), 667–724. <https://doi.org/10.1146/annurev-astro-081915-023426>
- Spergel, D., Gehrels, N., Baltay, C., Bennett, D., Breckinridge, J., Donahue, M., Dressler, A., Gaudi, B. S., Greene, T., Guyon, O., Hirata, C., Kalirai, J., Kasdin, N. J., Macintosh, B., Moos, W., Perlmutter, S., Postman, M., Rauscher, B., Rhodes, J., ... Zhao, F. (2015). *Wide-Field Infrared Survey Telescope-Astrophysics Focused Telescope Assets WFIRST-AFTA 2015 Report*. ArXiv:1503.03757 [Astro-Ph]. <http://arxiv.org/abs/1503.03757>
- van den Bergh, S. (1975). Stellar populations in galaxies. *Annual Review of Astronomy and Astrophysics*, *13*(1), 217–255. <https://doi.org/10.1146/annurev.aa.13.090175.00124575.001245>
- Walcher, J., Groves, B., Budavári, T., & Dale, D. (2011). Fitting the integrated spectral energy distributions of galaxies. *Astrophysics and Space Science*, *331*(1), 1–51. <https://doi.org/10.1007/s10509-010-0458-z>

Wei, P., Zou, H., Lin, L., Zhou, X., Liu, X., Kong, X., Ma, L., & Ma, S.-G. (2021). Spatially-resolved stellar population properties of the M51–NGC 5195 system from multi-wavelength photometric data. *Research in Astronomy and Astrophysics*, 21(1), 006. <https://doi.org/10.1088/1674-4527/21/1/6>

Whitmore, B. C., Chandar, R., Kim, H., Kaleida, C., Mutchler, M., Stankiewicz, M., Calzetti, D., Saha, A., O'Connell, R., Balick, B., Bond, H. E., Carollo, M., Disney, M. J., Dopita, M. A., Frogel, J. A., Hall, D. N. B., Holtzman, J. A., Kimble, R. A., McCarthy, P. J., ... Young, E. T. (2011). Using H $\alpha$  morphology and surface brightness fluctuations to age-date star clusters in M83. *The Astrophysical Journal*, 729(2), 78. <https://doi.org/10.1088/0004-637X/729/2/78>

## Chapter 2

# 2 Galactic Component Mapping of Galaxy UGC 2885 by Machine Learning Classification

## 2.1 Introduction

Galaxies contain stars, dust, and gas that make up larger bulge, bar, disc, and spiral arm structures. Nearby galaxies – galaxies in close enough proximity that their components can be resolved in high resolution imagery – are particularly useful for study of the spatial and temporal relationships. By studying galaxies, we can learn about the structure and evolution of our own Milky Way galaxy. Studying the evolution of galaxies is an important step to understanding how all matter formed. Further, by focusing on nearby galaxies we can apply our gained knowledge to more distant galaxies that we cannot observe in full detail (Bianchi et al., 2014; Kalirai, 2018). It is also known that deconstruction of spiral galaxies into their components plays a key role in understanding the nature of galactic evolution and the structural properties of components such as stars, dust, and gas (Lingard et al., 2020). Therefore, studying the spatial distribution of galactic components will broaden the understanding of the photometric properties of galaxies. By automating the process of galactic component analysis, the same methodologies can be applied to high-resolution imagery of similar nearby galaxies such as the Legacy ExtraGalactic Ultraviolet Survey (LEGUS) and the Physics at High Angular Resolution in Nearby Galaxies with the Hubble Space (PHANGSHST) survey (Calzetti et al., 2015; Lee et al., 2021).

Bulge-disc decomposition of galaxies – the photometric or spectroscopic separation of the bulge and disc regions of a spiral galaxy within digital imagery – is a well researched area of study (a recent example: Pak et al., 2021). However, to our knowledge, identification and mapping of all galactic components within a galaxy’s bulge and disc has not been done. Although an expert can perform visual inspection and identification of galaxy components, the automation of this process can shorten the time required to analyze digital imagery of galaxies. By classifying each component within digital imagery of spiral galaxies, the fine details we observe and identify are remnants of star formation history of the galaxy and can provide clues to the evolution and formation of matter within the galaxy (Peng et al.,

2002). The use of human visual interpretation is not a new idea, but no study has used this method for pixel-by-pixel component identification of high resolution imagery. Citizen science based projects such as Galaxy Zoo (Lingard et al., 2020; Lintott et al., 2008) work to expand upon the traditional bulge-disc classification by having participants identify components within digital imagery of galaxies. However, these studies focus on identifying substructures within the bulge and disc such as spiral arms and bars. The success of human visual interpretation to identify galaxy structures demonstrates the usefulness of visual observation of digital imagery. By surpassing the traditional bulge-disc decomposition and digging deeper into galactic structure, we can aid in the quantitative understanding of galactic component distribution and evolution (Lingard et al., 2020).

In recent years, there has been growing interest in machine learning (ML) for digital image analysis (Baron, 2019; Thanh Noi & Kappas, 2017). By supervising the identification of pixels that represent different classes, the machine learns the photometric characteristics of classes and can then automatically classify individual pixels in digital imagery based on the algorithm's acquired knowledge. Several recent studies have used pixel-based ML to classify galaxies. Hausen & Robertson (2020) use ML methods to classify morphologies or types of galaxies within a Hubble image by use of a pixel-based method. Bialopetravičius & Narbutis (2020) use ML to identify star clusters within a nearby nearly face-on galaxy. Both studies emphasize the need for automation of high resolution galaxy classification methods. However, their focus on morphology and star clusters neglects other significant galaxy components such as dust lanes and stellar populations that are not members of clusters. This knowledge gap is significant as pixel based classification of galactic populations is necessary for mapping complex distributions of galactic material and for better understanding the complex relationships within galaxies. Further, the upcoming Euclid (Laureijs et al., 2011) and Roman Space Telescope (Spergel et al., 2015) will make available more high-resolution Hubble-like data for nearby galaxies, making pixel-based mapping more feasible.

Texture analysis is a commonly used image processing technique in earth-based remote sensing. Several studies have explored the usefulness of texture for morphological analysis of galaxies (Au, 2006; Ntwaetsile & Geach, 2021; Pedersen et al., 2013; Schutter &

Shamir, 2015; Shamir, 2009; Shamir, 2021). Ntwaetsile & Geach (2021) find that texture analysis is particularly useful for radio galaxy morphology analysis and recommend that it be applied to large imaging surveys. Similarly, Shamir (2021) notes the usefulness of texture analysis for identifying outlier galaxies in optical imagery of galaxies. Both examples demonstrate the diverse applications of texture for astronomical imagery. However, to our knowledge texture has yet to be tested for identification of galactic components within digital imagery. Because texture is useful for summarizing imagery, we expect texture to be particularly useful for identifying the differences between the fine details of galaxy components within high resolution digital imagery.

A geographic information system (GIS) method not commonly used in classification of celestial phenomena is distance as calculated for the pixel contents of a digital imagery. Because components of galaxies are arranged based on spatial and temporal patterns, distance measures commonly used in GIS for Earth-based phenomena are compatible with these galaxies. In one particular instance of use of distance for astronomical research, Bialopetravičius & Narbutis (2020) make use of distance from galaxy center and spiral arms to observe the relationship between distance from spiral arms and galaxy center to the age of stars within galaxy M83. However, the distance measures are implemented for post classification analysis only. To our knowledge no study has incorporated per pixel distance measures within the galaxy plane to classify all galaxy components within digital imagery.

By addressing the below research objectives, we hope to better understand dynamics of nearly edge-on nearby galaxies and the most efficient method of identifying galactic components within digital imagery through machine learning classification:

- (1) Evaluate how machine learning can be used to classify galactic components,
- (2) Determine which input parameters are most useful,
- (3) Determine which machine learning method is most useful.

## 2.2 Methods

### 2.2.1 Study Area

Galaxy UGC 2885 or ‘Rubin’s Galaxy’ is an unusually large and late type (Sc) spiral galaxy sitting 79.1 Mpc or approximately 258,000,000 light years away (Hunter et al., 2013; Figure 2-1). Rubin’s Galaxy is at a suitable distance for observation of galactic components; at these distances, the stellar field is so dense that most pixels contain the integrated light from multiple stars. The massive size of UGC 2885 –approximately 44.4 kpc or 145,000 light years in diameter (Hunter et al., 2013) – makes it an interesting case study for mapping of the populations within it. By studying the distribution of galactic components within UGC 2885, we can better understand the nature of spatial and temporal patterns within the galaxy, as well as how massive galaxies differ from galaxies with more common properties. UGC 2885 is also inclined 74 (Hunter et al., 2013), meaning that it is nearly edge-on, where a completely edge-on galaxy is 90. The inclination of a galaxy is defined relative to the point of view of an observer.

At the center of UGC 2885, a supermassive black hole has been identified (Holwerda et al., 2021). The massive size of UGC 2885, the presence of a supermassive black hole, and the lack of star formation within the galaxy make UGC 2885 defy easy morphological classification. As Holwerda et al. (2021) note, classifying the small structures within the galaxy can contribute to a better understanding of UGC 2885 and similarly unusual galaxies. Ultimately, UGC 2885 is an optimal study area for galactic component mapping due to the large population of components within the galaxy. This particular galaxy also exhibits spatial and temporal gradients of stars. For instance, the galactic center is redder due to the large number of old stars within it. On the contrary, the spiral arms have a larger concentration of young stellar populations making them appear bluer.

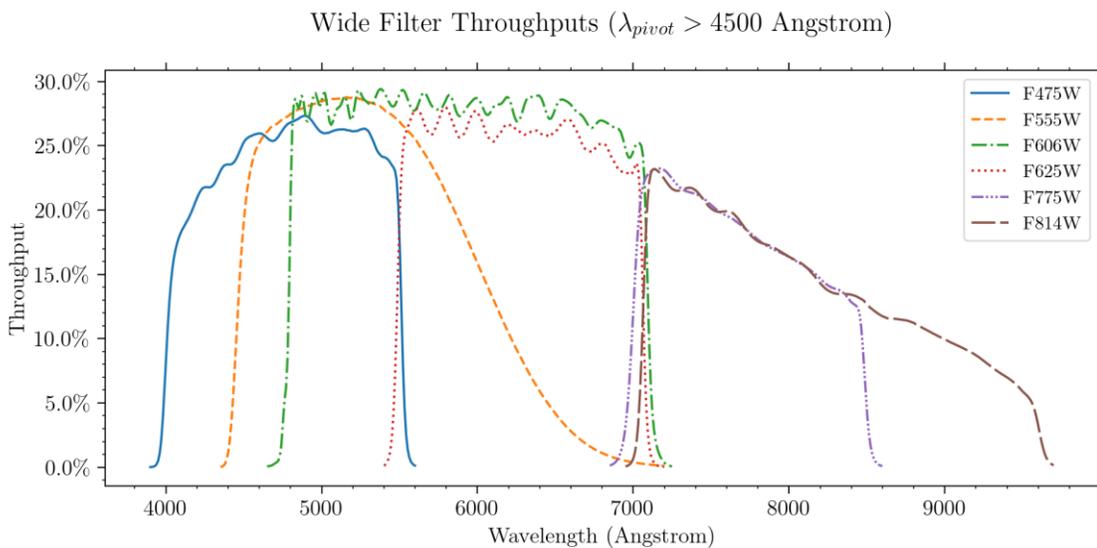


**Figure 2-1: HST colour composite map of UGC 2885 where the F475W band is displayed in blue, the F606W band is displayed in green, and the F814W band is displayed in red. This image is in celestial orientation; therefore, east is towards the left rather than to the right of the north direction. Parsecs refers to a measure of distance equaling approximately 3.26 light years. Image credit: NASA, ESA, and B.W. Holwerda (University of Louisville).**

### 2.2.2 Data

We use the Hubble Space Telescope (HST) multispectral digital imagery – the highest resolution imagery currently available in visible and near infrared wavelengths – for UGC 2885 in three wavelength bands. Holwerda et al. (2020) generously provide mosaics for all three wavelength bands. Observations in three wavelength bands – band F475W, F606W, and F814W – are available as part of HST program 15107, The Cluster Population of UGC 2885 (Holwerda, 2017). Figure 2-2 shows the filter wavelength ranges. Band F475W shows the blue-green (B) emission (Fukugita et al., 1996) and covers a broad range of

134.35 nm around a central wavelength of 475 nm. The F475W filter is the equivalent of the g' filter of the Sloan Digital Sky Survey system (Dressel et al., 2022). The widest band, F606W, shows visual (V) emission over range of 218.92 nm, with a central wavelength of 606 nm. The V band central wavelength is the equivalent to green wavelength band in remote sensing applications. Lastly, band F814W shows near infrared (I) emission over a wavelength range of 156.52 nm, with a central wavelength of 814 nm (Dressel, 2021). The F814W band's response peaks in the near infrared, and drops off steeply at the longer wavelengths.



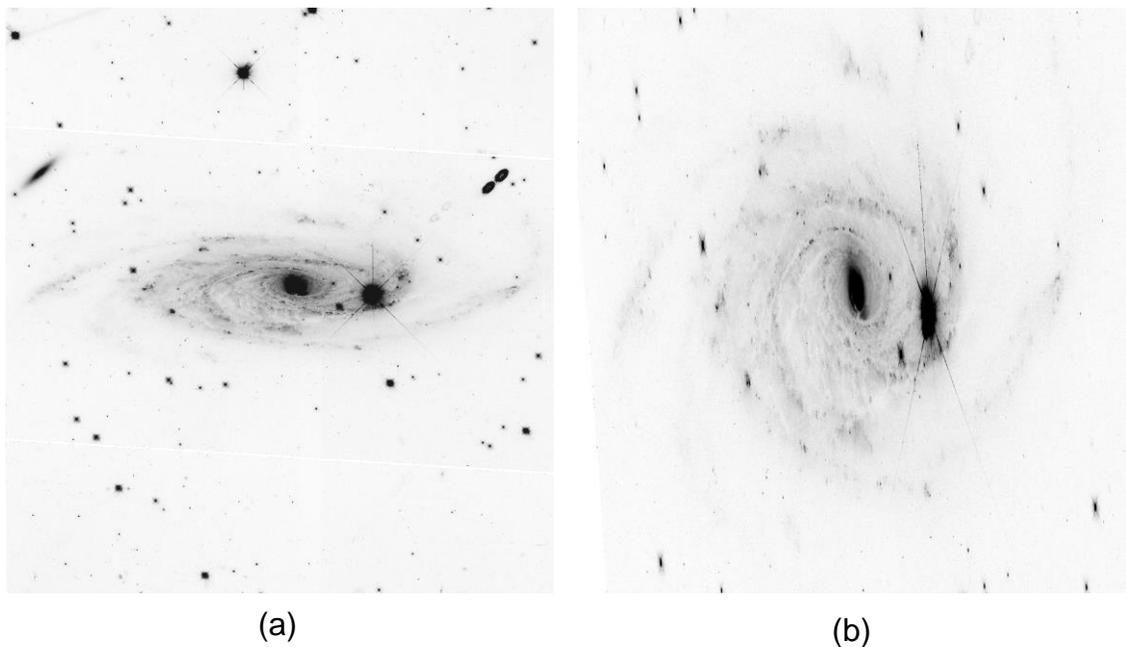
**Figure 2-2: HST wide band filter wavelength ranges. (Dressel et al., 2022)**

BVI images are significant for mapping stellar populations and other galactic components within nearby galaxies (Holwerda et al., 2020; Kiar et al., 2017). More specifically, the B band is useful for observing younger and hotter stars while the V and I bands are useful for identifying the cooler and redder stars (Kiar et al., 2017). The BVI bands can also be used to observe dust lanes throughout the galaxy; in particular, because the dust lanes are redder in wavelength, the dust lanes emit stronger in the V and I bands. Therefore, the HST BVI bands are useful for observation of the major galaxy components within the UGC 2885.

The HST mosaics (Holwerda et al., 2020) are in Flexible Image Transport System (FITS) format. We use FITS Liberator 3 software (ESA/ESO/NASA, 2021) to export the images

to Tag Image File Format (TIFF) using a logarithmic stretch. The logarithmic stretch is for visualization purposes only and does not change the original values of the HST imagery.

Due to the 74° inclination of UGC 2885, we also deproject the digital imagery to approximate a face on galaxy to calculate Euclidean distance as described in Section 2.3.4. (Figure 2-3). We use deprojection only for distance layer generation. To deproject the digital imagery, we calculate a trigonometric stretch ratio of  $1/\cos(74)$  and find a value of 3.628. Therefore, we stretch the digital imagery vertically by 363% using a raster graphics editor; see also Davis et al. (2012) for an excellent description of deprojection of similarly inclined galaxies. Due to the presence of foreground stars in the HST imagery, we also mask the 20 brightest stars within a three arcminute radius of UGC 2885 from Gaia Early Data Release 3 (EDR3) data (Gaia Collaboration et al., 2016; Gaia Collaboration et al., 2021).



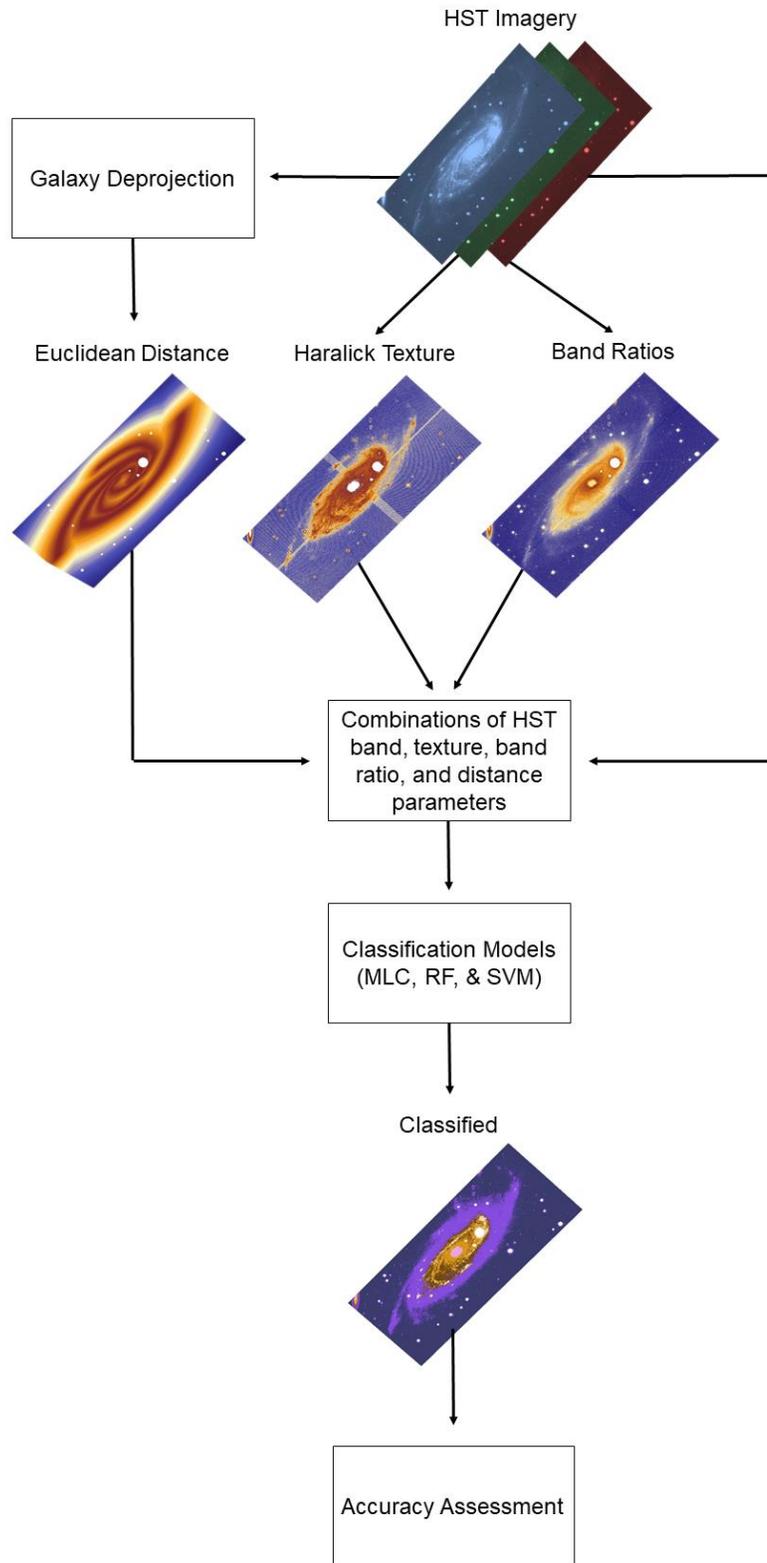
**Figure 2-3: UGC 2885 blue-green band imagery (F475W): a) original HST imagery showing the projected galaxy with an inclination of 74°; b) deprojected image of UGC 2885 – stretched vertically by 363%.**

## 2.3 HST Image Processing

### 2.3.1 Coordinate Transformation

For spatial analysis in a GIS, imagery must be transformed to an Earth-based coordinate system. We retrieve coordinates for features within the image using SAOImage DS9 (Version 8.0.1.) open source software (Joye & Mandel, 2003). HST images of UGC 2885 are transformed using the Helmert transformation (Farhadian & Clarke, 2020) available within the “Georeferencer” tool in an open source software QGIS Version 3.16.3. (QGIS 3.16., 2021). The Helmert transformation shifts and rotates the image through affine methods that preserve the collinearity and ratio of distance of the features in an image (Song et al., 2014). Testing other transformations in QGIS resulted in highly distorted imagery, leading us to use the Helmert transformation.

Figure 2-4 shows the HST image processing methodology following coordinate transformation. With the transformed HST imagery, we create an image stack in GIS software. From the HST raster composite, we produce textural features, band ratios, and distance layers. We further describe the methodology (Figure 2-4) in the following sections.



**Figure 2-4: Flowchart showing methods used in this study.**

### 2.3.2 Band Ratio

Band ratios are effective as they are useful at describing how stars emit light and how absorption affects that light. For instance, bluer stars are more affected by dust lane absorption so will have different ratio values than a red star affected by the same dust. We use several broadband ratios commonly used in astronomy to identify particular aspects of UGC 2885. Kiar et al. (2017) provide a good overview of HST band ratios. We specify similar band ratios for UGC 2885 in Table 2-1 and expect the ratios to emphasize different phenomena such as star clusters within the galaxy. We calculate band ratios by division of wavelength bands in a GIS software. Typically, the shorter wavelength band by is divided by the longer wavelength band when calculating simple ratios (i.e. B/V). Band ratios used in classification are described below.

**Table 2-1: Band ratios used for classification and their respective equations.**

<b>Flux Band Ratio</b>	<b>Fraction Band Ratio</b>	<b>Spectral Slope Band Ratio</b>
B-V	$B/(B+V+I)$	$(B/V)/(B+V+I)$
B-I	$V/(B+V+I)$	$(B/I)/(B+V+I)$
V-I	$I/(B+V+I)$	$(V/I)/(B+V+I)$

Before creating ratios, we calibrate the HST bands F475W, F606W, and F814W with their respective Flexible Image Transport System (FITS) calibration factors found in the Imager Header tab within FITS Liberator 3 software (ESA/ESO/NASA, 2021; Table 2-2). The process of calibration involves multiplying each pixel in a raster image by a certain value in order to convert the pixels into meaningful units; we convert the original 32-bit pixel range to Jansky units which are equal to  $10^{-26}$  Watts metre<sup>-2</sup> Hertz<sup>-1</sup>. For instance, we use Eq. (1) to calibrate the F475W blue-green band:

$$\text{Calibrated Raster} = \text{“F475W raster”} * 1.8782514\text{E-}07 \quad (1)$$

We repeat the above calculation for both the F606W and F814W bands. After calibration, we compute the flux, fraction, and spectral slope band ratios.

**Table 2-2: Calibration factors for the BVI imagery as found in FITS Liberator 3 Image Headers tab under “PHOTFNU”.**

<b>Band</b>	<b>Calibration Factor</b>
F475W (Blue-green)	1.8782514E-07 / Inverse sensitivity, Jy*sec/e-
F606W (Visual)	1.32242795E-07 / Inverse sensitivity, Jy*sec/e-
F814W (Near-Infrared)	3.2380001E-07 / Inverse sensitivity, Jy*sec/e-

Flux band ratios involve simple band division. Three combinations of flux band ratios are possible with HST data available for UGC 2885: B-V, B-I, and V-I. For the B-V ratio, we perform division of blue-green and visual bands (B/V) in a GIS software where the highest pixel values are those having a higher value in the blue-green band and a lower value in the visual band. The same concept applies to the B-I and V-I ratios.

To calculate fraction ratios, we divide a band by the flux sum of all three bands. For instance,  $B / (B + V + I)$  calculates the fraction of blue-green-to-total light emitted from UGC 2885. The fraction band ratios compare the band flux of a single band to the total brightness, therefore they identify the most prominent emission wavelengths for a given pixel. For instance, the B-fraction band ratio ( $B / (B + V + I)$ ) will identify the bluest sources within the galaxy, these being the young stars. B-fraction band ratio also identifies dust lanes. Dust is a better absorber of blue light than red light, so the dust appears darkest in the B band. The V band shows a wider range of visual light, therefore the V-fraction band ratio ( $V / (B + V + I)$ ) shows many galaxy components. In particular, it does a good job of emphasizing the structure of the galaxy in both the inner and outer disc that contain old and young stars. We can also make out clear dust lanes in the V-fraction band ratio image. Near-infrared (I) will have a larger emission of light for the old stars, dust lanes, and galaxy center, so the I fraction band ratio ( $I / (B + V + I)$ ) emphasizes the inner disc of the galaxy where the old stars are accumulated.

The final band ratio, spectral slope, calculates the spectral slope of two bands over the flux product of the three bands. For example,  $(B/V) / (B + V + I)$ , calculates the blue-green/visual slope. The spectral slope ratio accounts for any correlation between colour, as measured by band ratios, and overall brightness, as measured by the band sum, making

them akin to colour-magnitude diagrams used in astronomy. These quantities might be expected to correlate because dust within a galaxy will make the emergent light both fainter and redder. Colour and brightness are direct indicators of stellar age: young stars are brighter and bluer while older stars are dimmer and redder. The BV spectral slope ratio  $((B/V)/(B + V + I))$  shows the most detail within the inner disc whereas in the BI spectral slope ratio  $((B/I)/(B + V + I))$  image it is more difficult to distinguish the galaxy center. The VI spectral slope ratio  $((V/I)/(B + V + I))$  image places most emphasis on the galaxy center region.

### 2.3.3 Texture Features

Texture features imitate visual patterns we see in objects, area, and phenomena. We calculate Haralick Grey Level Co-occurrence Matrix (GLCM) textures (Haralick et al., 1973) for HST imagery of UGC 2885. The GLCM is produced from all pixel grey level values within a moving window of specified size and considers the grey levels of two pixels at a time, the reference pixel and the neighbouring pixel. For instance, a 5x5 window will produce a GLCM from 25 pixels. The number of grey levels chosen by the user determines the size of the GLCM, and the pixel values of the original imagery are scaled down to the chosen number of grey levels. On the position of the GLCM where the grey levels of the reference and neighbour pixel meet, 1 is added to that position. After the GLCM is produced, second order statistics are calculated based off of the contents of the GLCM (Haralick et al., 1973).

In galactic imagery, dust lanes often exhibit a rougher texture while the bright galaxy center and star clusters have a smooth appearance. We aim to determine if texture features can identify these differences and increase accuracy of classification. GLCM textures have been successfully tested for many remote sensing research applications of Earth-based phenomena (Ghasemian & Akhoondzadeh, 2018; Hall-Beyer, 2017; Wei et al., 2021; Zhang et al., 2014), making texture a promising prospect for astronomical research. Textures are excellent at rapidly summarizing the contents of an image, so are useful for processing the abundance of astronomical data for machine learning classification (Ntwaetsile & Geach, 2021).

We use eight Haralick textures (Haralick et al., 1973) available in a commercial remote sensing software; textures include angular second moment, contrast, correlation, dissimilarity, entropy, homogeneity, mean, and standard deviation. We specify 64 grey levels and a sliding window size of 5x5. Any grey level can be chosen, but we specify 64 as we find it suitable for the HST imagery. Further, we choose a window size of 5x5 pixels based on the measured pixel width of a typical star cluster within the HST BVI imagery. We describe textures below and show their appearance when calculated for HST band F606W in Figure 2-5.

Entropy texture calculates unevenness of the image grey levels respectively (Wei et al., 2021) and is represented by the following equation:

$$Entropy = - \sum_{i=1}^L \sum_{j=1}^L \hat{P}(i,j) \log[\hat{P}(i,j)] \quad (2)$$

where  $P(i,j)$  is the  $(i,j)$ th entry in a normalized GLCM. Entropy texture is high when an image has a large range of grey levels, therefore having unevenness.

Angular second moment calculates textures based on uniformity of the imagery. Eq. (3) describes how angular second moment is calculated:

$$Angular\ second\ moment = \sum_{i=1}^L \sum_{j=1}^L (\hat{P}(i,j))^2 \quad (3)$$

Images with a larger number of grey levels have smaller uniformity and therefore smaller values of angular second moment texture.

Homogeneity looks at the evenness or homogeneous nature of the spectral characteristics throughout an image and is calculated by the following Eq. (4):

$$Homogeneity = \sum_{i=1}^L \sum_{j=1}^L \frac{\hat{P}(i,j)}{1+(i-j)^2} \quad (4)$$

Mean calculates the average of the grey-levels in the GLCM local window and is defined by Eq. (5):

$$Mean = \sum_{i=1}^L \sum_{j=1}^L i * \hat{P}(i,j) \quad (5)$$

Contrast and dissimilarity are very similar to each other in that they both measure the spectral variation within the local GLCM window. However, they are different in that contrast incorporates the square root of the difference between i and j co-occurrence matrix where dissimilarity uses the absolute difference between i and j co-occurrence matrix. Contrast is defined as:

$$\text{Contrast} = \sum_{n=0}^{L-1} (i - j)^2 \sum_{i=1}^L \sum_{j=1}^L \hat{P}(i, j) \quad (6)$$

where high contrast indicates textures with sharp edges in an image. Dissimilarity is described by the following equation:

$$\text{Dissimilarity} = \sum_{i=1}^L \sum_{j=1}^L |i - j| \hat{P}(i, j) \quad (7)$$

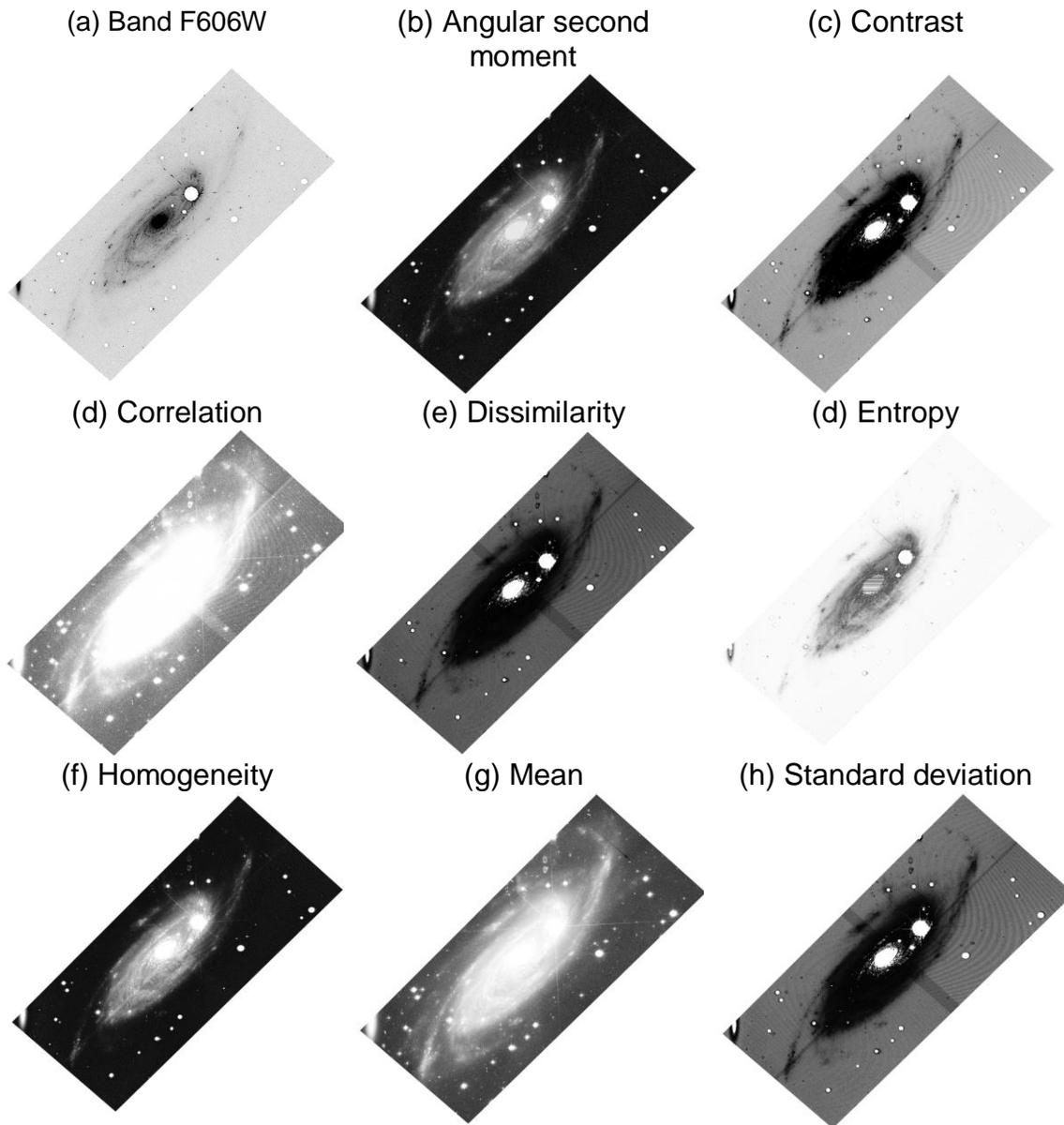
Standard deviation texture calculates the standard deviation or scattering of the local spectral information with respect to the mean. Areas where pixels have a small range of values have lower standard deviation while areas with high pixel ranges have higher standard deviation. Eq. (8) defines standard deviation:

$$\text{Standard Deviation} = \sum_{i=1}^L \sum_{j=1}^L (i - \mu)^2 \hat{P}(i, j) \quad (8)$$

Correlation is a measure of linear dependency of the spectral variation on local pixels within the GLCM window. High values show areas where noise or sharp changes are present in the image. Correlation is described by Eq. (9):

$$\text{Correlation} = \sum_{i=1}^L \sum_{j=1}^L ((i * j * \hat{P}(i, j) - \mu_x \mu_y) / \sigma_x \sigma_y) \quad (9)$$

where  $\mu_x$  and  $\mu_y$  are the means of  $P_x$  and  $P_y$ , and  $\sigma_x$  and  $\sigma_y$  are the standard deviations of  $P_x$  and  $P_y$ .



**Figure 2-5: Band F606W Haralick textures with 20 brightest foreground stars masked.**

#### 2.3.4 Distance from Spiral Arms and Galaxy Center

Distance information is useful in spiral galaxies that exhibit age gradients. Galaxies that are dominated by spiral density waves (Lin & Shu, 1964) are most compatible with measurements of distance. Density waves are thought to be present in UGC2885 (Canzian et al., 1993), meaning distance measures are expected to be useful for classification of galaxy components. Young stars typically form in the dense spiral arms and disperse as

they age forming the age gradient. The use of distance information can help to better understand the existence of age gradients within UGC2885.

Tracing of spiral arms has been used in galaxies with prominent dust lanes along their spiral arms (Shabani et al., 2018). However, galaxy UGC 2885 has some ambiguity in the spiral arm structure due to a lack of observable dust lanes in the optical HST imagery and the galaxy's 74 inclination. Therefore, we define the spiral arms by fitting logarithmic spirals, which are good approximations of the shape of spiral arms (Davis et al., 2012; Seigar & James, 1998). To ensure our tracing of spiral arms follows logarithmic structure, we perform a piece-wise fit of the spiral arms by manually overlaying logarithmic spirals onto the HST imagery and selecting segments. This is done in Desmos, a free online graphing program. When we identify a sufficient piece-wise fit, we plot points along the spiral arms, connecting the points using a parametric curve.

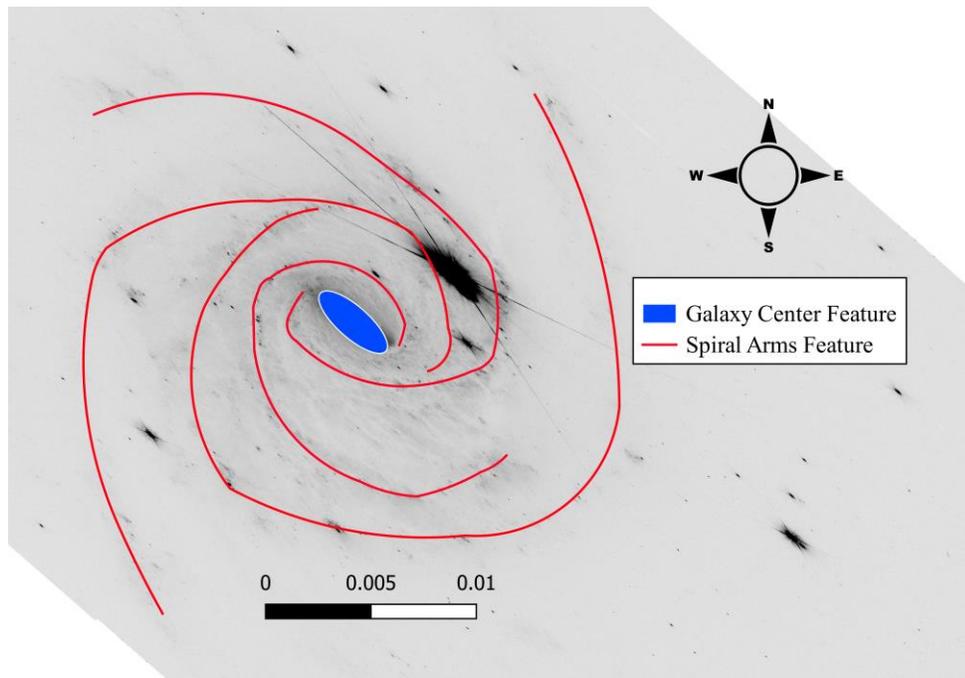
Many methods of spiral arm identification exist and there is presently no method that is agreed upon to be the most accurate. Arm identification is affected by the galaxy inclination adopted as well as the wavelength used to trace the arms (Davis et al., 2012). Dust lane identification is sometimes used: according to the spiral density wave theory (Lin & Shu, 1964), dust lanes are formed from the compression of gas by the spiral density waves, so are closely correlated with the true position of the spiral arms. However, the inclination of UGC2885 makes dust lanes difficult to identify within the HST imagery. Piece-wise fitting of spiral arms in particular involves some uncertainty, since it relies on visual estimates.

Uncertainty in the spiral arm positions affects the computation of distance to the arm centre and will thus affect classifications that uses the distance to arm center as a parameter. We suspect that the magnitude of this effect is small; a detailed quantification is left to further work.

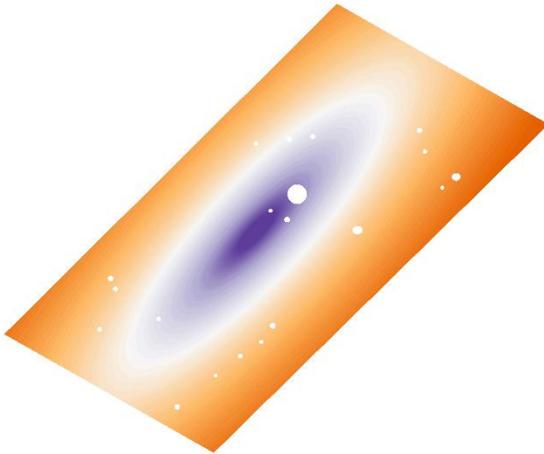
The tool used to compute Euclidean distance converts the vector spiral arm line features to raster by generating pixels along the spiral arm lines. Distance is therefore calculated for each cell in a specified extent to the closest raster pixel in the spiral arm line. Following distance calculation, the distance raster is clipped to match the extent of the HST imagery and foreground stars are masked. The interarm regions, those between the spiral arms, will

have distance calculated from the nearest spiral arm by way of the shortest distance. We also draw a polygon over the galaxy center in the deprojected HST imagery. The galaxy center feature is also converted to raster pixels, and Euclidean distance is calculated for each pixel to the nearest edge pixel in the galaxy center raster by way of the shortest distance. We end up with two distance layers in decimal degrees and multiply both rasters by 3600 to convert to distance in arcseconds. This step brings the raster values closer to the 32-bit range of values of the HST bands that we will input into classification along with the distance bands. We then reproject the distance layers to the original extent of the HST imagery. The spiral arm and galaxy center features as well as their respective distance layers are shown in Figure 2-6.

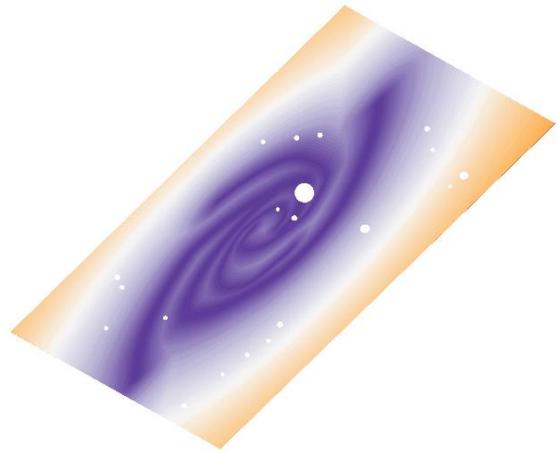
(a) Spiral arm and galaxy center features



(b) Galaxy center distance raster



(c) Spiral arm distance raster



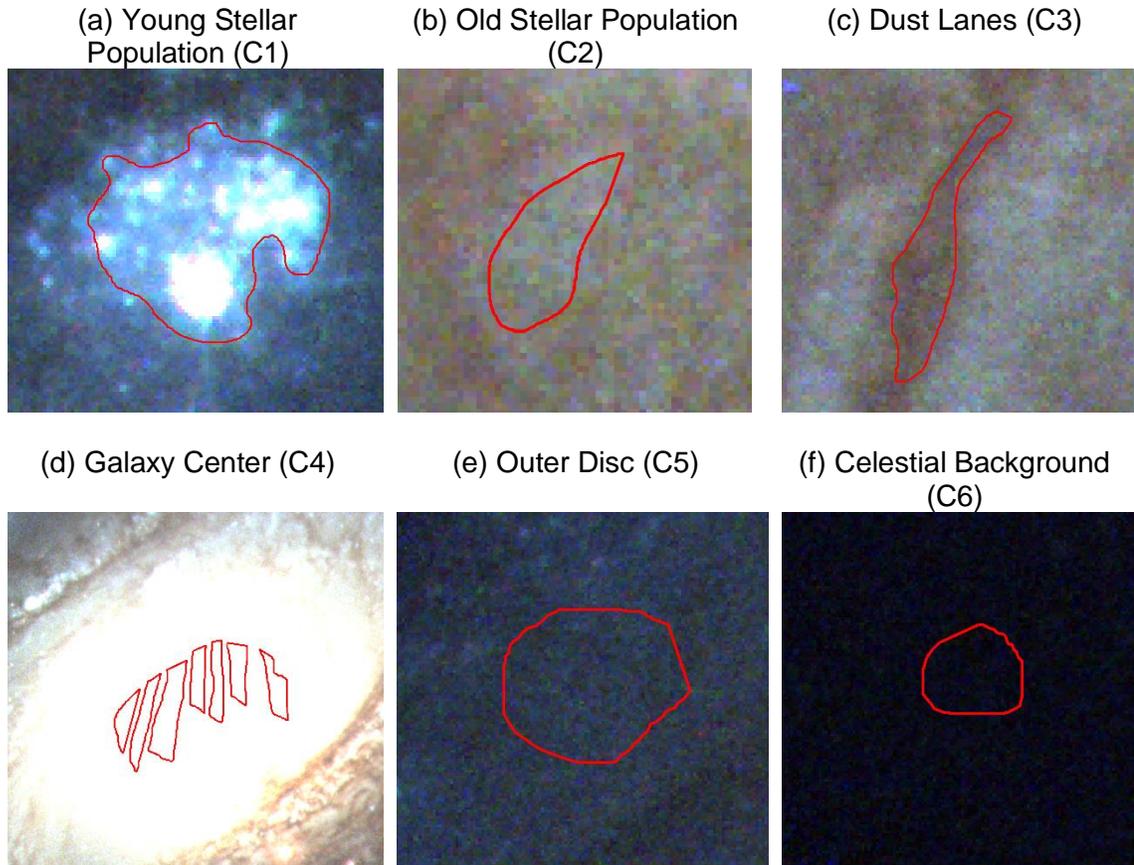
**Figure 2-6: Features used to calculate Euclidean distance and the corresponding Euclidean distance rasters: a) spiral arm and galaxy center features; b) galaxy center distance raster; c) spiral arm distance raster. Spiral arms and galaxy center features shown in red and blue respectively. The background image shows the deprojected UGC 2885 F475W blue-green band (Holwerda et al., 2020). Both distance rasters are reprojected to the galaxy's  $74^\circ$  inclination.**

## 2.4 Classification Scheme

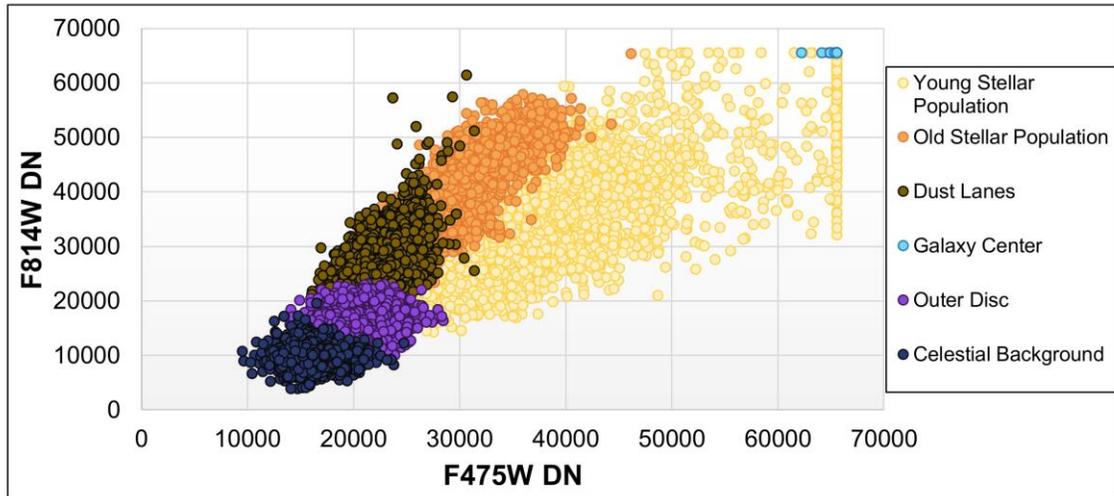
Although UGC 2885 is considered a nearby galaxy, it is not near enough for observation of individual stars. However, we can observe groups of stars called star clusters meaning that there is variation within galactic components we observe in the digital imagery. Because of UGC 2885's vast distance, we also have no access to 'ground truth' like we do for Earth-based phenomena. Therefore, we use our expert classification to train the models based off our theoretical understanding of galaxy components in the HST imagery. Along with HST imagery, we use the distance layers (galaxy center and spiral arm) and the band ratios to create training sites. The distance and band ratios act as complementary information to confirm the visual identification, and help to reduce subjectivity of training site creation.

To improve reproducibility we provide a guide to classification schemes used in this research. Figure 2-7 shows an example of how we defined the classification schemes within a GIS software. We decide on six classes based on their spectral values and visual appearance within the digital imagery: young stellar population (C1), old stellar population (C2), dust lanes (C3), galaxy center (C4), outer disc (C5), and celestial background (C6). Although the celestial background is not a part of the galaxy, we include it to avoid confusion with similar pixels within the galaxy.

We create training site polygons over areas of the digital imagery representing the six classes. Figure 2-8 visually demonstrates class separability between HST blue-green band F475W and near-infrared band F814W using a scatterplot of pixel values from each class. There is some class confusion due to the variation within galactic components in the HST imagery; the most confusion occurs between the old stellar populations and dust lanes that are similar in appearance in the HST imagery. Young and old stellar populations exhibit the widest range in pixel values. Because of saturation in the original HST imagery, the galaxy center and the young stellar population have saturated points concentrated in the top right of the graph.



**Figure 2-7: Examples of training site selection using the classification scheme with six classes: a) young stellar population; b) old stellar population; c) dust lanes; d) galaxy center; e) outer disc; f) celestial background. Background image is a RGB colour composite of HST band data of UGC 2885: F475W band is displayed in blue, the F606W band is displayed in green, and the F814W band is displayed in red.**



**Figure 2-8: Scatterplot of classification scheme pixels and their respective digital number (DN) values within the blue-green and near-infrared HST bands.**

## 2.5 Machine Learning Algorithms

To test the usefulness of machine learning for galactic component identification, we compare the performance of traditional Maximum Likelihood Classifier (MLC) model to the more powerful and increasingly popular Random Forest (RF) and Support Vector Machine (SVM) models. All three models are commonly used for image classification (Baron, 2019; Fluke & Jacobs, 2020; Lavallin & Downs, 2021; Maxwell et al., 2018).

### 2.5.1 Maximum Likelihood Classifier

The MLC model is capable of classifying pixels in an image into probability density functions based on their variance and covariance statistics (Foody et al., 1992; Norovsuren et al., 2019). In our case, MLC assigns each pixel in the imagery to one of the six classes specified. We performed MLC in a GIS. We can define MLC by the following equation (Richards & Jia, 1999):

$$D = \ln(a_c) - [0.5\ln(|COV_c|)] - [0.5(X - M_c)T(COV_c - 1)(X - M_c)] \quad (10)$$

where  $D$  is the likelihood,  $c$  is the particular class in question,  $COV_c$  is the covariance matrix for the class  $c$  pixels,  $X$  is a measurement vector for a specific pixel, and  $M_c$  is a mean vector of a class ( $c$ ).

## 2.5.2 Support Vector Machine

The second method of machine learning classification is SVM, a method that uses a hyperplane to define an optimal split between classes. An optimal split can be defined as one that separates the natural groupings in the samples while maintaining the maximal distance from support vectors, which are extreme samples within the data (Fluke & Jacobs, 2020). When training the SVM models, we use the value of 500 samples per class. We test SVM using several numbers of samples including 125, 250, 500, 750, and 1000. We find that the default value of 500 samples per class is sufficient, as there is no drastic change in accuracy when testing with the other numbers of samples. Shao and Lunetta (2012) find that SVM performs well with a low sample size of 20 pixels, although classification accuracy did increase when testing up to 800 pixels per class. SVM is not as sensitive to training sample sizes as RF (Thanh Noi & Kappas, 2017) so it makes sense the accuracy does not drastically change. We evaluate the accuracy of training sample sizes using overall accuracy statistic, which is a measure of the sum of the individual class accuracy (correctly classified pixels in each class) divided by the total number of pixels in the testing data.

## 2.5.3 Random Forest

The RF model is an ensemble algorithm that relies on a set of decision trees that each make a decision about the state of a sample. After processing a sample through its decision trees, the class or state of that sample is decided through a majority vote of the trees meaning that the class or state most commonly identified by the decision trees is assigned to the sample (Breiman, 1984; Breiman, 2001).

We performed RF in a GIS. From testing of models with 125, 250, 500, 750, and 1000 trees, we find that there is no notable difference in accuracy, leading us to use the 500 trees as recommended by Belgiu & Drăguț, (2016). Lawrence et al. (2006) note that the use of 500 trees allows the model to stabilize errors before all trees are processed. We test tree depths of 5, 15, 30, 80, and 100. Using tree depths of 5, 80, and 100 resulted in lower model accuracy whereas a maximum depth of 15 and 30 trees results in the highest accuracy; therefore, we choose 30 trees. We also use 1000 samples per class to ensure that a sufficient number of pixels are included in training. Because there is some variation within galactic

components, this is particularly important to ensure we are able to train the model on all interclass differences.

To analyze the importance of the parameters, we run the RF algorithm in R Studio (Version 1.1.463; RStudio Team, 2020) using all 38 layers. The GIS program used for MLC, SVM, and RF classifications is not capable of processing a mean decrease in Gini coefficient (MDG) importance plot. Therefore, we make use of the “randomForest” package in R programming language to produce a MDG plot ranking the importance of all layers using the “varImpPlot()” function. The higher the mean decrease in Gini, the more important the parameter is for classification (Koo et al., 2021). MDG also identifies natural subgroups from analysis of all parameters.

#### 2.5.4 Classification Groups

From the analysis of the 38 parameters created, we identify the following groups of classifications:

- (1) HST bands
- (2) Most important textures
- (3) Less important textures
- (4) All eight Haralick textures
- (5) Distance and HST bands
- (6) Band flux ratios
- (7) Fraction band ratios
- (8) Spectral slope ratios
- (9) Top important MDG layers

We perform each classification three times: once for MLC, once for RF, and once for SVM. As a baseline comparison, we classify the original HST BVI imagery. Because there are

eight textures in total, we choose to classify the ones identified as most important by the MDG plot. Further, we also classify less important textures to see how the accuracy changes. Along with these, we classify with all eight Haralick textures (Haralick et al., 1973) to determine whether more or less textural information is useful. We also test classification of flux, fraction, and spectral slope band ratios to determine how band ratios contribute to classification of galaxy components. Finally, we classify the most important layers within the top subgroups as identified by the RF MDG plot.

## 2.6 Accuracy Assessment

To assess the accuracy of our model, we split the polygon sites described in Section 2.4 into a training set and test set. After testing of 90/10, 80/20, and 70/30 splits, we find that there is no notable difference in accuracy; therefore, we choose to use a 70/30 split representing 70% and 30% of the total number of polygons. The number of polygons and pixels per data set is shown in Table 2-3. To increase the confidence of our model, we perform classification twice, alternating the sets used for training and testing and averaging the two accuracies.

**Table 2-3: Number of polygons and pixels within 70% and 30% of the polygon dataset.**

	<b>70% of Dataset</b>	<b>30% of Dataset</b>
<b>Number of Polygons</b>	455	195
<b>Number of Pixels</b>	170784	120605

To analyze the prediction power of the MLC, RF, and SVM models, we calculate overall accuracy (OA), user's accuracy (UA), producer's accuracy (PA), and F1 score. OA is defined by the below equation:

$$OA = \text{correctly classified pixels} / \text{total number of pixels in the test set} \quad (11)$$

Overall accuracy is a commonly used statistic in remote sensing map analysis and is useful as a simple measure of the proportion of correctly classified pixels in a map (Shao et al., 2019).

The second accuracy statistic UA is described by the equation:

$$UA = TP / TP + FP \quad (12)$$

where TP is the true positive and FP is the false positive in the confusion matrix. UA represents the class accuracy for the rows of the confusion matrix. The UA metric is useful for measuring the errors of commission (Congalton & Green, 2019). Similarly, PA defines the class accuracy for the columns of the confusion matrix and is represented by Eq. (14):

$$PA = TP / TP + FN \quad (13)$$

The PA metric looks at the errors of omission making it useful for knowing what samples have been omitted from being correctly classified (Congalton & Green, 2019). Both the UA and PA metrics take into account confusion matrix error. Because OA does not account for error, reporting on the UA and PA values is ideal for ensuring the confusion matrix is summarized properly.

The fourth and final accuracy metric, the F1 Score, analyzes both the user's and producer's accuracy statistics of a confusion matrix. The F1 Score for an individual class is calculated as the mean of user's and producer's accuracy by the equation:

$$F1 = 2 * (UA * PA) / (UA + PA) \quad (14)$$

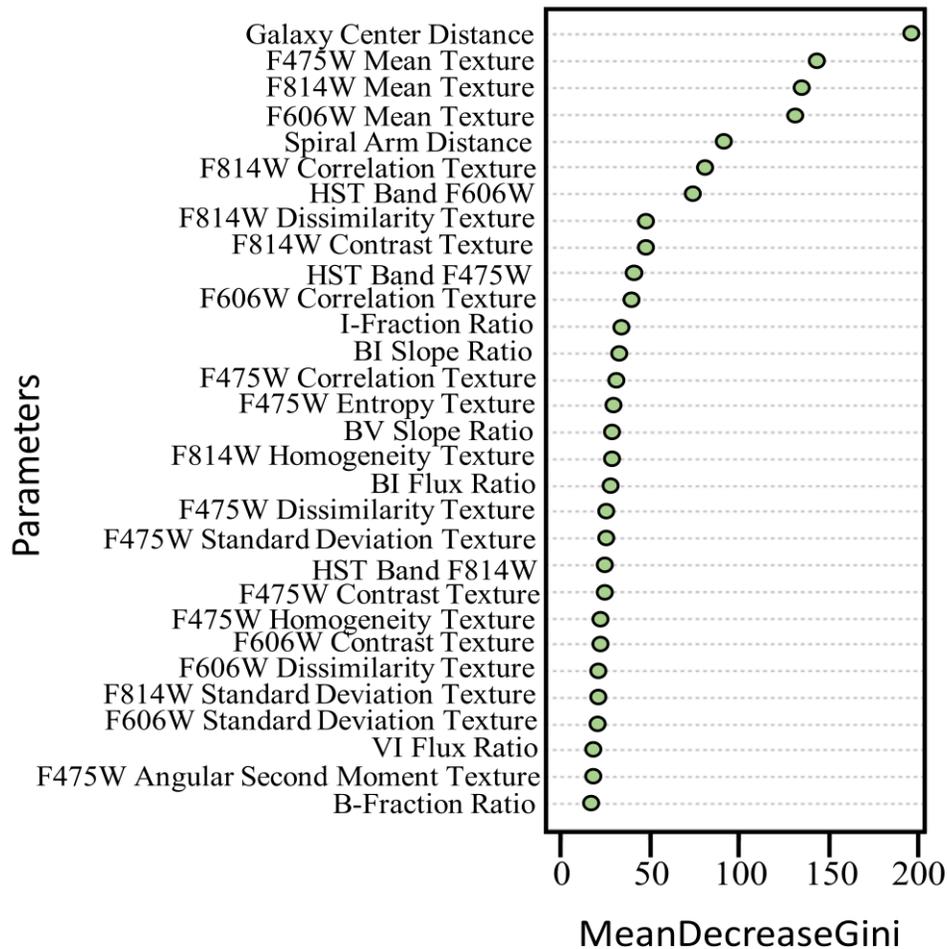
We also evaluate the F1 Score for the overall map, so we average the individual class F1 Scores (Goutte & Gaussier, 2005).

## 2.7 Results

### 2.7.1 Parameter Importance

The Mean Decrease Gini (MDG) plot from RF classification in Figure 2-9 identifies several groups of importance from the input parameters. Overall, the galaxy center distance is the

most important of the 38 total parameters and forms its own group. We also find that Mean texture parameters are the most important textures and form their own subgroup. The third group contains the spiral arm distance, band F814W (infrared) Correlation texture, and the HST band F606W (visual). We perform classification of these three groups (top seven MDG layers). Because both distance parameters and all three Mean texture parameters are within the top three subgroups, we conclude that these are the most useful parameters for galaxy component classification. Angular Second Moment is the least important texture according to the MDG plot. Therefore, we also perform classifications of Mean textures and Angular Second Moment textures as the most and least important textures respectively.

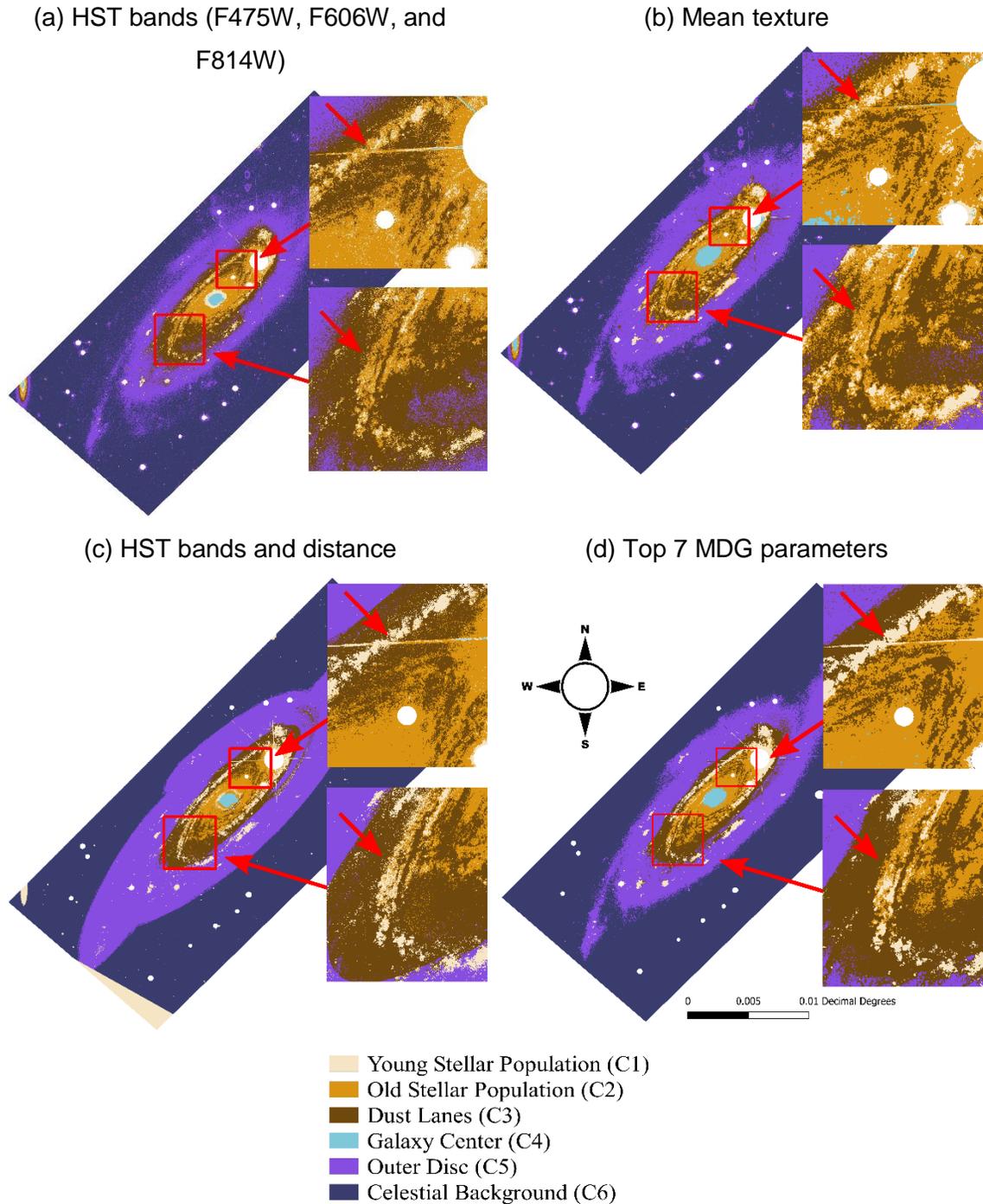


**Figure 2-9: MDG plot of the top 30 important input parameters.**

Figure 2-10 shows maps of classifications with the top performing layers from the MDG plot as compared to the classification with HST bands, which we use as a baseline. Because SVM model performs slightly better than the MLC and RF models, we compare SVM classifications when using 70% of the polygon dataset for training. We do not show the classifications using 30% of the polygon dataset for training as there are no noteworthy differences. Classifications with distance layers – HST bands and distance, and top seven MDG layers – results in the highest accuracies, but also exhibits sharp edges throughout the image. On the contrary, classification with mean textures and classification with HST bands exhibit smoother changes in galaxy component membership throughout the map. This contrast is due to the nature of the Euclidean distance rasters where the distance is discrete rather than a continuous surface. However, classification with the top seven layers incorporates both texture and distance along with HST band F606W, reducing effects of the distance layers. The main pitfalls of the use of distance in classification are the lack of pixels classified as old stellar population and the overemphasis of the dust lanes and outer disc. Use of mean texture appears to improve upon classification with HST bands as it does a better job of identifying old stellar populations within the digital imagery. Classification with only the HST bands tends to overemphasize the dust lanes. Similarly, young stellar populations are better defined within the mean texture classification.

Comparison of all maps reveals that classification of mean texture and classification of the top seven MDG parameters are better at classifying the old stellar populations within the inner disc of the galaxy. However, in the classification of the top seven MDG parameters, the discrete effects of the distance parameters are stronger in the outer disc region as shown by the arrow in the upper inset image (Figure 2-10). There is a sudden change in galaxy component membership going from young stellar population to dust lanes to outer disc. In the original HST BVI imagery, this is not the case and that there is more diversity of galaxy components within this same region, similar in appearance to the classification with mean textures. The same effect can be seen in the lower inset map where mean texture has the most diversity of galaxy components and therefore best resembles the original HST BVI imagery. However, the classification with HST bands, HST bands and distance, and the top seven MDG parameters does not show the same diversity of galaxy components within the region identified by the arrow in the lower inset map (Figure 2-10). Therefore, although

we find that the classification with top seven MDG parameters results in the highest accuracy, the classification of mean textures is best at classifying the small-scale changes in galaxy component membership throughout the digital imagery.



**Figure 2-10: SVM classifications using 70% of the polygon set as training and 30% of the polygons as testing: a) HST bands (F475W, F606W, and F814W); b) Mean texture; c) HST bands and distance; d) Top 7 MDG parameters. We include inset maps, represented by red squares, to show the small details within the inner disc of the galaxy.**

We include the confusion matrices of the most accurate classification of SVM with the top seven MDG layers in Table 2-4 and Table 2-5. A confusion matrix compares the pixels in each class within the test set and the classification results of the pixels within the classified imagery. The diagonal represents the correctly classified pixels for their respective classes and is coloured in green. The column and row coloured in orange represents the UA and PA respectively. We colour the OA statistic in blue. The row and column titled ‘Total’ represent the total number of pixels in each class from the original reference data (test set) and the classified image respectively. Misclassifications are the boxes not along the diagonal, excluding the Total, UA, and PA rows and columns as well as the OA. We bold the boxes that represent the most confusion, meaning that these classes have higher rates of misclassification between them in the form of misclassified pixels.

Table 2-4 shows SVM classification of the top seven layers when using 70% of the polygon set for training while Table 2-5 shows SVM classification of the top seven layers when using 30% of the polygon set for training. From Table 2-4, we notice that there is the most confusion between classes C2 (old stellar population) and C3 (dust lanes). This agrees with our expectations as discussed in Section 2.4. Similarly, class C3 (dust lanes) and C5 (outer disc) share confusion due to their spectral similarities. Within the inner disc, it is easier to distinguish between the dust lanes and other galaxy components. However, in the outer disc where there are less galaxy components, the HST imagery is darker making it more difficult to visually separate the components within it. Class C1 (young stellar population) and C2 (old stellar population) share a bit of confusion within both tables. However, C1 performs better due to its bluer appearance. The least confusion is present in the C4 (galaxy center) and C6 (celestial background) classes as these are the most spectrally unique parts of the HST imagery.

**Table 2-4: Confusion matrix of SVM classification of the top seven layer using 70% of the polygon set for training and 30% for testing. The green shading shows the number of correctly classified pixels for each class, orange shading shows the PA and UA row and column, and the blue shading shows the OA statistic. Bolded numbers identify the areas of the most confusion between classes.**

		<i>Reference data by expert interpretation</i>							
		C1	C2	C3	C4	C5	C6	Total	UA
<i>ML Classification</i>	C1	9856	55	3	0	9	0	9923	0.99
	C2	185	9579	<b>696</b>	0	0	0	10460	0.92
	C3	46	<b>875</b>	10270	0	<b>800</b>	0	11991	0.86
	C4	52	113	0	6279	0	0	6444	0.97
	C5	0	0	<b>578</b>	0	49622	0	50200	0.99
	C6	0	0	0	0	9	31578	31587	1.00
	Total	10139	10622	11547	6279	50440	31578	120605	
	PA	0.97	0.90	0.89	1.00	0.98	1.00		0.96

**Table 2-5: Confusion matrix of SVM classification of the top seven layer using 30% of the polygon set for training and 70% for testing. The green shading shows the number of correctly classified pixels for each class, orange shading shows the PA and UA row and column, and the blue shading shows the OA statistic. Bolded numbers identify the areas of the most confusion between classes.**

		<i>Reference data by expert interpretation</i>							
		C1	C2	C3	C4	C5	C6	Total	UA
<i>ML Classification</i>	C1	27696	222	46	0	6	0	27970	0.99
	C2	<b>653</b>	16788	<b>2073</b>	0	0	0	19514	0.86
	C3	90	<b>2417</b>	27229	0	<b>4707</b>	0	34443	0.79
	C4	0	0	0	9105	0	0	9105	1.00
	C5	0	0	531	0	36281	0	36812	0.98
	C6	0	0	0	0	113	42827	42940	1.00
	Total	28439	19427	29879	9105	41107	42827	170784	
	PA	0.97	0.86	0.91	1.00	0.88	1.00		0.94

## 2.7.2 Galaxy Component Classification Performance

We present the user's accuracy (UA), producer's accuracy (PA), and F1 Scores in Table 2-6, Table 2-7, and Table 2-8. The ML models are best at predicting galaxy center and celestial background as they exhibit lower rates of confusion than the remaining classes

(Figure 2-8); this is due to the brightness of the galaxy center and the darkness of the celestial background. Class C5, the outer disc, also exhibits high prediction power. According to Table 2-6, 2-7, and 2-8, the old stellar population (C2) and dust lane (C3) classes have the lowest classification accuracy due to the spectral similarities between these classes. For instance, Figure 2-8 (Section 2.4) demonstrates the confusion between these three classes due to the visual similarities and mixing within the inner disc of the galaxy due to the limited resolution of the HST imagery. The young stellar population (C1) has slightly higher UA and PA values as this class is more spectrally unique than both the old stellar population and dust lanes; by spectrally unique, we refer to the brightness and blueness of the young stars within the galaxy. In contrast, the old stars and dust lanes tend to be redder and dimmer than the young stars.

Overall, the MLC model best classified the outer disc (C5) and celestial background (C6) classes for each accuracy statistics (Table 2-6, 2-7, 2-8). MLC has low classification accuracy for the old stellar population class (C2). In comparison to the RF and SVM models, MLC class accuracies are comparative to the RF class accuracies more so than to the SVM class accuracies. For instance, in Table 2-6, MLC and RF have the same average UA of 0.80 for young stellar population (C1) classification. For producer's accuracy in Table 2-7, MLC model outperforms RF accuracy of classification of dust lanes (C3) and outer disc (C5). The SVM model consistently outperforms individual class accuracies of the MLC model.

**Table 2-6: Average of user’s accuracy for each class in the MLC, RF, and SVM models. The bolded numbers represent the model with the highest user’s accuracy for each class. SVM model outperforms the MLC and RF models.**

Model	User’s Accuracy Average					
	C1	C2	C3	C4	C5	C6
MLC	0.80	0.64	0.68	0.69	0.89	0.88
RF	0.80	0.70	0.65	<b>0.87</b>	0.91	0.93
SVM	<b>0.84</b>	<b>0.77</b>	<b>0.73</b>	0.83	<b>0.92</b>	<b>0.94</b>

**Table 2-7: Average of producer’s accuracy for each class in the MLC, RF, and SVM models.3 The bolded numbers represent the model with the highest producer’s accuracy for each class. SVM and RF models outperform the MLC.**

Model	Producer’s Accuracy Average					
	C1	C2	C3	C4	C5	C6
MLC	0.75	0.70	0.76	0.83	0.79	0.89
RF	<b>0.82</b>	0.71	0.70	<b>1.00</b>	0.77	<b>0.96</b>
SVM	0.80	<b>0.72</b>	<b>0.78</b>	<b>1.00</b>	<b>0.81</b>	<b>0.96</b>

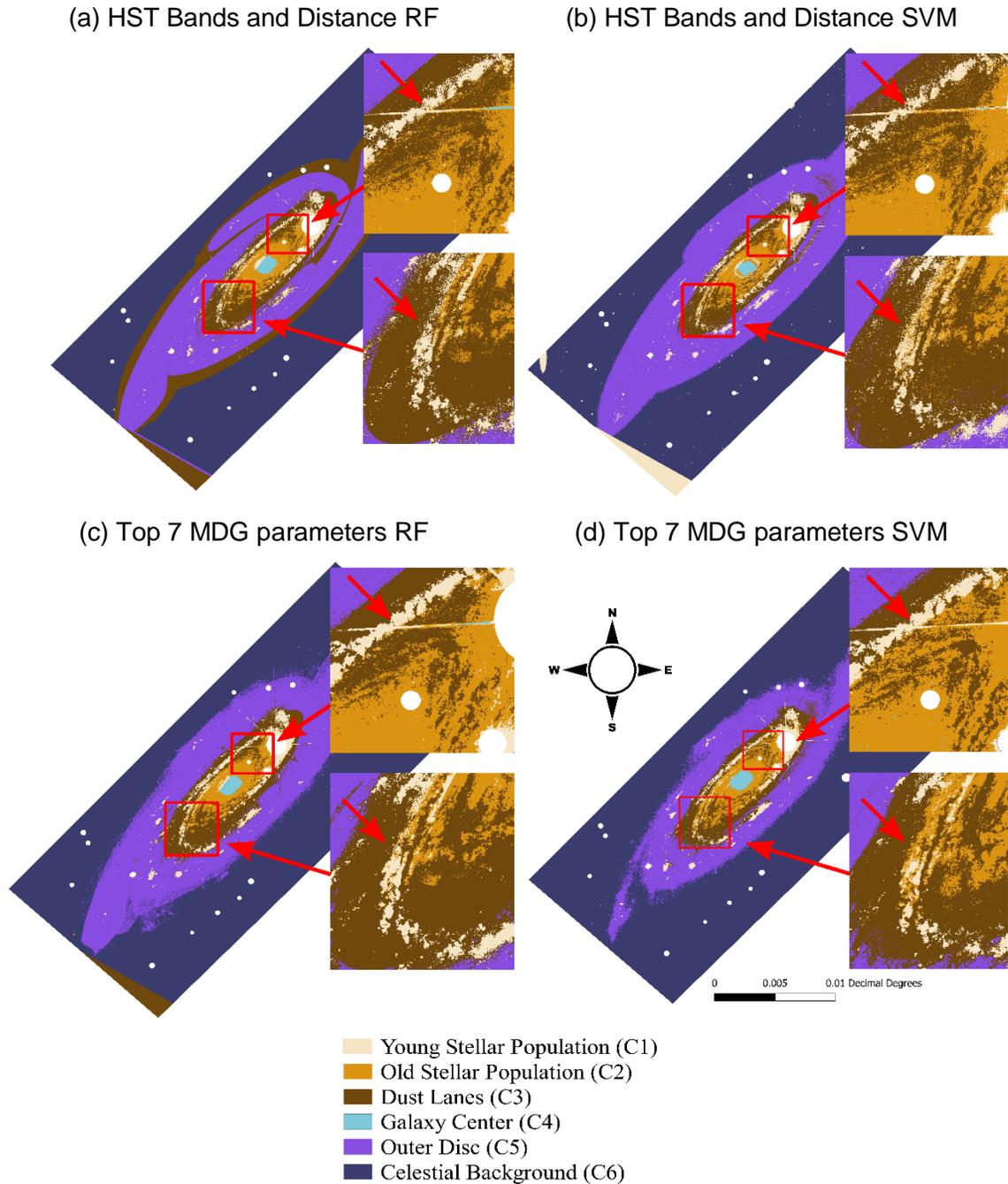
**Table 2-8: Average of F1 Score for each class in the MLC, RF, and SVM models.3 The bolded numbers represent the model with the highest F1 score accuracy for each class. SVM model outperforms the MLC and RF models.**

Model	F1 Score Average					
	C1	C2	C3	C4	C5	C6
MLC	0.77	0.65	0.71	0.75	0.82	0.89
RF	0.79	0.68	0.67	<b>0.93</b>	0.81	<b>0.95</b>
SVM	<b>0.81</b>	<b>0.72</b>	<b>0.73</b>	0.90	<b>0.85</b>	<b>0.95</b>

Figure 2-11 shows maps of the top four classifications as determined by average of overall accuracy statistics and F1 Scores (Table 2-9). For the highest accuracy classifications reported in Figure 2-11 and Table 2-9, MLC performs with a similar accuracy to the RF and SVM models, but still falls short by about 2% for both OA and F1 score accuracies. From observation of the maps, we note that RF and SVM classifications appear quite similar. The main difference between the resulting maps is that the HST bands and distance classifications create sharper transitions as shown by the arrows in the upper and lower inset maps. The SVM classifications show more small details within the galaxy. Therefore,

the RF model tends to emphasize the distance layers the most in both classification types shown in Figure 2-11. We find that both RF and SVM are useful methods of classification for digital imagery of galaxies.

Visually, the best models in Figure 2-11 are those using the top seven MDG parameters. This agrees with the OA and F1 Score as the models using top seven layers result in slightly higher (1%) accuracies than the HST band and distance models. The galaxy center is best identified by the SVM classification with top seven MDG parameters; in the other three classifications in Figure 2-11, there is some visual confusion between the class membership of the galaxy center and young stellar population. SVM classification with top seven MDG layers is also best at identifying the old stellar population within the inner disc of the galaxy, although confusion with dust lanes remains an issue. Ultimately, the arrows indicated inside the inset maps demonstrate that there is not much difference between the classifications of HST bands and distance and of the top seven MDG parameters. The accuracies of each model, as reported in Table 2-9, agree with this finding. 3.3. Model performance summary.



**Figure 2-11: Top performing classifications: a) HST bands and distance RF; b) HST bands and distance SVM; c) Top 7 MDG parameters RF; d) Top 7 MDG parameters SVM. All maps shown are the product of classifications using 70% of the polygon set as training and 30% as testing. We include inset maps, represented by red squares, to show the fine details within the inner disc of the galaxy.**

**Table 2-9: Top performing classifications from Figure 2-11 and their corresponding accuracy statistics.**

	<b>HST Bands and Distance RF</b>	<b>HST Bands and Distance SVM</b>	<b>Top 7 MDG parameters RF</b>	<b>Top 7 MDG parameters SVM</b>
<b>OA</b>	94.7%	94.5%	95.4%	95.4%
<b>F1 Score</b>	93.5%	93.3%	94.4%	94.6%

### 2.7.3 Model Performance Summary

We perform nine classifications for each ML model: HST bands, mean texture, angular second moment texture, all textures, distance and HST bands, band flux ratios, fraction band ratios, spectral slope ratios, and top seven MDG layers. Overall, the RF and SVM models result in higher performance than the traditional MLC model with SVM being slightly more successful at predicting galactic components over the nine classifications performed (Table 2-10). The MLC model performs particularly well with the spectral slope band ratio classification with an average OA of 85.3%. For the same spectral slope classification, RF and SVM models perform with an average OA of 48.1% and 66.7% respectively. Otherwise, RF and SVM consistently outperform the MLC model. Therefore, we suggest that RF and SVM models be used for galaxy component classification.

**Table 2-10: Average accuracies for each model.**

<b>Model</b>	<b>Overall accuracy</b>	<b>F1 Score</b>
<b>MLC</b>	80.5%	80.4%
<b>RF</b>	82.6%	80.9%
<b>SVM</b>	<b>84.9%</b>	<b>82.6%</b>

## 2.8 Discussion

In this study, we use MLC, RF, and SVM machine learning models to classify galaxy component membership within HST digital imagery of UGC 2885. Along with HST imagery, we input band ratios derived from HST imagery, textural features derived from

HST imagery, and distance layers into the classification to determine the most effective method of class membership prediction. We identify six classes within the digital imagery: young stellar population. To analyze accuracy of galaxy component membership, we make use of PA and UA statistics. The classes with the most confusion are young stellar population and old stellar population as well as old stellar population and dust lanes as they share similar spectral appearance. The classes with the least confusion are galaxy center, outer disc, and celestial background. To analyze accuracy of galaxy component membership, we make use of PA and UA statistics. The classes with the most confusion are young stellar population and old stellar population as well as old stellar population and dust lanes as they share similar spectral appearance. The classes with the least confusion are galaxy center, outer disc, and celestial background. This is expected as these classes are unique in spectral appearance. Among the MLC, RF, and SVM machine learning models used, SVM results in the highest accuracy of galaxy component classification between both PA and UA statistics. SVM also results in the highest accuracy confirmed by the high accuracies that classification with distance and textures yields. Finally, a combination of HST bands, texture, and distance results in the highest accuracy. The combined power of several types of information is optimal for galaxy component classification within digital imagery. The success of the SVM and RF models and relatively poor performance of MLC is expected and agrees with results from other recent studies comparing machine learning algorithms in both remote sensing and astronomy (Ghayour et al., 2021; Wang et al., 2021).

One limitation of our study is the lack of ‘ground truth’ reference data. Due to this lack of reference data, we use expert visual interpretation to train the models. Since UGC 2885 is 71 Mpc away, some galaxy components are difficult to distinguish. Because of this, we are only able to train the model on pixels whose class membership we are certain. This means that large portions of the galaxy are ignored in accuracy assessment, likely contributing to the high accuracy results. To improve the reliability of accuracy assessment, we recommend classifying galaxies that have some sort of reference data available. Further, galaxies in closer proximity may be more suitable due to the higher resolution of imagery, making it easier to identify galaxy components.

One manual step in our study is the tracing of spiral arms. Although this method has been used for several purposes in spiral galaxy research (Scheepmaker et al., 2009; Shabani et al., 2018; Bialopetravičius & Narbutis, 2020), it would be infeasible to manually trace spiral arms of many galaxies if using our model to automate rapid galaxy component classification. Research is being done to automate spiral arm fitting (Davis & Hayes, 2014; Bekki, 2021), so the addition of an automated model for spiral arm fitting into our machine learning model would drastically decrease processing time. Additionally, deprojection of galaxy imagery is unnecessary for spiral arm fitting models due to the low dependence on disc inclination angle (Bekki, 2021). Due to the manual tracing of the spiral arms, there may be some uncertainties present in the final classification of the galaxy components. Future work can contribute to improving the uncertainty of the distance calculations. For instance, the development of methods that increase accuracy of identification of the true spiral arm center.

For future study of galactic components, we find that use of textures may improve classification accuracy. We recommend that texture analysis be experimented with further to explore its full potential for astronomical research. We recommend that texture also be tested on imagery of other deep space celestial phenomena such as irregular galaxies that exhibit tonal variation and patterns within digital imagery. One limitation of texture is that it is not as effective when applied to low resolution imagery. We recommend texture be used for high resolution HST imagery or for nearby celestial phenomena. For instance, textural features would be useful for upcoming telescopes such as the Euclid and Roman Space Telescope, as well as the operating James Webb Space Telescope that produce high resolution imagery comparable to that of HST (Gardner et al., 2006; Laureijs et al., 2011; Spergel et al., 2015). Along with texture, we also recommend the use of distance measures for classification of galaxy components, in conjunction with other parameters such as texture and HST bands to achieve the optimal accuracy.

The machine learning models used present their own set of limitations. One downfall of the SVM algorithm is that it ignores training samples that do not support the hyperplane (Foody & Mathur, 2006). This might cause classes with a wide range of pixel values to be classified poorly. Presence of extraneous features may negatively affect the performance

of the model. We recommend that the features be analyzed to identify any outlier features to avoid a significant decrease in accuracy of the model (Baron, 2019). MLC is not able to handle data with a non-normal distribution as it attempts to define a unique probability density function for each class (Otukey & Blaschke, 2010). To combat this, data with normal distribution should be used for MLC classification. The HST data used here has normally distributed flux so is compatible with the MLC algorithm. A disadvantage to the RF algorithm is that it cannot handle datasets with imbalanced training samples (Dalponte et al., 2013). Astronomical data has noise present and the ML methods used here fail to account for that. ML algorithms that account for uncertainties of both features and labels have only recently been developed (e.g., Reis et al., 2019). To take astronomical noise into account, probability distribution functions are created for the features and labels. This improves accuracy of RF classification by 10–30% (Reis et al., 2019). Currently, RF models in widely-used packages and GIS programs are not capable of accounting for uncertainty in input data. Incorporating ML algorithms that account for uncertainty would be a natural extension of our model and would improve accuracy of classification.

Our ML model is successful at classifying galaxy components within a nearby spiral galaxy, UGC 2885. Dissecting these fine structural details within galaxies is important for understanding formation and evolution of galaxies (Lingard et al., 2020; Peng et al., 2002). UGC 2885 exhibits an extended disc with a sparse population of young stars that we classify using visible HST imagery. We also notice that a large portion of the stellar matter we can see is located in the inner disc. Further research on galaxies of different spiral forms and life cycle stages is needed to fully understand the secular spatial and temporal changes of galaxy component distribution. Classification using ours or similar models helps to automate that process.

## 2.9 Conclusions

In this study, we proposed a machine learning (ML) approach for galaxy mapping of UGC 2885 using high-resolution Hubble Space Telescope (HST) digital imagery. We compare three ML models: maximum likelihood classifier (MLC), random forest (RF), and Support Vector Machine (SVM). ML is successful at mapping galaxy components: RF and SVM models are found to have the strongest classification power whereas MLC performance is

slightly inferior. The ML models successfully classify all identified components within the digital imagery, with the most confusion shared between the dust lanes and old stellar populations within the galaxy. The young stellar population, galaxy center, outer disc, and celestial background are the best classified by the ML models and therefore have the least confusion. From analysis of parameter importance, distance and mean textural parameters are the most important for galaxy component classification. The best performing models were those using the top seven mean decrease Gini parameters, a combination of distance, textural features derived from HST imagery, and HST digital imagery data, making this method particularly important. Further research could determine the full potential of textural analysis for study of galaxies and other celestial phenomena. These findings are relevant for the soon to be launched Euclid and Roman Space Telescope, as well as the operational James Webb Space Telescope, as these telescopes will provide an abundance of high-resolution data similar to the HST data used in this study. Our research demonstrates that the automation of mapping the fine galaxy component structures within digital imagery is feasible.

## 2.10 References

- Au, K. (2006). *Inferring galaxy morphology through texture analysis* (Publication No. AAT3216277) [Doctoral dissertation, Carnegie Mellon University]. ProQuest Dissertations And Theses.
- Baron, D. (2019). Machine learning in astronomy: A practical overview. ArXiv:1904.07248 [Astro-Ph]. <http://arxiv.org/abs/1904.07248>
- Bekki, K. (2021). Quantifying the fine structures of disk galaxies with deep learning: Segmentation of spiral arms in different Hubble types. *Astronomy & Astrophysics*, 647, A120. <https://doi.org/10.1051/0004-6361/202039797>
- Belgiu, M., & Drăguț, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, 114, 24–31. <https://doi.org/10.1016/j.isprsjprs.2016.01.011>
- Bialopetravičius, J., & Narbutis, D. (2020). Study of star clusters in the M83 galaxy with a convolutional neural network. *The Astronomical Journal*, 160(6), 264. <https://doi.org/10.3847/1538-3881/abbf53>
- Bianchi, L., Kang, Y., Hodge, P., Dalcanton, J., & Williams, B. (2014). The role of ultraviolet imaging in studies of resolved and unresolved young stellar populations. M31 and M33. *Advances in Space Research*, 53(6), 928–938. <https://doi.org/10.1016/j.asr.2013.08.024>
- Breiman, L. (1984). *Classification and regression trees*. Belmont, Calif: Wadsworth International Group.
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Calzetti, D., Lee, J. C., Sabbi, E., Adamo, A., Smith, L. J., Andrews, J. E., Ubeda, L., Bright, S. N., Thilker, D., Aloisi, A., Brown, T. M., Chandar, R., Christian, C., Cignoni, M., Clayton, G. C., Silva, R. da, Mink, S. E. de, Dobbs, C., Elmegreen,

- B. G., ... Wofford, A. (2015). Legacy extragalactic UV survey (LEGUS) with the *Hubble Space Telescope*. I. Survey description. *The Astronomical Journal*, 149(2), 51. <https://doi.org/10.1088/0004-6256/149/2/51>
- Canzian, B., Allen, R.J., & Tilanus, R.P.J. (1993). Spiral structure of the giant galaxy UGC 2885 – H-alpha kinematics. *Astrophysical Journal*, 406, 457. <https://doi.org/10.1086/172457>.
- Congalton, R. G., & Green, K. (2019). *Assessing the accuracy of remotely sensed data: Principles and practices* (Third edition). CRC Press.
- Dalponte, M., Orka, H. O., Gobakken, T., Gianelle, D., & Naeset, E. (2013). Tree species classification in boreal forests with hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing*, 51(5), 2632–2645. <https://doi.org/10.1109/TGRS.2012.2216272>
- Davis, B. L., Berrier, J. C., Shields, D. W., Kenefick, J., Kenefick, D., Seigar, M. S., Lacy, C. H. S., & Puerari, I. (2012). Measurement of galactic logarithmic spiral arm pitch angle using two-dimensional fast Fourier transform decomposition. *The Astrophysical Journal Supplement Series*, 199(2), 33. <https://doi.org/10.1088/0067-0049/199/2/33>
- Davis, D.R., & Hayes, W.B. (2014). SpArcFiRe: Scalable automated detection of spiral galaxy arm segments. *Astrophysical Journal*, 790, 87. <https://doi.org/10.1088/0004-637X/790/2/87>.
- Dressel, L. (2021). *Wide Field Camera 3 Instrument Handbook* (Version 13.0). Baltimore: STScI. <<https://hst-docs.stsci.edu/wfc3ihb>>.
- Dressel, L. (2022). *Wide Field Camera 3 Instrument Handbook* (Version 14.0). Baltimore: STScI. <<https://hst-docs.stsci.edu/wfc3ihb>>.
- ESA/ESO/NASA. (2021). *FITS Liberator 3*. <[https://esahubble.org/projects/fits\\_liberator/download\\_v30/](https://esahubble.org/projects/fits_liberator/download_v30/)>.

- Farhadian, R., & Clarke, B. R. (2022). A note on the Helmert transformation. *Communications in Statistics - Theory and Methods*, 51(15), 5258–5264. <https://doi.org/10.1080/03610926.2020.1836223>
- Fluke, C. J., & Jacobs, C. (2020). Surveying the reach and maturity of machine learning and artificial intelligence in astronomy. *WIREs Data Mining and Knowledge Discovery*, 10(2). <https://doi.org/10.1002/widm.1349>
- Foody, G. M., Campbell, N. A., Trodd, N. M., & Wood, T. F. (1992). Derivation and Applications of Probabilistic Measures of Class Membership from the Maximum-Likelihood Classification. *Photogrammetric Engineering & Remote Sensing*, 58(9), 1335–1341.
- Foody, G. M., & Mathur, A. (2006). The use of small training sets containing mixed pixels for accurate hard image classification: Training on mixed spectral responses for classification by a SVM. *Remote Sensing of Environment*, 103(2), 179–189. <https://doi.org/10.1016/j.rse.2006.04.001>
- Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., Brown, A. G. A., Vallenari, A., Babusiaux, C., Bailer-Jones, C. A. L., Bastian, U., Biermann, M., Evans, D. W., Eyer, L., Jansen, F., Jordi, C., Klioner, S. A., Lammers, U., Lindegren, L., Luri, X., Mignard, F., Milligan, D. J., ... Zschocke, S. (2016). The Gaia Mission. *Astronomy & Astrophysics*, 595, A1. <https://doi.org/10.1051/0004-6361/201629272>
- Gaia Collaboration, Brown, A. G. A., Vallenari, A., Prusti, T., de Bruijne, J. H. J., Babusiaux, C., Biermann, M., Creevey, O. L., Evans, D. W., Eyer, L., Hutton, A., Jansen, F., Jordi, C., Klioner, S. A., Lammers, U., Lindegren, L., Luri, X., Mignard, F., Panem, C., ... Zwitter, T. (2021). Gaia Early Data Release 3: Summary of the contents and survey properties. *Astronomy & Astrophysics*, 649, A1. <https://doi.org/10.1051/0004-6361/202039657>
- Gardner, J. P., Mather, J. C., Clampin, M., Doyon, R., Greenhouse, M. A., Hammel, H. B., Hutchings, J. B., Jakobsen, P., Lilly, S. J., Long, K. S., Lunine, J. I.,

Mccaughrean, M. J., Mountain, M., Nella, J., Rieke, G. H., Rieke, M. J., Rix, H.-W., Smith, E. P., Sonneborn, G., ... Wright, G. S. (2006). The James Webb Space Telescope. *Space Science Reviews*, 123(4), 485–606.  
<https://doi.org/10.1007/s11214-006-8315-7>

Ghasemian, N., & Akhoondzadeh, M. (2018). Introducing two Random Forest based methods for cloud detection in remote sensing images. *Advances in Space Research*, 62(2), 288–303. <https://doi.org/10.1016/j.asr.2018.04.030>

Ghayour, L., Neshat, A., Paryani, S., Shahabi, H., Shirzadi, A., Chen, W., Al-Ansari, N., Geertsema, M., Pourmehdi Amiri, M., Gholamnia, M., Dou, J., & Ahmad, A. (2021). Performance Evaluation of Sentinel-2 and Landsat 8 OLI Data for Land Cover/Use Classification Using a Comparison between Machine Learning Algorithms. *Remote Sensing*, 13(7), 1349. <https://doi.org/10.3390/rs13071349>

Goutte, C., & Gaussier, E. (2005). A probabilistic interpretation of precision, recall and F-score, with implication for evaluation. In D. E. Losada & J. M. Fernández-Luna (Eds.), *Advances in Information Retrieval*, 3408, 345–359. Springer Berlin Heidelberg. [https://doi.org/10.1007/978-3-540-31865-1\\_25](https://doi.org/10.1007/978-3-540-31865-1_25)

Hall-Beyer, M. (2017). Practical guidelines for choosing GLCM textures to use in landscape classification tasks over a range of moderate spatial scales. *International Journal of Remote Sensing*, 38(5), 1312–1338.  
<https://doi.org/10.1080/01431161.2016.1278314>

Haralick, R. M., Shanmugam, K., & Dinstein, I. (1973). Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-3(6), 610–621. <https://doi.org/10.1109/TSMC.1973.4309314>

Hausen, R., & Robertson, B. (2020). Morpheus: A deep learning framework for pixel-level analysis of astronomical image data. *The Astrophysical Journal Supplement Series*, 248(1), 20. <https://doi.org/10.3847/1538-4365/ab8868>

- Holwerda, B.W. (2017). *The cluster population of UGC 2885*. HST Proposal id.15107. Cycle 25.
- Holwerda, B.W., Chandar, R., Barmby, P., Peeples, M.S., Bailin, J., Ford, K., Gorski, M., Bridge, J., Jacques, A., & Steele, B. (2020). UGC2885 ‘‘Rubin’s Galaxy’’; Hubble observations of a gentle giant spiral galaxy. *Bulletin of the American Astronomical Society*, 52(1), 726.
- Holwerda, B. W., Wu, J. F., Keel, W. C., Young, J., Mullins, R., Hinz, J., Ford, K. E. S., Barmby, P., Chandar, R., Bailin, J., Peek, J., Pickering, T., & Böker, T. (2021). Predicting the spectrum of UGC 2885, Rubin’s Galaxy with machine learning. *The Astrophysical Journal*, 914(2), 142. <https://doi.org/10.3847/1538-4357/abffcc>
- Hunter, D. A., Elmegreen, B. G., Rubin, V. C., Ashburn, A., Wright, T., Józsa, G. I. G., & Struve, C. (2013). Star formation in two luminous spiral galaxies. *The Astronomical Journal*, 146(4), 92. <https://doi.org/10.1088/0004-6256/146/4/92>
- Joye, W. A., & Mandel, E. (2003). New features of SAOImage DS9. *Astronomical Data Analysis Software and Systems XII ASP Conference Series*, 295, 489–492.
- Kalirai, J. (2018). Scientific discovery with the James Webb Space Telescope. *Contemporary Physics*, 59(3), 251–290. <https://doi.org/10.1080/00107514.2018.1467648>
- Kiar, A. K., Barmby, P., & Hidalgo, A. (2017). Deconstructing a galaxy: Colour distributions of point sources in Messier 83. *Monthly Notices of the Royal Astronomical Society*, 472(1), 1074–1087. <https://doi.org/10.1093/mnras/stx2037>
- Koo, B.-K., Baek, J.-W., & Chung, K.-Y. (2021). Weight Feedback-Based Harmonic MDG-Ensemble Model for Prediction of Traffic Accident Severity. *Applied Sciences*, 11(11), 5072. <https://doi.org/10.3390/app11115072>
- Laureijs, R., Amiaux, J., Arduini, S., Auguères, J.-L., Brinchmann, J., Cole, R., Cropper, M., Dabin, C., Duvet, L., Ealet, A., Garilli, B., Gondoin, P., Guzzo, L., Hoar, J., Hoekstra, H., Holmes, R., Kitching, T., Maciaszek, T., Mellier, Y., ... Zucca, E.

- (2011). *Euclid Definition Study Report*. ArXiv:1110.3193 [Astro-Ph].  
<http://arxiv.org/abs/1110.3193>
- Lavallin, A., & Downs, J. A. (2021). Machine learning in geography—Past, present, and future. *Geography Compass*, 15(5). <https://doi.org/10.1111/gec3.12563>
- Lawrence, R. L., Wood, S. D., & Sheley, R. L. (2006). Mapping invasive plants using hyperspectral imagery and Breiman Cutler classifications (randomForest). *Remote Sensing of Environment*, 100(3), 356–362.  
<https://doi.org/10.1016/j.rse.2005.10.014>
- Lee, J. C., Whitmore, B. C., Thilker, D. A., Deger, S., Larson, K. L., Ubeda, L., Anand, G. S., Boquien, M., Chandar, R., Dale, D. A., Emsellem, E., Leroy, A. K., Rosolowsky, E., Schinnerer, E., Schmidt, J., Turner, J., Van Dyk, S., White, R. L., Barnes, A. T., ... Williams, T. G. (2021). *The PHANGS-HST Survey: Physics at High Angular resolution in Nearby Galaxies with the Hubble Space Telescope*. ArXiv:2101.02855 [Astro-Ph]. <http://arxiv.org/abs/2101.02855>
- Lingard, T. K., Masters, K. L., Krawczyk, C., Lintott, C., Kruk, S., Simmons, B., Simpson, R., Bamford, S., Nichol, R. C., & Baeten, E. (2020). Galaxy Zoo Builder: Four-component photometric decomposition of spiral galaxies guided by citizen science. *The Astrophysical Journal*, 900(2), 178.  
<https://doi.org/10.3847/1538-4357/ab9d83>
- Lintott, C. J., Schawinski, K., Slosar, A., Land, K., Bamford, S., Thomas, D., Raddick, M. J., Nichol, R. C., Szalay, A., Andreescu, D., Murray, P., & Vandenberg, J. (2008). Galaxy Zoo: Morphologies derived from visual inspection of galaxies from the Sloan Digital Sky Survey. *Monthly Notices of the Royal Astronomical Society*, 389(3), 1179–1189. <https://doi.org/10.1111/j.1365-2966.2008.13689.x>
- Maxwell, A. E., Warner, T. A., & Fang, F. (2018). Implementation of machine-learning classification in remote sensing: An applied review. *International Journal of Remote Sensing*, 39(9), 2784–2817.  
<https://doi.org/10.1080/01431161.2018.1433343>

- Norovsuren, B., Tseveen, B., Batomunkuev, V., Renchin, T., Natsagdorj, E., Yangiv, A., & Mart, Z. (2019). Land cover classification using maximum likelihood method (2000 and 2019) at Khandgait valley in Mongolia. *Earth and Environmental Science*, 381(1).
- Ntwaetsile, K., & Geach, J. E. (2021). Rapid sorting of radio galaxy morphology using Haralick features. *Monthly Notices of the Royal Astronomical Society*, 502(3), 3417–3425. <https://doi.org/10.1093/mnras/stab271>
- Otukei, J. R., & Blaschke, T. (2010). Land cover change assessment using decision trees, support vector machines and maximum likelihood classification algorithms. *International Journal of Applied Earth Observation and Geoinformation*, 12, S27–S31. <https://doi.org/10.1016/j.jag.2009.11.002>
- Pak, M., Lee, J. H., Oh, S., D'Eugenio, F., Colless, M., Jeong, H., & Jeong, W.-S. (2021). Stellar populations of spectroscopically decomposed bulge–disk for S0 galaxies from the CALIFA survey. *The Astrophysical Journal*, 921(1), 49. <https://doi.org/10.3847/1538-4357/ac1ba1>
- Pedersen, K. S., Stensbo-Smidt, K., Zirm, A., & Igel, C. (2013). Shape index descriptors applied to texture-based galaxy analysis. *2013 IEEE International Conference on Computer Vision*, 2440–2447. <https://doi.org/10.1109/ICCV.2013.303>
- Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H.-W. (2002). Detailed structural decomposition of galaxy images. *The Astronomical Journal*, 124(1), 266–293. <https://doi.org/10.1086/340952>
- QGIS 3.16. (2021). *Geographic Information System Developers Manual*. [docs.qgis.org/3.16/en/docs/developers\\_guide/index.html](https://docs.qgis.org/3.16/en/docs/developers_guide/index.html).
- Reis, I., Baron, D., & Shahaf, S. (2018). Probabilistic Random Forest: A machine learning algorithm for noisy data sets. *The Astronomical Journal*, 157(1), 16. <https://doi.org/10.3847/1538-3881/aaf101>

- Richards, J. A., & Jia, X. (1999). *Remote sensing digital image analysis*. Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-662-03978-6>
- RStudio Team. (2020). RStudio: Integrated Development for R; RStudio: Boston, MA, USA. [www.rstudio.com/](http://www.rstudio.com/)
- Scheepmaker, R. A., Lamers, H. J. G. L. M., Anders, P., & Larsen, S. S. (2009). The spatial distribution of star and cluster formation in M 51. *Astronomy & Astrophysics*, *494*(1), 81–93. <https://doi.org/10.1051/0004-6361:200811068>
- Schutter, A., & Shamir, L. (2015). Galaxy morphology—An unsupervised machine learning approach. *Astronomy and Computing*, *12*, 60–66. <https://doi.org/10.1016/j.ascom.2015.05.002>
- Seigar, M. S., & James, P. A. (1998). The structure of spiral galaxies—II. Near-infrared properties of spiral arms. *Monthly Notices of the Royal Astronomical Society*, *299*(3), 685–698. <https://doi.org/10.1046/j.1365-8711.1998.01779.x>
- Shabani, F., Grebel, E. K., Pasquali, A., D’Onghia, E., Gallagher, J. S., Adamo, A., Messa, M., Elmegreen, B. G., Dobbs, C., Gouliermis, D. A., Calzetti, D., Grasha, K., Elmegreen, D. M., Cignoni, M., Dale, D. A., Aloisi, A., Smith, L. J., Tosi, M., Thilker, D. A., ... Pellerin, A. (2018). Search for star cluster age gradients across spiral arms of three LEGUS disc galaxies. *Monthly Notices of the Royal Astronomical Society*, *478*(3), 3590–3604. <https://doi.org/10.1093/mnras/sty1277>
- Shamir, L. (2009). Automatic morphological classification of galaxy images. *Monthly Notices of the Royal Astronomical Society*, *399*(3), 1367–1372. <https://doi.org/10.1111/j.1365-2966.2009.15366.x>
- Shamir, L. (2021). Automatic identification of outliers in Hubble Space Telescope galaxy images. *Monthly Notices of the Royal Astronomical Society*, *501*(4), 5229–5238. <https://doi.org/10.1093/mnras/staa4036>

- Shao, G., Tang, L., & Liao, J. (2019). Overselling overall map accuracy misinforms about research reliability. *Landscape Ecology*, *34*(11), 2487–2492.  
<https://doi.org/10.1007/s10980-019-00916-6>
- Shao, Y., Lunetta, R.S. (2012). Comparison of support vector machine, neural network, and CART algorithms for the land-cover classification using limited training data points. *ISPRS J. Photogramm. Remote Sens.*, *70*, 78–87.  
<https://doi.org/10.1016/j.isprsjprs.2012.04.001>.
- Song, Z., Zhou, S., & Guan, J. (2014). A novel image registration algorithm for remote sensing under affine transformation. *IEEE Transactions on Geoscience and Remote Sensing*, *52*(8), 4895–4912. <https://doi.org/10.1109/TGRS.2013.2285814>
- Spiegel, D., Gehrels, N., Baltay, C., Bennett, D., Breckinridge, J., Donahue, M., Dressler, A., Gaudi, B. S., Greene, T., Guyon, O., Hirata, C., Kalirai, J., Kasdin, N. J., Macintosh, B., Moos, W., Perlmutter, S., Postman, M., Rauscher, B., Rhodes, J., ... Zhao, F. (2015). *Wide-Field Infrared Survey Telescope-Astrophysics Focused Telescope Assets WFIRST-AFTA 2015 Report*. ArXiv:1503.03757 [Astro-Ph]. <http://arxiv.org/abs/1503.03757>
- Thanh Noi, P., & Kappas, M. (2017). Comparison of random forest, k-nearest neighbor, and support vector machine classifiers for land cover classification using Sentinel-2 imagery. *Sensors*, *18*(2), 18. <https://doi.org/10.3390/s18010018>
- Wang, C., Bai, Y., Yuan, H., Wang, S., & Liu, J. (2021). Machine learning applied to STAR-GALAXY-QSO classification of the Javalambre-Photometric Local Universe Survey. ArXiv:2106.12787 [Astro-Ph]. <http://arxiv.org/abs/2106.12787>
- Wei, L., Wang, K., Lu, Q., Liang, Y., Li, H., Wang, Z., Wang, R., & Cao, L. (2021). Crops fine classification in airborne hyperspectral imagery based on multi-feature fusion and deep learning. *Remote Sensing*, *13*(15), 2917.  
<https://doi.org/10.3390/rs13152917>

Zhang, J., Li, P., & Wang, J. (2014). Urban built-up area extraction from Landsat TM/ETM+ images using spectral information and multivariate texture. *Remote Sensing*, 6(8), 7339–7359. <https://doi.org/10.3390/rs6087339>

## Chapter 3

### 3 Age Classification of Spiral Galaxy M83 by Supervised Machine Learning

#### 3.1 Introduction

Galaxies are massive structures composed of stars, dust, and gas that exhibit different spatial and temporal patterns. Spiral galaxies – those having distinctive arm-like features – in particular may be characterized by spiral density waves that move through the galaxy and excite star formation in their path (Lin and Shu 1964). As the waves continue to glide through the galaxy, these young star forming regions are left behind to age and disperse over millions of years. Mapping age of stellar populations in spiral galaxies can contribute to better understanding how these processes take place and over what time scale. Because we cannot observe these galaxies over millions or billions of years, we must make inferences by observing galaxies in different life stages.

Different galactic processes produce unique wavelength emissions. Regions containing young stars are characterized by continuum emission in ultraviolet light and spectral line emission due to ionized gas (such as the Hydrogen-alpha line H-alpha at 657 nm), while older stars' emission peaks at redder wavelengths and is unable to ionize nearby gas. Spectral energy distribution (SED) model fitting is the process of comparing the observed SED of astronomical phenomena to known SEDs (Walcher et al. 2011). SED model fitting has been proven useful for estimation of stellar properties such as age and metallicity and can be done for individual clusters or for entire images of galaxies. Pixel-based age maps of entire galaxies by SED modelling are proven successful, but typically have a lower resolution than that of Hubble Space Telescope imagery or other high resolution telescopes. One disadvantage of the use of imagery from Hubble Space Telescope is the lack of a complete set of wavelength bands needed to reproduce a higher resolution age map. Imagery in several wavelength bands may not be available for some galaxies as telescopes have different sensing capabilities. SED-fitting is limited by its degeneracies, uncertainty, and is time consuming (Walcher et al. 2011), so developing a model that decreases the effects of these limitations would benefit astronomical research.

Sánchez-Gil et al. (2019) derive pixel-based age maps of young stellar populations up to 12 million years in age for a set of spiral galaxies using a hierarchical Bayesian model. The authors use imagery from several telescopes in the H-alpha, ultraviolet, and infrared wavelengths to make age maps on a pixel by pixel basis. They calculate the H-alpha/far ultraviolet ratio for each galaxy, which can identify regions of star formation. Stellar age can also be estimated from the ratio as the H-alpha emission decreases before the far ultraviolet emission decreases as a star forming region ages. This is relevant because ultraviolet emission from young OB stars ionizes the surrounding hydrogen gas to the emission line of H-alpha (657 nm). Evidence of age gradients within several galaxies is found (Sánchez-Gil et al. 2019), contributing to a better understanding fine age patterns of young stars within the spiral.

High-resolution HST pixel analysis of stellar properties have been explored but do not attempt to classify each pixel based on selection of training sites within the imagery. For instance, pixel color magnitude diagram (pCMD) modelling labels each pixel based on the corresponding color magnitude diagram (Lee et al. 2011, 2012). This allows for a visual interpretation of typical color magnitude diagrams that describe the brightness of stars in different wavelengths. Lee et al. (2011, 2012) make use of Hubble Space Telescope imagery of galaxy M51 and its companion galaxy NGC 5195 and perform high-resolution analysis to derive the pCMD characteristics of each pixel. The pixels are grouped into 4x4 bins so that the pixel brightness in different wavelength bands is more robust. Each pixel within galactic imagery is an average of the emission from the area of sky the pixel encompasses. Therefore pixel values reflect the average stellar age within the pixel extent, making pixels closer together more related. The pixel binning slightly decreases the resolution of the age map, but does address some uncertainty in the imagery. The pCMD describes the red (age >1 billion years) and blue sequence (age <1 billion years) membership of the pixels, so estimation of fine age distinctions is not the goal of the work by Lee et al. (2011, 2012). They find that the distribution of the stellar populations in M51 is consistent with the density wave theory where the younger stars are more concentrated on the spiral arms and the older stars are more dispersed in the stellar field. NGC 5195, on the other hand, has a larger population of old stars with some evidence of star formation

where the two galaxies meet. The pCMD model is a useful method for deriving stellar properties for a variety of galaxies.

Wei et al. (2020) also perform an analysis of galaxies M51 and NGC 5195 by fitting synthetic SEDs to the spectral emission from the pixels within the galaxy system. The authors make use of 28 multi-band images in ultraviolet to mid infrared wavelengths from several telescopes, and produce pixel-based maps of stellar properties with a pixel resolution of 1.7 arcsecond, much lower than the resolution of Hubble Space Telescope imagery. They find that M51 exhibits a stronger age gradient in the regions closer to the galaxy center (inner disc) and weaker age gradients in the outer disc.

From the literature review, we identify a lack of pixel-based age estimation of nearby and well-resolved galaxies. Developing pixel-based maps of stellar properties of nearby and well-resolved galaxies is a well-studied method, but often produces lower resolution maps than that of Hubble Space Telescope. The use of a stellar fitting model requires imagery in many wavelength bands which often are not available from a single telescope. SED model fitting is proven to be extremely useful for stellar property estimation within galaxies, including estimation of age. However, models that achieve high-resolution age estimation of galaxies spanning from stellar populations with finer estimation of ages in millions of years to ages in billions of years would provide an in-depth view of galaxy mechanics of formation and evolution.

In this paper, we classify age of young stellar populations in spiral galaxy M83 with a pixel-based supervised machine learning classification model. Using the M83 age map produced by Sánchez-Gil et al. (2019), we train random forest and support vector machine learning models. We strive to develop a method of stellar age estimation that requires fewer images and decreased processing time. This would allow for the use of high-resolution imagery from Hubble Space Telescope and James Webb Space Telescope (Gardner et al. 2006) as well as from other high-resolution telescopes such as The Cosmological Advanced Survey Telescope for Optical and ultraviolet Research (CASTOR) proposed by the Canadian Space Agency (Côté et al. 2019). We identify the following research objectives:

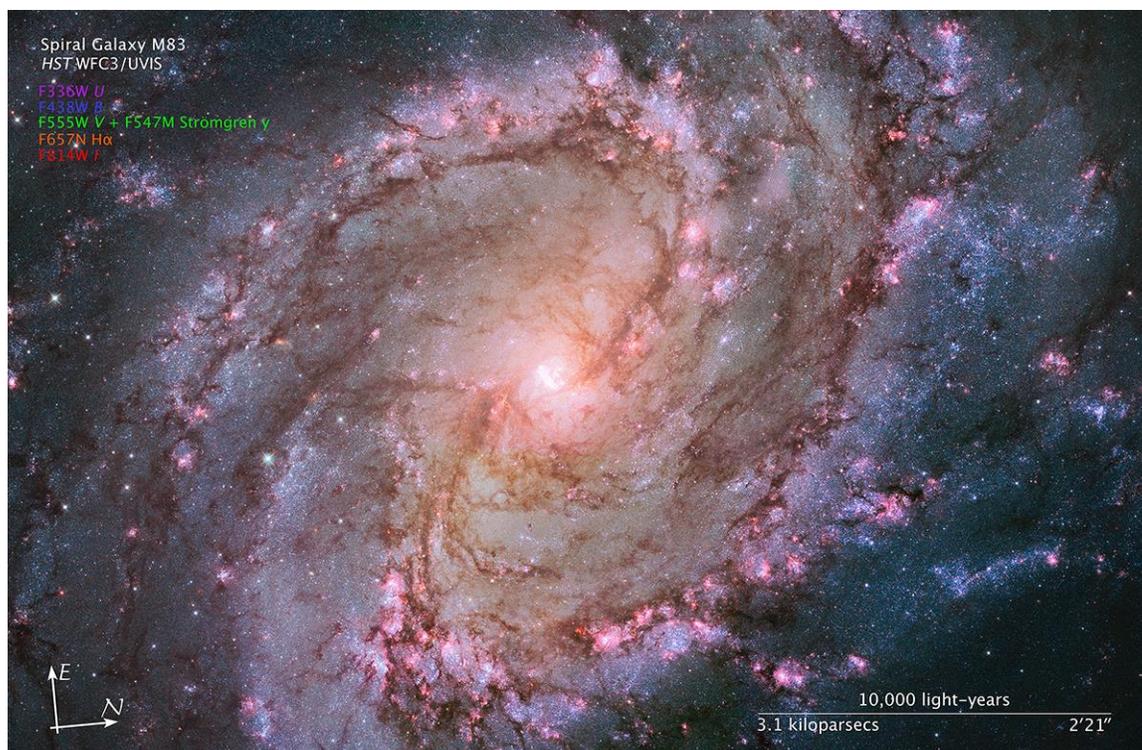
1. Determine how machine learning can be used to classify age of stellar populations within galaxies,
  - a. Which parameters are most useful or important for age classification?
  - b. Which stellar ages are best classified?
  - c. How can lower resolution age maps be used as reference data to classify age of stellar populations within galaxies using high resolution Hubble Space Telescope imagery?

Achieving these objectives can contribute to more rapid estimation of stellar properties with less data required, while maintaining high pixel resolution.

## 3.2 Methods

### 3.2.1 Study Area

In this paper we study M83, a grand design spiral galaxy with two prominent spiral arms. M83 is a well-studied nearby galaxy at a distance of about 15 million light years with a nearly face on inclination of  $24^\circ$ , making the galaxy components within it well resolved (Sánchez-Gil et al. 2019). Figure 3-1 shows regions of star formation as the pink bubbles; these regions are indicators of star formation as the young and hot stars ionize the surrounding hydrogen gas, producing the Hydrogen-alpha line H-alpha at 657 nm. Age gradients are thought to exist within galaxy M83 due to the spiral density waves dominating it (Sánchez-Gil et al. 2019). As the waves move they excite star formation in matter they pass through. Eventually the density wave may leave the star forming regions where the young stars age and slowly disperse. M83 is an excellent study area for observing the stages of star formation and evolution (Blair et al. 2014), prompting us to study it in detail.



**Figure 3-1. Hubble Space Telescope map of galaxy M83: Purple: ultraviolet band (F336W), Blue: blue band (F438W), Green: visual band (F555W) + Strömgren y (F547M), Orange: H-alpha (F657N), Red: near-infrared (F814W). Image credit: NASA, ESA, and the Hubble Heritage Team (STScI/AURA)**

### 3.2.2 Data

#### 3.2.2.1 M83

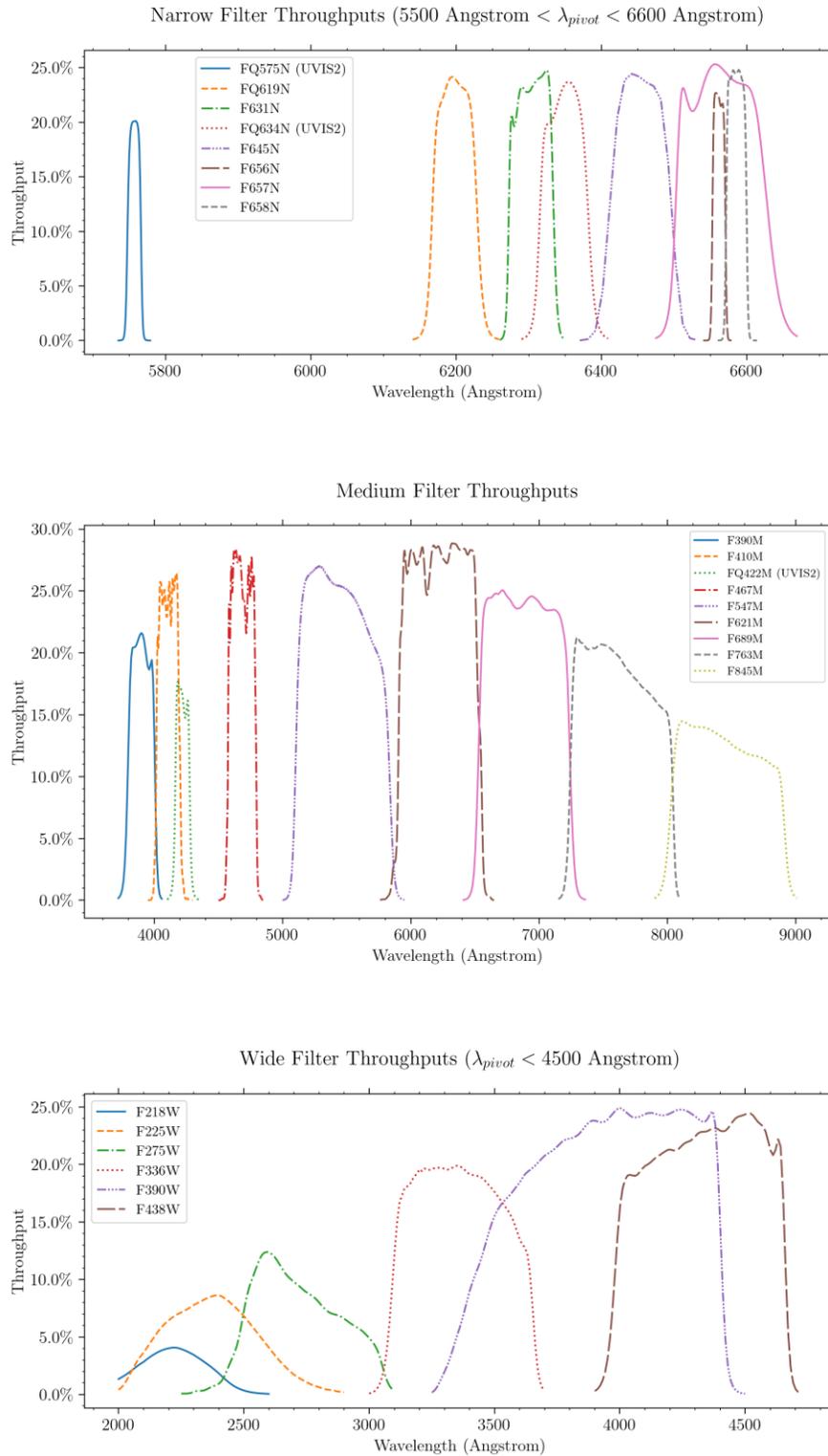
We retrieve M83 mosaics produced by Blair et al. (2014) from the Mikulski Archive for Space Telescopes website. We use M83 mosaics in filters F438W (blue), F555W-F547M (visual), F814W (near infrared), and F657N (H-alpha) as taken by the Wide Field Camera 3 (WFC3) aboard the HST. These mosaics are available as part of the Early Release Science Program 11360 (PI: R. O'Connell) and cycle 19 HST General Observer Program 12513 (PI: W. Blair) and have a resolution of 0.0396'' (arcsecond) per pixel. Blue, visual, and near infrared bands are useful for observation of stars and analysis of their properties such as age (Kiar, Barmby, and Hidalgo 2017) whereas H-alpha is an indicator of very young and hot OB stars, and therefore useful for identifying regions of star formation (Sánchez-

Gil et al. 2019). Wavelength ranges for the M83 bands are reported in Table 3-1 and visual representations are shown in Figure 3-2. Figure 2-2 shows the filter wavelength range for F814W.

The HST image mosaics are composed of images taken in seven sky positions, called fields. Field 1 of the F555W-547M image is taken with the F555W wavelength filter and the remainder of the fields are taken with the F547M wavelength filter. Only one F555W image is used to prevent any emission-line effects (Blair et al., 2014). The F555W image is calibrated to match the scale of the F547M Fields.

**Table 3-1: HST wavelength bands for M83 and their respective wavelength ranges.**

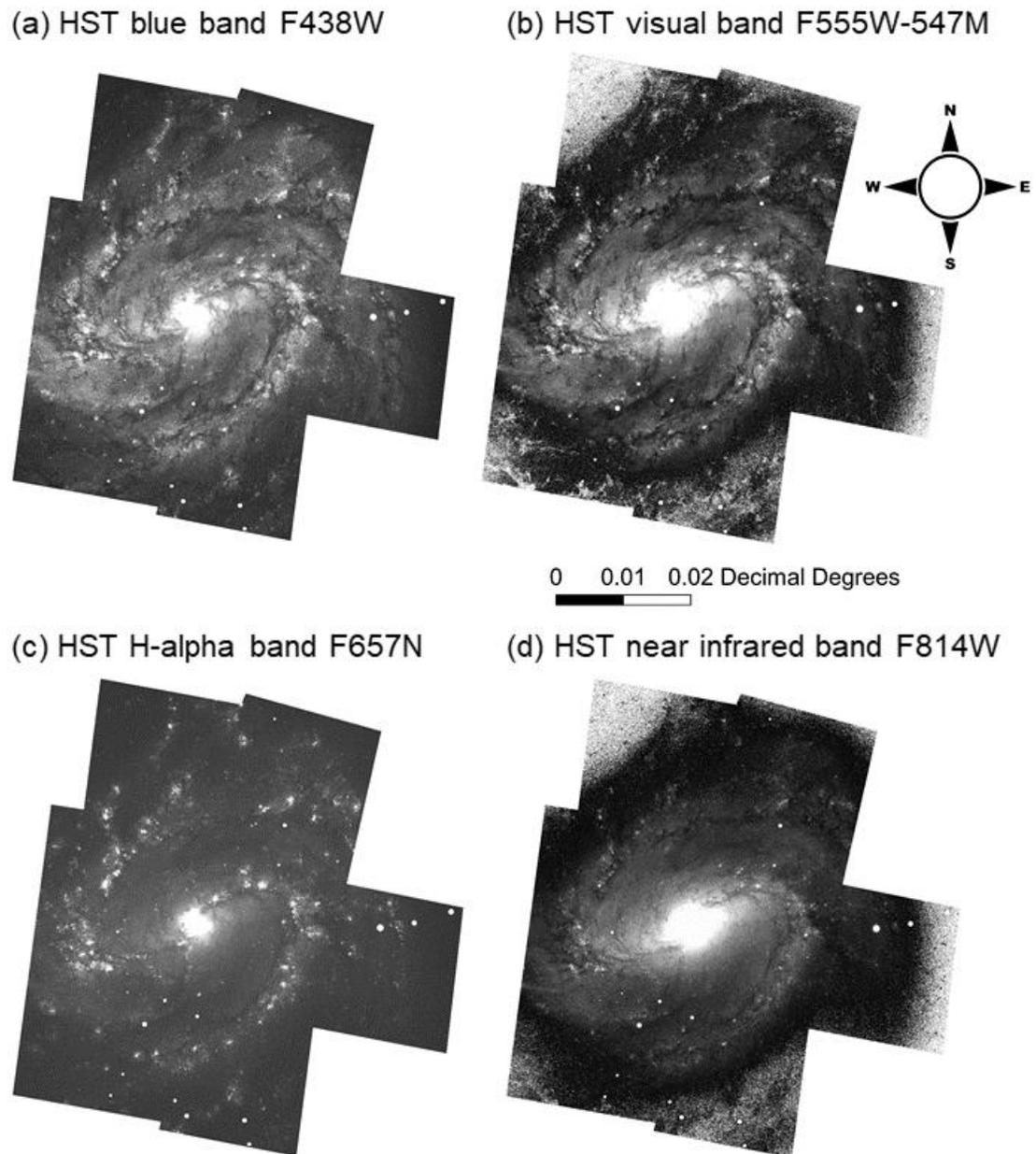
<b>Wavelength band</b>	<b>Width (nm)</b>
F438W	61.47
F547M	65
F657N	12.1
F814W	156.52



**Figure 3-2: HST WFC3 narrow, medium, and wide band filter wavelength ranges. (Dressel et al., 2022)**

### 3.2.2.2 HST Image Processing

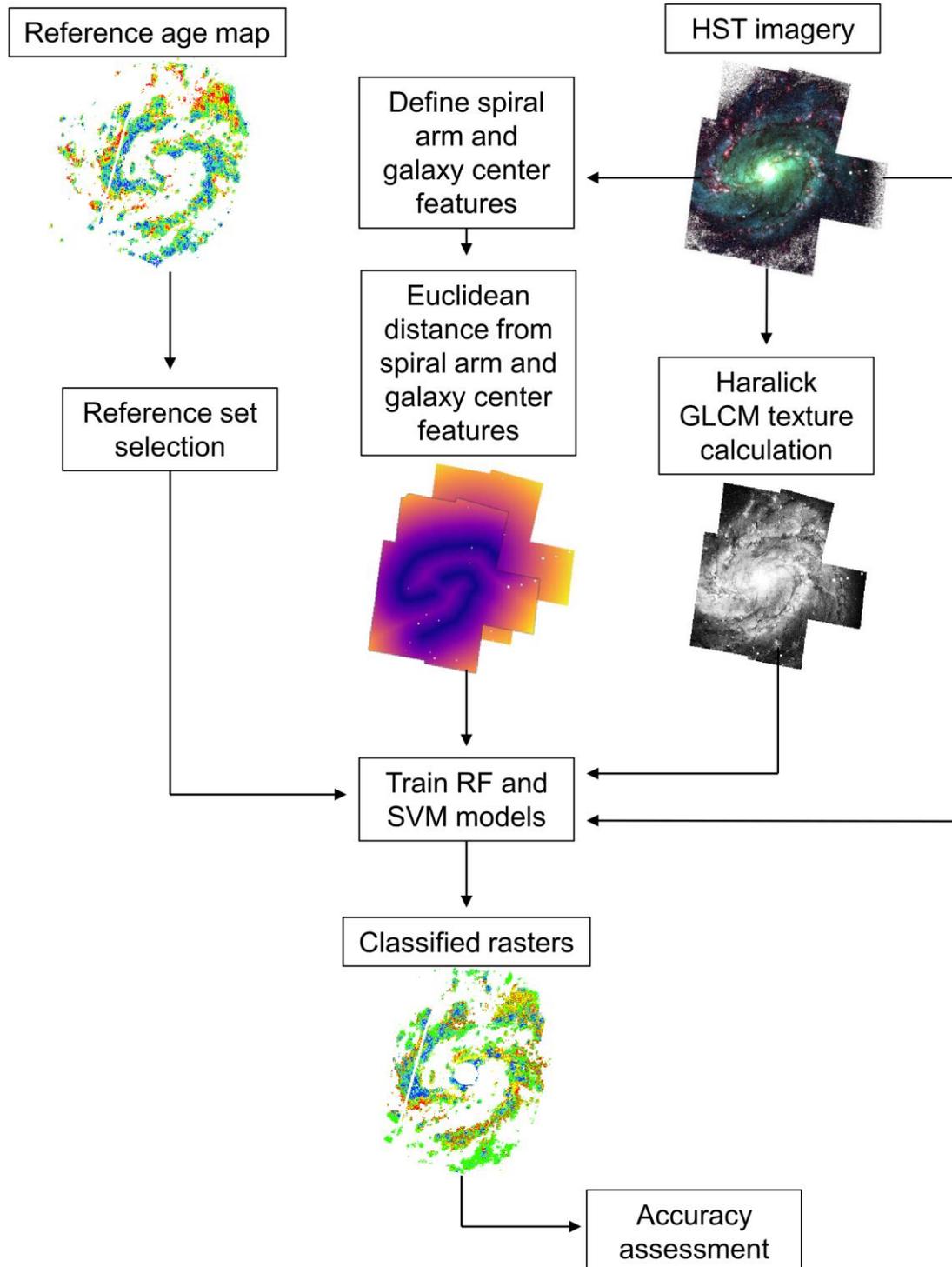
The mosaics retrieved are in Flexible Image Transport System (FITS) file format. We use FITS Liberator 3 (ESA/ESO/NASA 2021) software to convert the FITS mosaics to Tag Image File Format (TIFF) for import into a Geographic Information System (GIS). The exported mosaics are geometrically corrected in open source software QGIS Version 3.16.3. (QGIS 3.16. 2021) using the “Georeferencer” tool. We select ground control points from the original FITS images using the open source SAOImageDS9 software Version 8.0.1. (Joye and Mandel 2003) that supports astronomical coordinate systems found in the FITS imagery. Geometric correction of the imagery alters the orientation of the image to Earth coordinates where North is upwards and East is to the right. In order to avoid distortion within the HST imagery, we use a Helmert transformation that performs simple shifting and rotating of the image (Farhadian and Clarke 2020; Song et al. 2014). We then combine the four geometrically corrected images shown in Figure 3-2 into a composite raster for further processing in a GIS. Foreground stars projected onto galaxy are also identified and masked from the imagery.



**Figure 3-3. HST imagery of M83 geometrically corrected to Earth orientation (East to the right): a) HST blue band F438W; b) HST visual band F555W-547M; c) HST H-alpha band F657N; d) HST near infrared band F814W.**

### 3.3 Parameter Creation

Following HST image processing, we create parameters from the geometrically corrected HST raster stack. Figure 3-3 describes the methodology to be discussed in the remainder of the methods section. We produce Euclidean distance from the traced spiral arms and galaxy center features. We also produce textural features calculated from the HST imagery. Different combinations of parameters are input into trained machine learning models, and accuracy assessment is performed on the classified results.

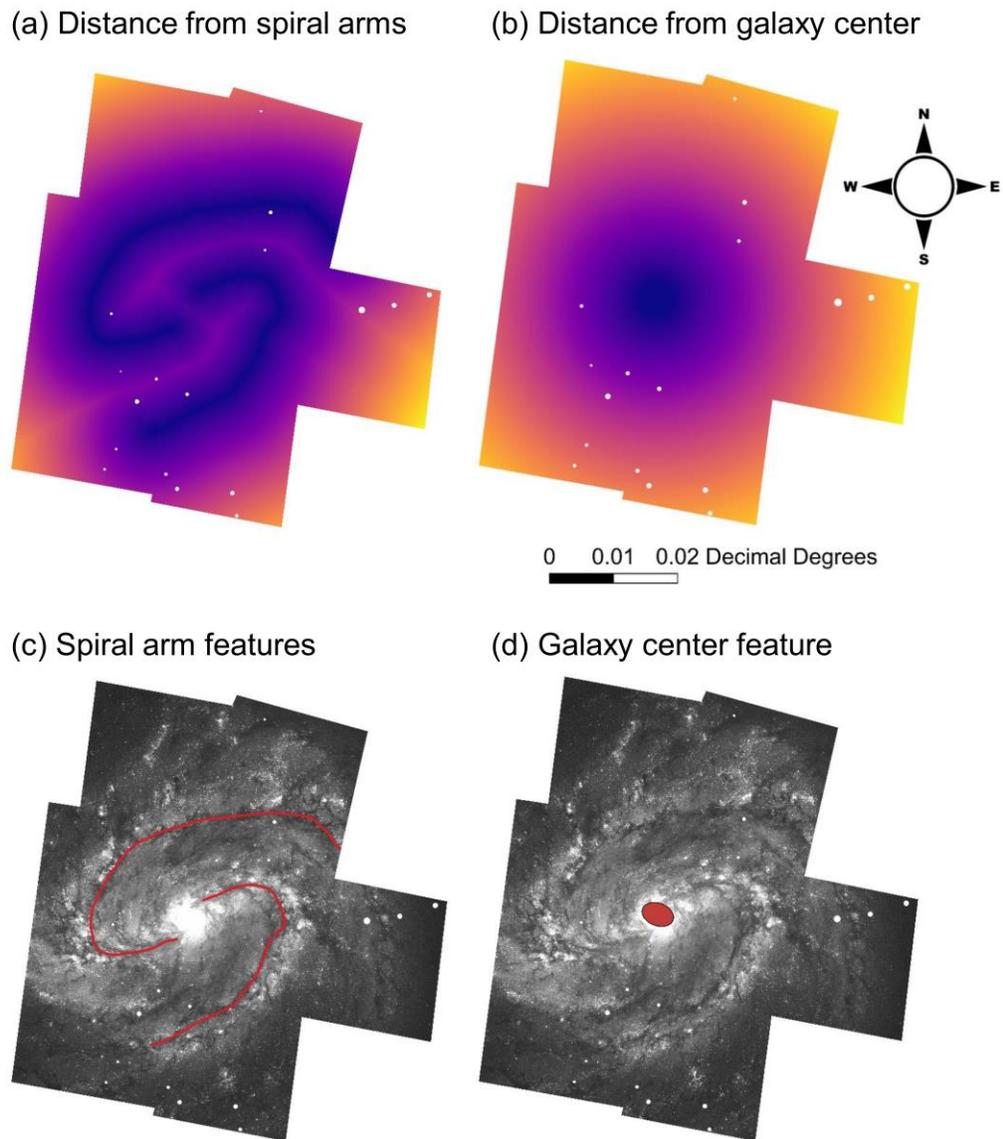


**Figure 3-4. Flowchart of methodology used in this study**

### 3.3.1 Euclidean Distance

Age gradients are thought to exist within galaxy M83 due to the spiral density waves dominating the galaxies (Abdeen et al. 2022). As the waves move they excite star formation in matter they pass through. Eventually the density wave may leave the star forming regions where the young stars age and slowly disperse. Euclidean distance, a commonly used GIS method, may be useful in identifying age gradients and classifying age of stars.

To define the spiral arms, we follow methods described by Shabani et al. (2018) by tracing along the dust lanes that run through the inner edge of the spiral arms. We use the blue band imagery (F438W) to identify the dust lanes as the dust lanes are darkest and most apparent in the blue band. The dust lanes are prominent in HST imagery of M83 and are therefore good approximations of the spiral arm locations. After defining the spiral arms and galaxy center for M83, we calculate Euclidean distance for each pixel in the raster images based on the linear distance to the nearest feature (Figure 3-4). The GIS used to calculate Euclidean distance first converts the spiral arm and galaxy center feature to pixels. It next calculates the linear distance from each pixel in the HST band raster extent to the nearest pixel in the spiral arm and galaxy center features. We mask the foreground stars from the distance rasters. As mentioned in Chapter 2, the manual tracing of the spiral arms may not be representative of the true center of the arms, so some uncertainty is present in the calculations of distance. In M83, there are multiple dust lanes within the Southern arm. We choose the dust lane closer in proximity to the young stellar regions, however, other nearby dust lanes could also be traced.



**Figure 3-5. Euclidean distance features drawn over the HST imagery and their respective distance calculations: a) Euclidean distance from spiral arm features; b) Euclidean distance from galaxy center feature; c) Spiral arm line feature traced over the HST blue band image; d) Galaxy center polygon feature drawn over galaxy center. Background image is the HST blue band (F438W) shown in grayscale**

### 3.3.2 Texture

Image texture calculation is used to identify texture of natural phenomena. Textural feature extraction is a common remote sensing method used for a variety of purposes: examples include land cover classification (Rajendran et al. 2020), and crop yield prediction (Yang et al. 2022). In Kwik et al. (2022), we found texture particularly useful for classification of galaxy components in UGC 2885, a nearby spiral galaxy at a distance of 79.1 megaparsecs (Mpc; Hunter et al. 2013). For age classification, texture may also be useful. Younger star clusters have more pixel variation than older star clusters (Whitmore et al. 2011) so are lumpier than the older stars and stellar field (Jensen et al. 1981). Therefore, we expect texture to successfully differentiate between the age classes.

Here we calculate texture on the HST imagery of galaxy M83. Because M83 is much closer than UGC 2885, the stars and star clusters are better resolved. One goal of this study is to test how texture performs on nearby and well-resolved star forming galaxy M83 and how texture can help to classify age of stellar populations. We calculate eight gray level co-occurrence matrix (GLCM) Haralick textures (Haralick, Dinstein, and Shanmugam 1973) for the M83 imagery and test to see which may be most useful for age classification: angular second moment, contrast, correlation, dissimilarity, entropy, homogeneity, mean, and standard deviation. From the texture rasters we also mask the foreground stars projected onto the galaxy.

The GLCM is calculated from pixel values within a moving window. Two pixels are considered at a time: the reference pixel and the neighbouring pixel. A GLCM counts how often different combinations of pixel values occur by looking at the relationship between the reference and neighbouring pixel, so the algorithm runs through all possible reference and neighbouring pixel combinations within the window in order to complete the GLCM. After completion of the GLCM, second order statistics are calculated from its contents (Haralick, Dinstein, and Shanmugam 1973) and the resulting value is assigned to the central pixel of the moving window. The algorithm continues until all pixels within the image have been assigned a new value. We measure the diameter of star clusters within the galaxies to determine the window size and choose an 11x11 window, meaning that 121 pixels are considered for GLCM creation. The original pixel values within the imagery are

also scaled down to 64 gray levels prior to GLCM calculation. We discuss the eight Haralick textures below and show the results of the texture calculations on the HST blue band (F438W) in Figure 3-5.

Angular second moment (ASM) texture calculates uniformity of pixel values within an image. ASM texture is calculated on the GLCM using Eq. (1):

$$\text{Angular second moment} = \sum_{i=1}^L \sum_{j=1}^L (\hat{P}(i, j))^2 \quad (1)$$

Variable  $P(i, j)$  in Eq. (3) represents the probability of the  $(i, j)$ th entry on the GLCM, meaning that it describes how often the reference and neighbouring will have pixel values of  $i$  and  $j$ . ASM will calculate the highest values for windows with only a few gray levels.

Entropy texture calculates chaos or unevenness of an image. Higher values of entropy will therefore indicate more chaos within the moving window. Eq. (2) shows the entropy calculation:

$$\text{Entropy} = - \sum_{i=1}^L \sum_{j=1}^L \hat{P}(i, j) \log[\hat{P}(i, j)] \quad (2)$$

Contrast and dissimilarity texture both calculate spectral variation in the imagery. Imagery with high spectral variation will result in higher values of contrast. We expect the smoother stellar field and inter-arm regions to have lower values of contrast. The rougher spiral arms and star forming regions should have higher values of contrast due to the spectral variations in these regions. Contrast is defined by Eq. (3):

$$\text{Contrast} = \sum_{n=0}^{L-1} (i - j)^2 \sum_{i=1}^L \sum_{j=1}^L \hat{P}(i, j) \quad (3)$$

where the diagonal of the GLCM is assigned a value of 0 (when  $i$  and  $j$  are equal). The larger the difference between the gray levels of  $i$  and  $j$ , the higher the value.

Dissimilarity calculates spectral variation with the absolute difference between  $i$  and  $j$  co-occurrence matrix rather than the difference squared as calculated by contrast in Eq. (4) and is calculated as:

$$\text{Dissimilarity} = \sum_{i=1}^L \sum_{j=1}^L |i - j| \hat{P}(i, j) \quad (4)$$

Homogeneity is similar to contrast and dissimilarity where it calculates the difference or contrast between  $i$  and  $j$  gray levels in the GLCM. Rather than the diagonal values being the lowest as with contrast and dissimilarity, the diagonal is the highest with a value of 1 for homogeneity calculation. Moving away from the diagonal, the values decrease. Eq. (5) shows the homogeneity calculation:

$$\text{Homogeneity} = \sum_{i=1}^L \sum_{j=1}^L \frac{\hat{P}(i,j)}{1+(i-j)^2} \quad (5)$$

Mean texture uses the mean descriptive statistic. Rather than calculating the average of the pixel values in the local window, mean texture calculates the average of values within the GLCM (Eq. 6):

$$\text{Mean} = \sum_{i=1}^L \sum_{j=1}^L i * \hat{P}(i,j) \quad (6)$$

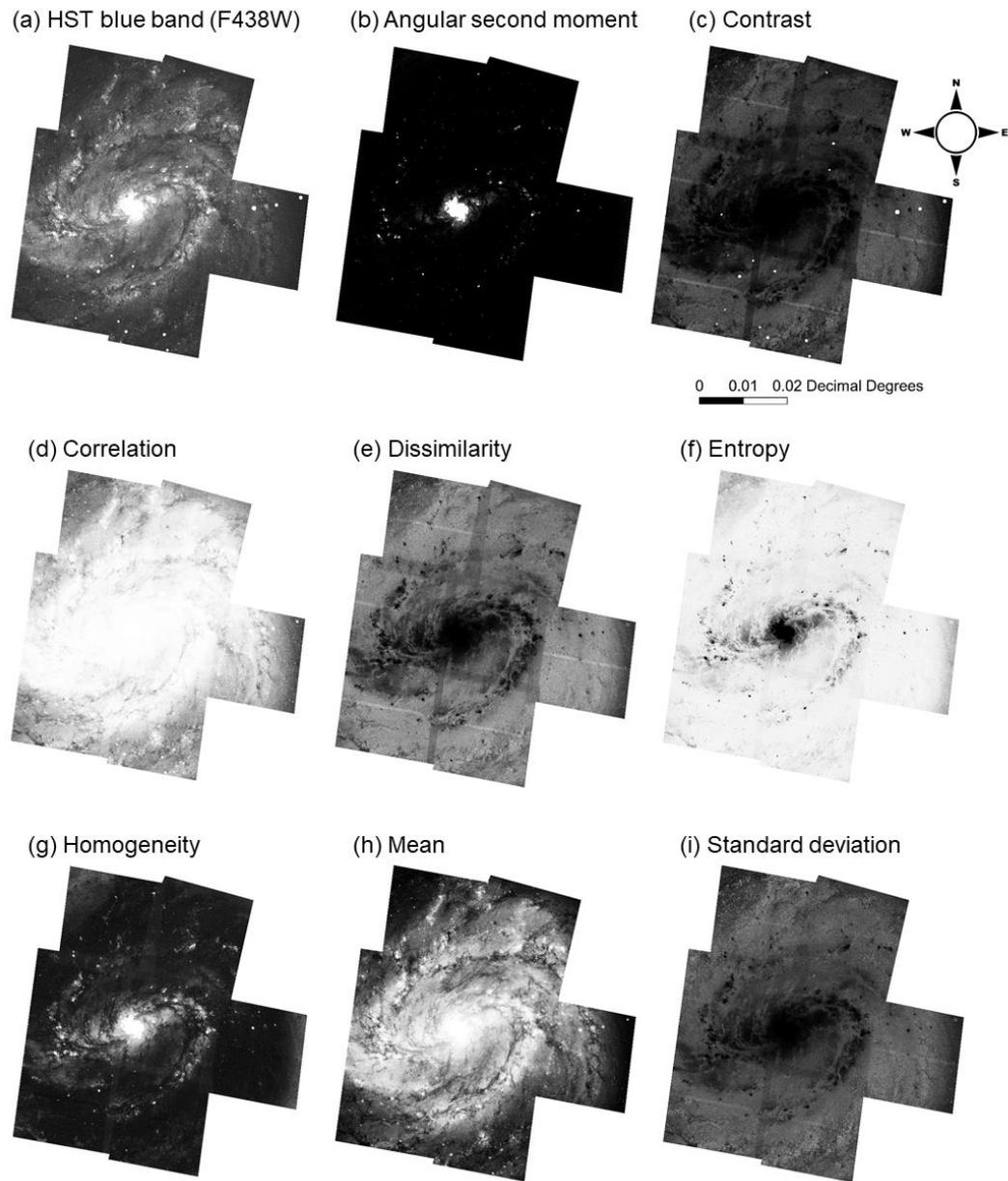
Similar to mean texture, standard deviation texture uses the standard deviation descriptive statistic. Standard deviation texture measure the variation of values on either side of the mean value. Eq. (7) describes the equation used for standard deviation texture calculation:

$$\text{Standard Deviation} = \sum_{i=1}^L \sum_{j=1}^L (i - \mu)^2 \hat{P}(i,j) \quad (7)$$

The final texture, correlation, calculates the relationship of neighbouring pixels in the GLCM. If there is high correlation, it means that the neighbouring pixels are similar or linear in nature. Typically, the closer neighbouring pixels are, the more correlated they will be. Therefore, when using large moving windows, correlation might be lower than if using a small moving window on the image. The equation for correlation texture, Eq. 8, is described as:

$$\text{Correlation} = \sum_{i=1}^L \sum_{j=1}^L ((i * j * \hat{P}(i,j) - \mu_x \mu_y) / \sigma_x \sigma_y) \quad (8)$$

where  $\mu_x$  and  $\mu_y$  are the average of  $P_x$  and  $P_y$ , and  $\sigma_x$  and  $\sigma_y$  are the standard deviations of  $P_x$  and  $P_y$ .

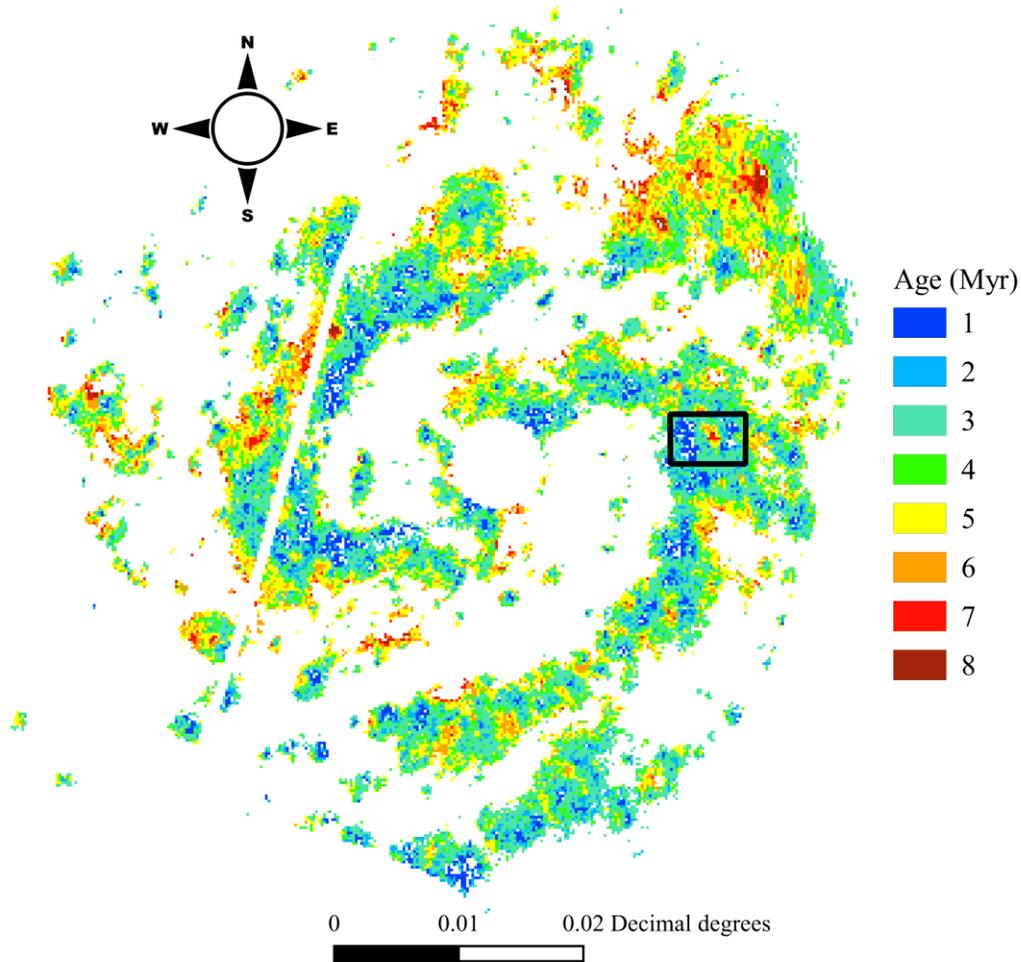


**Figure 3-6. Haralick textures calculated from HST blue (F438W) imagery of M83: a) HST blue band image (F438W); b) angular second moment texture; c) contrast texture; d) correlation texture; e) dissimilarity texture; f) entropy texture; g) homogeneity texture; h) mean texture; i) standard deviation texture. White color represents higher values and black represents lower values.**

### 3.4 Reference Data

We use the M83 age map from Sánchez-Gil et al. (2019) as an approximate ‘ground truth’ or reference data (Figure 3-6). Sánchez-Gil et al. (2019) estimate ages of young stars up to nine million years within galaxy M83. The age map is produced using H-alpha ( $H\alpha$ ) and far ultraviolet (FUV) emission imagery of M83 and a hierarchical Bayesian model. As a young star ages, H-alpha emission from star forming regions drops off more quickly than FUV emission. Therefore, the H-alpha/FUV ratio is found to be useful for age estimation (Sánchez-Gil et al. 2019). The age map does not estimate age of the interarm regions as there are smaller concentrations of young stellar populations outside of the main spiral arms.

To create reference sites from the age map for classification, the age map must be geometrically corrected to overlay on the HST imagery. We perform the same procedure as discussed in Section 3.2.2, performing geometric correction in the open source software QGIS Version 3.16.3. (QGIS 3.16., 2021) using the “Georeferencer” tool. We use the youngest regions in the age map (1 and 2 Myr; Figure 3-6) as ground control points to match to the young star forming regions indicated by the H-alpha emission in the HST imagery. Geometric correction using the Helmert transformation used for geometric correction of the HST imagery was not successful in matching the age map to the HST imagery. Instead we opt to use the thin plate spline transformation that warps the local area surrounding the control points (Zhou et al. 2013). This allows us to more precisely match the young stellar populations in the age map to the regions with high H-alpha emission in the HST imagery. The geometrically corrected age map is shown in Figure 3-6. The black box within Figure 3-6 indicates the location of the inset shown in Figure 3-7.



**Figure 3-7. Original age map of young stellar populations within M83 (Sánchez-Gil et al. 2019). Age is in millions of years (Myr). Blue pixels represent the youngest regions while red pixels represent the oldest. The black box indicates the location of the inset shown in Figure 3-7.**

### 3.4.1 Classification Scheme

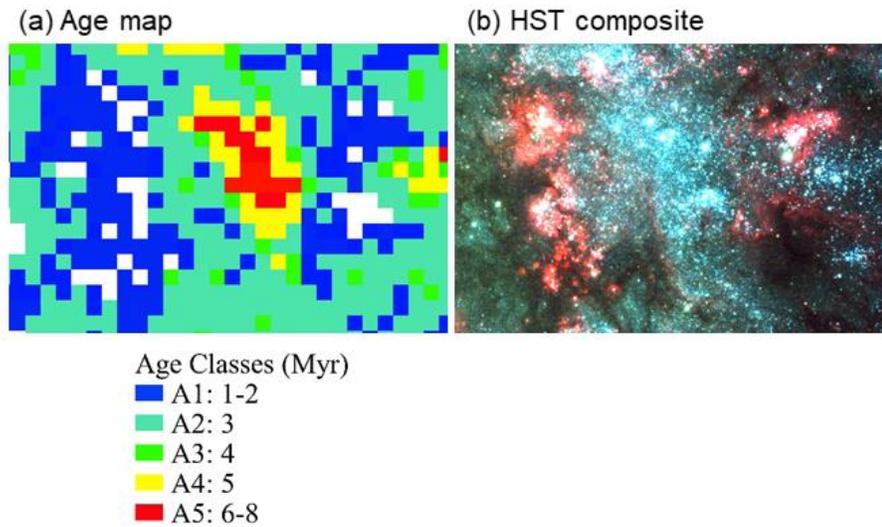
From the age map, we derive several age classes to be used for classification. Figure 3-7 compares the HST image and age map resolution from the same region of M83 (black box in Figure 3-6) that is 0.0040065 by 0.0060895 degrees in size as measured on the geometrically corrected imagery. For galaxy M83, one arcsecond as projected on the sky is equal to about 22 parsecs within M83 (Blair et al. 2014; Sánchez-Gil et al. 2019). Therefore, this region represents 317.24 x 482.24 parsecs according to our geometrically

corrected imagery. Due to their low frequency within the age map, the youngest pixels (1 and 2 Myr) are grouped together into class A1. We also group the oldest pixels (6, 7, and 8 Myr) together into their own age class A5 for the same reason. The intermediate aged pixels (3, 4, and 5 Myr) are more common within the age map, so each get their own class. In total, five age classes are chosen (Table 3-1).

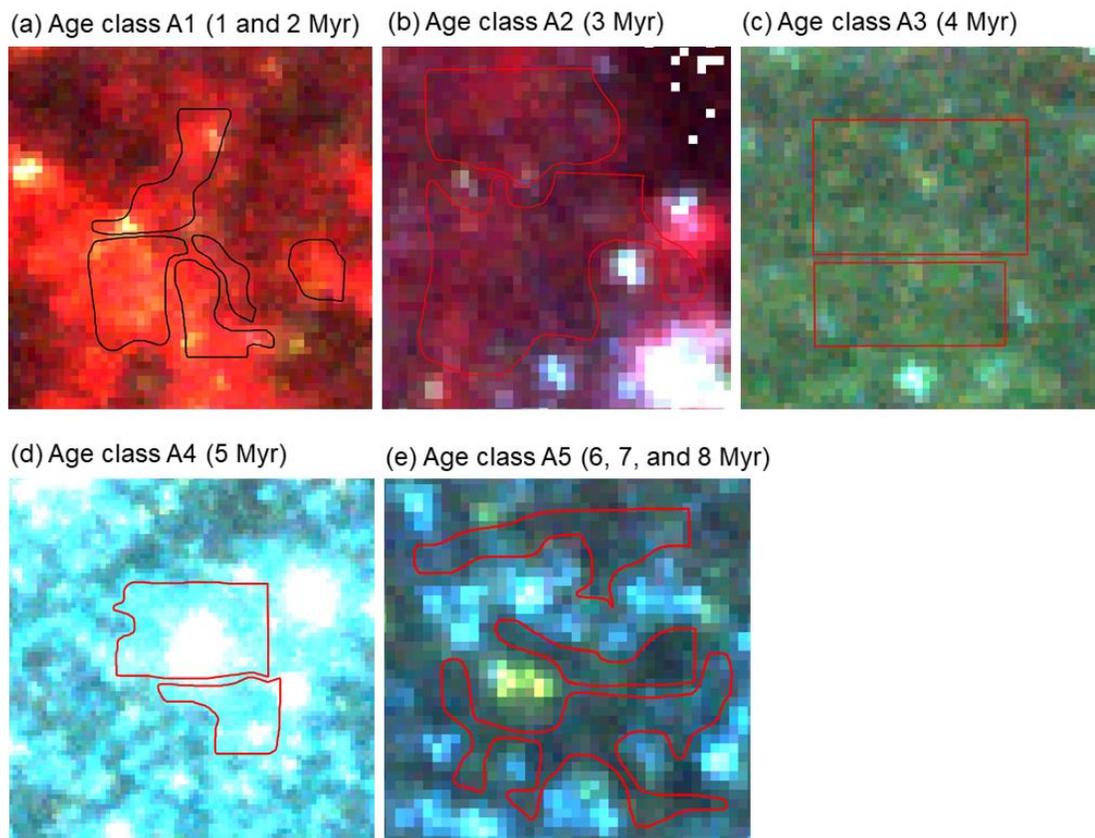
To prevent bias in reference site selection, we randomly select pixels by generating random samples within the M83 age map extent. We examine the HST imagery underneath the age map pixels and trace polygons over groups of pixels in the HST imagery that are best representative of the age map pixel. We do this to account for differences in the resolution of the age map pixels (1.5") versus the HST imagery (0.0396"). Larger pixels will be an average of the band flux of the smaller HST pixels. Figure 3-8 shows some examples of the pixel tracing done over the HST imagery.

**Table 3-2. Chosen classification scheme of age map pixels.**

<b>Age Map Pixels (Myr)</b>	<b>Class ID</b>
1 and 2	A1
3	A2
4	A3
5	A4
6, 7, and 8	A5



**Figure 3-8. Pixel resolution comparison for the same region of the HST imagery and the M83 age map as shown on the black box in Figure 3-6: a) HST band imagery; b) reference age map. The HST imagery is an RGB color composite where R – F657N, G – F555W-547M, and B – F438W. The images represent an area of approximately 317.24 x 482.24 parsecs as measured on the geometrically corrected imagery. The red “bubbles” within the HST imagery is H-alpha emission showing the young star forming regions.**



**Figure 3-9. Training sites drawn over color composite of HST imagery of M83: a) age class A1 (1 and 2 Myr); b) age class A2 (3 Myr); c) age class A3 (4 Myr); d) age class A4 (5 Myr); e) age class A5 (6, 7, and 8 Myr). Colour composite RGB assignment: R – F657N, G – F555W-547M, B – F438W.**

### 3.5 Machine Learning Classification

In Kwik et al. (2022) we find that the traditional maximum likelihood classifier model is not as useful for machine learning classification of galaxy imagery. Therefore, we choose to compare the more successful ML models, random forest (RF) and support vector machine (SVM) that had high accuracy of galaxy component classification.

#### 3.5.1 Support Vector Machine

The SVM model identifies support vectors within the training data and defines a hyperplane separating the classes defined within the training data. Support vectors are data

points closest to a hyperplane that define the maximal margin between the hyperplane and support vectors. The advantage of the SVM algorithm is that it is excellent at handling data samples whose values do not exhibit linear patterns (Sen et al. 2022). The SVM model was originally described in (Cortes and Vapnik 1995).

We perform parameter selection testing on the HST imagery using 50, 100, 250, 500, 750, and 1000 samples. 250 samples results in an OA of 3% less than classification with 500 samples, whereas 750 and 1000 samples did not result in drastic increases in accuracy from 500 samples. Therefore, we choose to use 500 samples for SVM classification, as increasing the sample number to 750 or 1000 is not cost effective. SVM model classification is performed in a GIS.

### 3.5.2 Random Forest

Random forest (RF) is a powerful machine learning model that makes use of a set or ‘forest’ of decision trees to process and make decisions regarding samples (Breiman 2001). RF is advantageous in that it produces highly accurate results without long processing times (Sen et al. 2022).

We perform RF classification in a GIS and test 50, 100, 250, 500, and 750 trees. There is no drastic difference in overall accuracy (<1%) when testing the number of trees, with 500 trees having a slightly higher accuracy than the rest of the tests. Therefore we use 500 trees as recommended by Belgiu and Dragut (2016). Setting the number of trees to 500 allows the model error to stabilize (Lawrence et al. 2006), so choosing a lower number of trees may introduce more error into the final classification result.

After testing of the number of trees, we perform testing of the tree depth. For this, the number of trees is set to the chosen 500. We test 15, 30, and 50 tree depths. A tree depth of 15 results in the highest overall accuracy and the shortest processing time. Because we find no increase in accuracy from 15 to 30 and 50 trees, we do not test a higher number of trees.

Mean decrease in Gini coefficient (MDG) is a useful statistic calculated from RF classification that measures the importance of variables input into a classification model

(Koo et al. 2021). Because MDG analysis is not available in the GIS used to perform classification, we run MDG analysis in R Studio Version 1.1.463 (RStudio Team, 2020) using the “randomForest” package. Input for mean decrease Gini analysis is a raster of all 38 parameters: 4 HST imagery + 2 distance rasters + 32 textural features. In case the machine learning models have difficulty handling parameters with different value ranges, we perform normalization of the 38 input parameters by dividing each by their maximum pixel value, creating a 0 – 1 raster value scale for all parameters. Normalization is done to prevent the overemphasis of certain parameters in classification results.

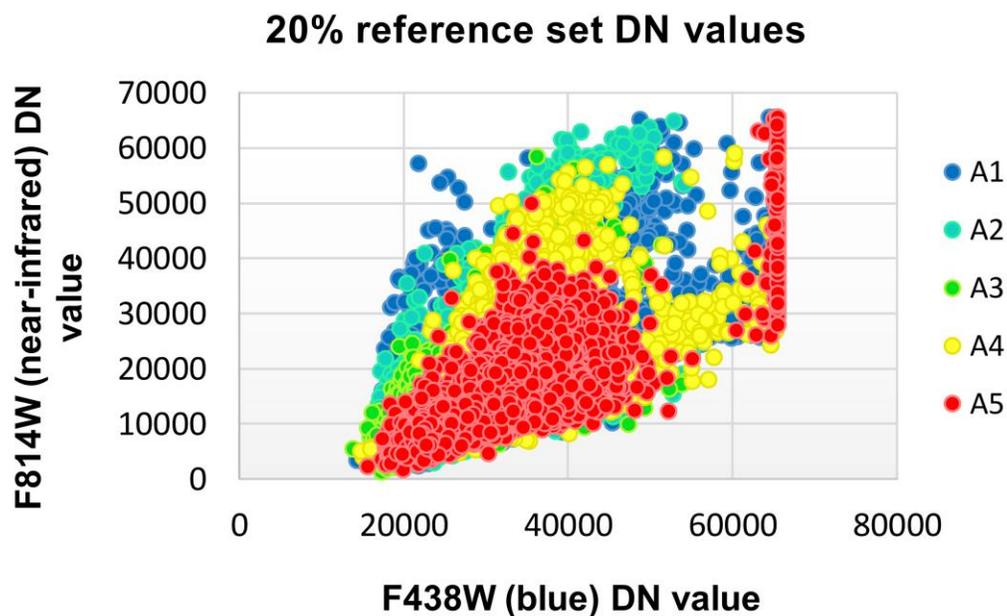
After classification we mask the HST mosaics to the same extent as the age map. To do this, we convert the age map raster to vector and use this to mask the classified result in a GIS. We can ignore the other portions of the galaxy in the classified rasters because they are not included in the age map used for accuracy assessment.

### 3.6 Accuracy Assessment

For accuracy assessment, we convert the polygon feature reference sites discussed in Section 3.4.1 to point features; this is done by generating a point on every pixel in the HST imagery, and masking the point feature file to the extent of the reference set polygon features. The GIS uses the values of the pixels beneath the point features to classify the imagery. We test 70/30, 80/20, and 90/10 splits of the points to determine which split results in the highest accuracy. We find that using an 80/20 split results in the highest accuracy of classification. Table 3-2 shows the number of points for each age class within the 80% and 20% splits representing 75,299 and 18,825 respectively, with the total number of reference points being 94,124. The digital number values for the five age classes within the 20% split are shown in Figure 3-9. There is quite a bit of overlap amongst the age classes, however, the scatterplot does show general patterns where the oldest ages (A5) are the dimmest pixels within the young stellar regions (low DN values), and the youngest ages (A1) are the brightest (high DN values). Pixel saturation can be seen in Figure 3-9 where many pixels create a linear feature at the 65535 DN value in the F438W image.

**Table 3-3. Number of M83 pixels for each class within the 80% and 20% splits of the reference set.**

Age class	80% of reference set points	20% of reference set points
A1: 1-2 Myr	14,541	3,637
A2: 3 Myr	13,646	3,481
A3: 4 Myr	16,152	4,003
A4: 5 Myr	15,580	3,814
A5: 6-7-8 Myr	15,380	3,890



**Figure 3-10. Scatterplot showing the original digital number (DN) values of the HST imagery (0 – 65535) for the age classes A1 to A5. We show the DN values for the 20% reference set with 18,825 pixel samples.**

We use overall accuracy (OA), F1 score, producer’s accuracy (PA), and user’s accuracy (UA) statistics to assess accuracy of classification. We analyse both individual class and overall F1 score. Because the age map is a much lower resolution (1.5”) than the HST imagery (0.0396”), we do not expect to achieve classification accuracies in the 90% range.

To calculate accuracy statistics, we use a confusion matrix that compares the classified result of the samples and the true state of the samples as known by the reference data. The

diagonal of the confusion matrix represents the correctly classified samples. All other boxes within the confusion matrix represents incorrectly classified samples.

Overall accuracy (OA) is defined as the sum of correctly classified pixels for each class divided by the total number of pixels:

$$OA = \text{sum of correctly classified pixels} / \text{total number of pixels in the test set}$$

(9)

Because the training and testing samples for each age class in Table 3-2 are balanced, we expect OA to be representative of the true accuracy.

Producer's accuracy (PA) and user's accuracy (UA) account for the omission and commission error respectively. These metrics are useful in conjunction with OA as they account for the confusion matrix error that OA overlooks. PA describes the pixels or samples omitted from being correctly classified and is described as:

$$PA = TP / TP + FN$$

(10)

where TP represents the true positives and FN represents the false negatives within the confusion matrix. TP is the correctly classified pixels in the confusion matrix whereas FN are pixels incorrectly labelled as being misclassified.

Error of commission or UA describes the likelihood that the classification result of a specific pixel or sample is representative of its true 'ground' state (Story and Congalton 1986). In our case, the likelihood that the age of a specific pixel is correctly classified. User's accuracy is defined by Eq. (11):

$$UA = TP / TP + FP$$

(11)

FP is the samples that have been classified as a particular class, but are not actually that class in reality.

F1 score calculates an accuracy statistic based on both PA and UA and is defined by Eq. (12) showing the calculation for the individual class F1 score:

$$F1 = 2 * (UA * PA) / (UA + PA)$$

(12)

We also calculate the overall F1 score, which is the average of the individual class F1 score (Goutte and Gaussier 2005).

### 3.6.1 Classification Groups

We perform classification of the general groups shown in Table 3-3. After MDG importance analysis, the most important textures and most important parameters for age classification will be identified. We perform RF and SVM classification for all classification groups. RF and SVM classification is first performed using 80% of the reference set for model training and 20% for accuracy testing. To increase confidence of our model, we perform RF and SVM classifications again using 20% of the reference set for training and 80% for accuracy testing, and then average the resulting accuracies.

**Table 3-4. Chosen classification groups.**

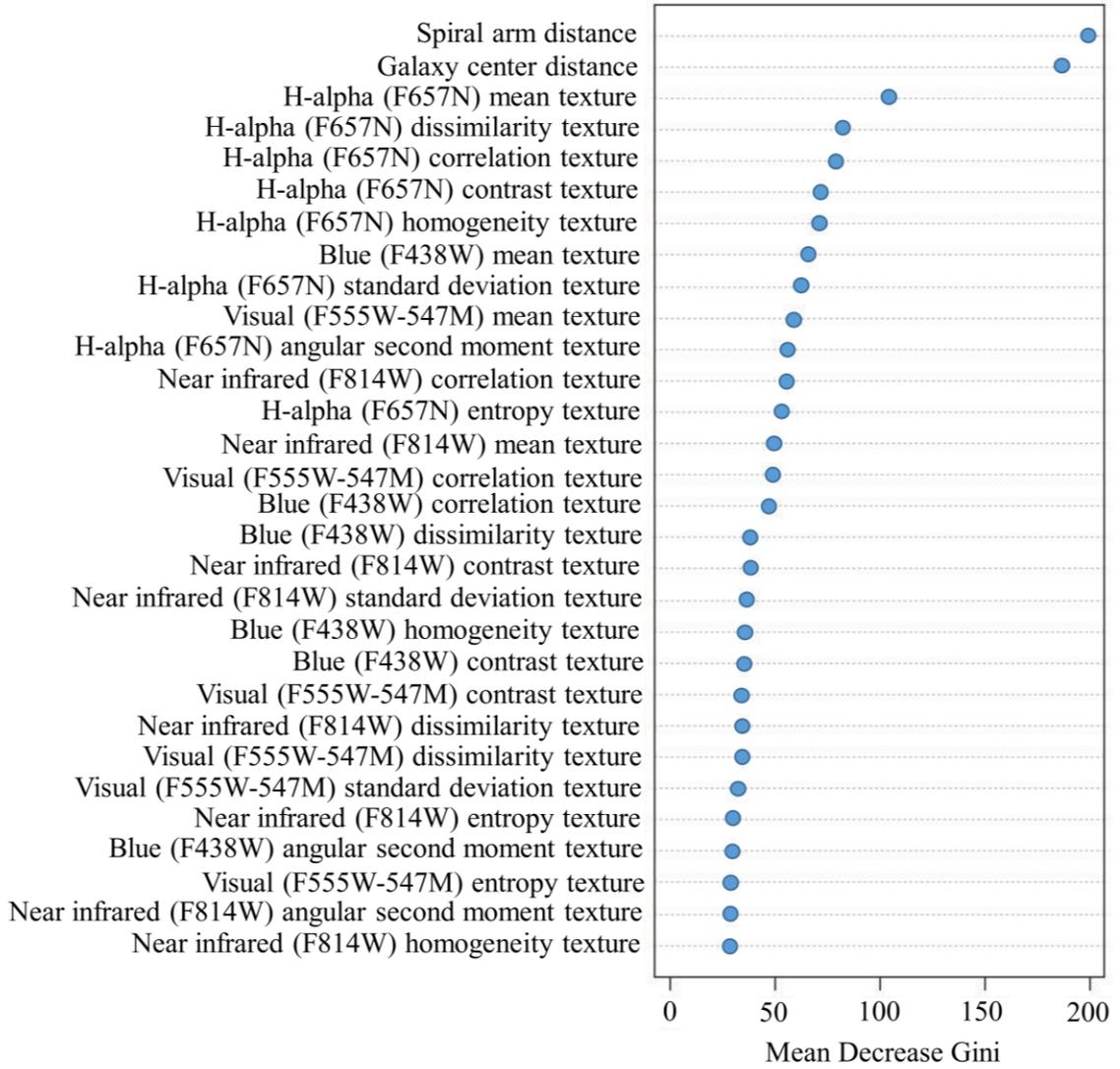
<b>Classification Groups</b>
HST images (F438W, F555W-547M, F657N, and F814W)
HST images (F438W, F555W-547M, F657N, and F814W) + distance from spiral arms and galaxy center
Most important textures
Most important textures + distance from spiral arms and galaxy center
Top MDG parameters

## 3.7 Results

In this section we report on parameter importance, age classification performance, and model performance to answer our research objectives.

### 3.7.1 Parameter Importance

We combine all 38 parameters into a raster stack (4 HST images + 32 texture features + 2 Euclidean distance rasters) and input the raster stack into a random forest model using the R “randomForest” package to derive the mean decrease Gini (MDG) importance measure for all parameters as discussed in Section 3.5.2. Distance from galaxy center and spiral arm parameters are identified as being the most important by the mean decrease Gini (MDG) analysis (Figure 3-10). As the third most important, H-alpha (F657N) mean texture comprises its own group. The remaining H-alpha textural features are also found to be important, but do not make their own distinct group within the MDG plot. From the MDG plot, we derive the classification groups in Table 3-4. The most important texture is identified as mean texture, so we perform several classifications with mean texture parameters: mean texture parameters for all HST images; mean texture for all HST images and distance parameters; H-alpha (F657N) mean texture and distance parameters. We perform classification with the HST images to compare how the textural and distance parameters improve or inhibit classification. Because distance parameters and H-alpha mean texture parameters are found to be outliers in the MDG plot, we also perform classification with these top three MDG parameters. After the top 16 parameters, the MDG importance has little variation between parameters, so we choose to perform classification using the top 16 parameters. Because of the steep slope of MDG of parameters beyond parameter 16, including them would introduce redundant information into the model, and the accuracy would not increase drastically. Lastly, we perform classification of the top five H-alpha (F657N) textures from the MDG plot to test how successfully the textural information from the H-alpha (F657N) band alone is able to classify age.



**Figure 3-11. MDG results for the top 30 parameters.**

**Table 3-5. Chosen classification groups.**

<b>Classification Groups</b>
HST images (F438W, F555W-547M, F657N, and F814W)
HST images (F438W, F555W-547M, F657N, and F814W) + distance from spiral arms and galaxy center
Mean texture for all HST images (F438W, F555W-547M, F657N, and F814W)
Mean texture for all HST images (F438W, F555W-547M, F657N, and F814W) + distance from spiral arms and galaxy center
H-alpha (F657N) textures
MDG top 3 parameters
MDG top 5 H-alpha (F657N) textures
MDG top 16 parameters

We find that the top performing classifications are those that incorporate mean texture, distance from spiral arms, and distance from galaxy center parameters. Classification with mean textures of the four HST imagery bands (F438W, F555W-547M, F657N, and F814W) and both distance parameters results in the highest classification accuracies out of all chosen classification groups. The RF and SVM overall map accuracies (OA) for the mean textures and distance parameters model are 98.63% and 95.09% respectively (average of the OA of classifications trained on 20% of the reference set and 80% of the reference set). Within the confusion matrix shown in Table 3-5, the italicized numbers on the diagonal show the number of correctly classified pixels in each age class, while the bolded numbers represent the overall map accuracy (OA). Columns named UA show the UA statistic for each age class and rows named PA show the PA statistic for each age class. Classification using the top three MDG parameters results in the second highest classification OA of 98.31% and 92.23% for the RF and SVM models. Such a high classification accuracy in both of the aforementioned classifications was unexpected due to the differences in resolution of the age map and Hubble Space Telescope imagery.

**Table 3-6. Confusion Matrix of best performing classification: RF classification model with mean textures (F438W, F555W-F547M, F657N, and F814W) and distance from spiral arm and galaxy center. Top: classification using 80% of training data for model training and 20% for model testing; bottom: classification using 20% of training data for model training and 80% for model testing. The italicized numbers on the diagonal show the number of correctly classified pixels in each age class. The bolded numbers represent the overall map accuracy (OA). Columns named UA are the UA for each age class and rows named PA are the PA for each age class.**

		<i>Age Map Reference Data</i>						
		<b>A1</b>	<b>A2</b>	<b>A3</b>	<b>A4</b>	<b>A5</b>	<b>Total</b>	<b>UA</b>
<b>ML</b> <b>Classification</b>	<b>A1</b>	<i>3618</i>	24	0	6	1	3649	0.9915
	<b>A2</b>	18	<i>3447</i>	0	28	9	3502	0.9843
	<b>A3</b>	1	6	<i>3925</i>	69	6	4007	0.9795
	<b>A4</b>	0	4	58	<i>3696</i>	16	3774	0.9793
	<b>A5</b>	0	0	20	15	<i>3858</i>	3893	0.9910
	<b>Total</b>	3637	3481	4003	3814	3890	18825	
	<b>PA</b>	0.9948	0.9902	0.9805	0.9691	0.9918		<b>0.9851</b>

		<i>Age Map Reference Data</i>						
		<b>A1</b>	<b>A2</b>	<b>A3</b>	<b>A4</b>	<b>A5</b>	<b>Total</b>	<b>UA</b>
<b>ML</b> <b>Classification</b>	<b>A1</b>	<i>14459</i>	44	0	1	5	14509	0.9966
	<b>A2</b>	78	<i>13554</i>	14	29	17	13692	0.9899
	<b>A3</b>	1	42	<i>15827</i>	262	9	16141	0.9805
	<b>A4</b>	3	6	252	<i>15273</i>	99	15633	0.9770
	<b>A5</b>	0	0	59	15	<i>15250</i>	15324	0.9952
	<b>Total</b>	14541	13646	16152	15580	15380	75299	
	<b>PA</b>	0.9944	0.9933	0.9799	0.9803	0.9915		<b>0.9876</b>

The excellent performance of classification models using distance from spiral arms and galaxy center demonstrates their importance for estimating age of stellar populations within galaxies. The addition of distance parameters to the mean texture classification improves accuracy by approximately 24%. Although mean texture is not as important for age estimation according to the MDG plot (Figure 3-10), when used in conjunction with distance parameters it results in the highest accuracies in the 98% range. Mean texture summarizes the average of the pixels in a particular window; in the case of galaxy M83, the window size is 11x11 pixels, meaning it summarizes the average of 121 pixels. Since

the age map pixels represent the flux average of the higher resolution HST pixels, this may explain why mean texture is so efficient at estimating age of the age map reference data.

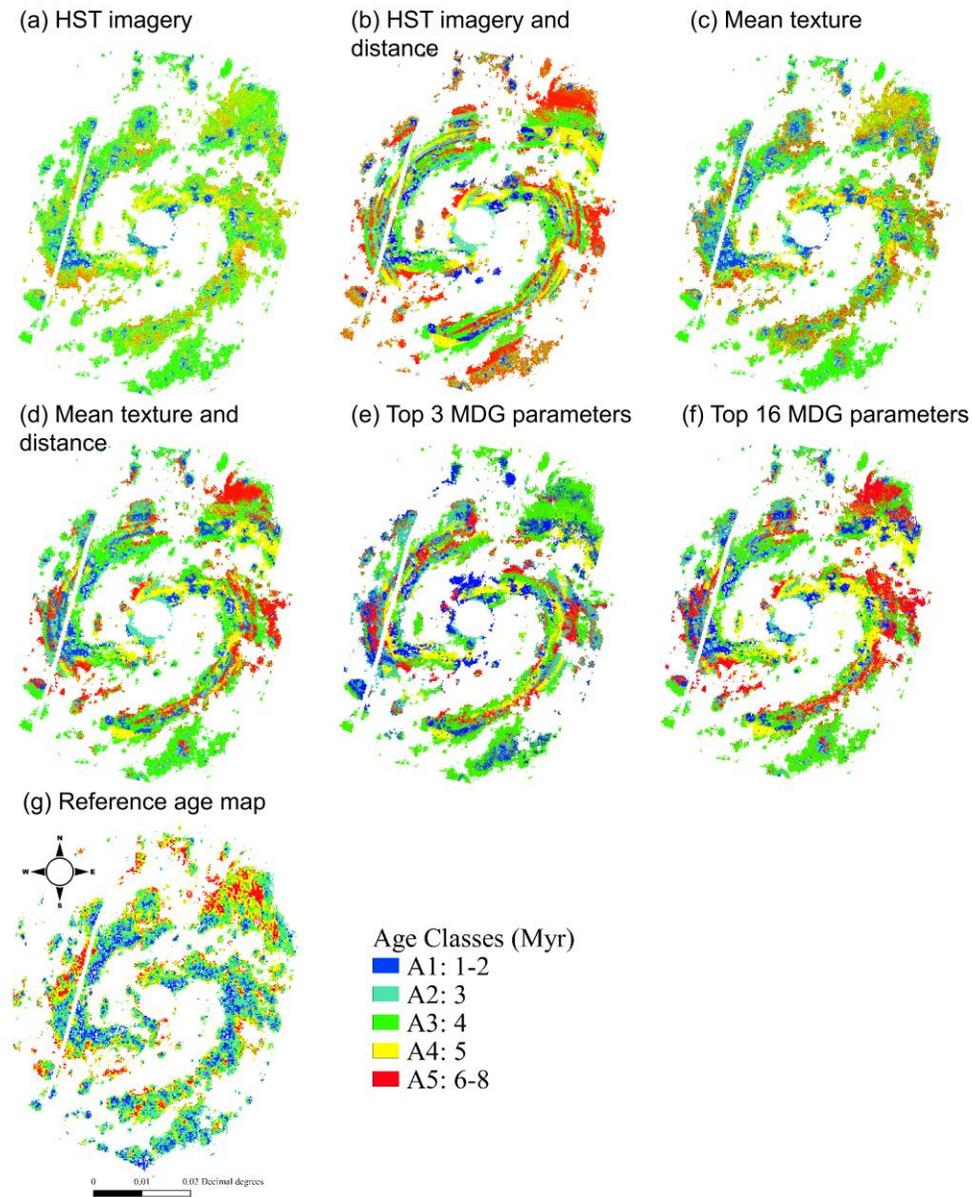
The third most accurate classification, classification of top 16 MDG parameters results in RF classification OA of 97.71% and a SVM classification OA of 93.22%. Because the RF classification of the top 16 MDG parameters (97.71%) is the third highest accuracy result following RF classification of mean textures and distance (98.63%) and MDG top three parameters (98.31%), MDG top 16 parameters is a viable option for age estimation. However, the use fewer parameters optimizes the classification process, so classification with distance and mean texture parameters is more cost effective.

The fourth most accurate classification with HST imagery and distance from spiral arms and galaxy center results in an averaged overall accuracy of 96.08% for RF and 71.44% for SVM. This result is evidence again for a strong relation between the age of stellar populations and the distance from spiral arms and galaxy center within galaxy M83. The disadvantage of this classification method using HST imagery and distance parameters is the inconsistency of the model accuracies, with a difference of about 25% between the RF and SVM models.

RF classifications are shown in Figure 3-11 alongside the age map reference data. We compare the following classification results: mean textures, mean textures and distance parameters, top three MDG parameters, top 16 MDG parameters, and HST imagery and distance parameters. The most notable differences are between the classifications without the use of distance parameters and those with the use of distance parameters. Distance seems to be overemphasized as classifications with distance result in a distinct linear pattern of pixel groupings. However, classifications with distance parameters are able to follow the general pattern of the M83 age map (Sánchez-Gil et al. 2019): younger stars are located on the inner regions (closer to the galaxy center) of the spiral arms and older stars are located on the outer regions (further from the galaxy center) of the spiral arms.

Although the distance parameters result in the highest accuracy, they do not resemble the reference age map as well as classification with mean textures (Figure 3-11). The mean texture classification results in an overall accuracy of 74.78% for RF and 74.62% for SVM.

This accuracy is in between the lowest (48.17%) and highest (98.63%) classification accuracies. The mean texture classification results might be more representative of the young stellar regions than classification with distance parameters is, but is limited by the lower resolution reference age map.



**Figure 3-12. RF classification results are compared to show the differences in classification when using different parameters: a) HST imagery classification b) HST imagery and distance classification; c) mean textures classification; d) mean textures and distance classification; e) top 3 MDG parameters classification; f) top 16 MDG parameters classification; g) reference age map. All maps are classification results using 80% of the reference set for training and 20% of the reference set for testing.**

### 3.7.2 Age Classification Performance

Age classification statistics are reported in Table 3-6. Statistics are averaged over all classifications for each model. For UA, PA, and F1 score statistics, the youngest stellar populations (A1) have consistently high accuracies. This is expected as the youngest regions are typically characterized by the highest H-alpha emission (F657N) whereas the intermediate aged stars (A2, A3, and A4) are similar in appearance to each other and are undergoing the process of H-alpha gas shedding. The oldest stars in class A5 have little to no H-alpha emission and are located further from the center of the spiral arms. Overall, age classes A2, A3, A4, and A5 have similar accuracies, demonstrating their visual similarities.

In Figure 3-11 (Section 3.7.1), we observe that each classification result is able to define class A1 relatively well. For the rest of the classes we can see the opposite: the locations of the age regions are not as definitive. Although distance is identified as being important by the MDG analysis and by the accuracy statistics, the resulting maps lack visual similarity to the age map. Use of distance results in rigid line features and creates an unnatural appearance. From the maps in Figure 3-11, the classification with HST imagery and distance parameters performs the most poorly in terms of both visual appearance and accuracy. Models that include textural features in classification have a higher accuracy than those using HST imagery, so texture (particularly mean texture) is a better determinant of stellar age when used in the machine learning classification model.

**Table 3-7. Age class UA, PA, and F1 score statistics for RF and SVM models.**

**Accuracy statistics are averaged over all eight classifications for each model.**

Model	Age Class Averaged User's Accuracy (UA) in %				
	A1	A2	A3	A4	A5
RF	90.55	75.34	73.43	71.76	77.28
SVM	88.32	68.69	68.24	65.46	71.53
Model	Age Class Averaged Producer's Accuracy (PA) in %				
	A1	A2	A3	A4	A5
RF	85.33	72.26	77.55	76.41	72.45
SVM	82.42	66.88	71.37	70.57	66.80
Model	Age Class Averaged F1 score in %				
	A1	A2	A3	A4	A5
RF	87.75	73.58	75.32	73.67	74.26
SVM	85.15	67.57	69.70	67.71	68.78

### 3.7.3 Model Performance

We find that the RF model outperforms the SVM model for age classification according to the overall map accuracy (OA; Table 3-7). OA and overall map F1 score averages for each classification are in excellent agreement, so we only report OA statistics. The jump in accuracy from classifications with 48% OA to those with 98% OA can be attributed to the addition of distance parameters to the classification models as discussed in Sections 3.7.1 and 3.7.2. For classifications without distance parameters, RF and SVM classification accuracies are comparable. In terms of the individual class accuracies in Table 3-6, the RF model again consistently outperforms the SVM model, meaning that the RF model is a better classifier of age of young stellar populations within galaxy M83.

One disadvantage of the RF model is its overemphasis of the distance layers due to the appearance of artificial line features within the classified raster (Figure 3-11). The best performing classification with mean textures and distance parameters shows the same artificial line pattern for both the RF and SVM model. However, the textural features are better represented by the SVM model, making the age classes appear more natural rather than the rigid lines shown in the RF mean textures and distance map. Because of the differences in resolution between the HST imagery and the age map, we cannot say for certain whether the RF or SVM model is a better classifier of age of young stellar populations.

For classifications with distance parameters, there is a difference of at least 3% OA between RF and SVM models. Classification using both HST imagery and distance has a difference of almost 25% in classification accuracy between the RF and SVM models. Therefore, we do not recommend the use of HST imagery and distance parameters together to perform age classification as the results are inconsistent and may produce unreliable results among different machine learning models. When performing age classification with distance parameters and mean textures calculated for F438W, F555W-F547M, F657N, and F814W bands, the accuracy difference of approximately 3% between the RF and SVM algorithms suggests robustness of age classification among different ML models. Use of other ML and deep learning models should be explored further.

**Table 3-8. All classifications and their overall accuracy (OA) statistics for RF and SVM classification models. The OA reported is average of the classifications trained on 80% and 20% of the total reference set. The top 5 highest accuracies are bolded.**

<b>Classification</b>	<b>RF OA (%)</b>	<b>SVM OA (%)</b>
HST imagery	48.17	48.66
HST imagery + distance	96.08	71.44
Mean textures	74.78	74.62
Mean textures + distance	<b>98.63</b>	<b>95.09</b>
F657N textures	52.25	49.53
MDG top 3 parameters	<b>98.31</b>	92.23
MDG top 5 F657N textures	48.48	47.88
MDG top 16 parameters	<b>97.71</b>	93.22

### 3.8 Discussion

The focus of this research was on the development of a machine learning model that automatically classified age of stellar populations within Hubble Space Telescope imagery (HST). We find that distance parameters are the most important for age estimation and suggest they be used in conjunction with mean textural features. Although the use of distance parameters results in high accuracies, mean texture classification is more visually similar to the reference age map. This may indicate that the lower resolution reference age map is inhibiting the accuracy results of the mean texture classification. It is possible that the mean texture age classification is the most accurate, but a higher resolution reference age map is needed to perform further investigation. Classification accuracies are lower for classifications without distance parameters (48 – 75% overall map accuracy), while classifications including distance parameters are typically above 90% overall map accuracy, with one outlier having 71.44% accuracy. The use of higher resolution reference data could enhance accuracy of our classification.

Use of textural features for high resolution astronomical imagery is promising. However, textural features are not effective when computed from lower resolution imagery. Texture will be more relevant for imagery from the James Webb Space Telescope, Nancy Roman Space Telescope, Euclid mission, and The Cosmological Advanced Survey Telescope for Optical and ultraviolet Research (Gardner et al. 2006; Laureijs et al. 2011; Spergel et al. 2015; Côté et al. 2019). However, determining the optimal window size of each image

needs to be considered. For M83, we performed visual inspection of the imagery and identified star clusters to define the optimal window size for defining the spatially related pixels. Manual inspection of many images takes time so developing a method to automate this step would enhance the applicability of textural features. A semivariogram analysis of images could be used to determine the window size that best represents the pixel neighbourhoods (Hall-Beyer 2017). When studying galaxies at similar distances and with imagery from the same telescope or a telescope with similar resolution, the same window size can be used.

Overall, more research needs to be conducted to determine the full potential of textural analysis for age classification. One problem mentioned in the results is the lack of a high resolution reference age map to compare to the mean texture result. Our classification with mean textures might prove to be quite accurate for age classification of young stellar populations when compared to an age map of higher resolution. Different textures might be more or less useful for different types of data, so testing of many textures as done in this paper is useful for identifying the most useful textures.

The reference data age map (Sánchez-Gil et al. 2019) estimates age of stellar populations using the relationship between the far ultraviolet and H-alpha emission. For galaxy M83, high resolution HST imagery in far ultraviolet (FUV) wavelengths is not available for the full galaxy extent. Having FUV wavelength data may increase the robustness of the model and its fidelity to the reference data. The Cosmological Advanced Survey Telescope for Optical and ultraviolet Research (CASTOR), proposed by the Canadian Space Agency, would capture high resolution imagery in the FUV wavelengths. CASTOR's resolution is similar to that of HST (Côté et al. 2019), so loss of pixel information when used in conjunction with other telescopes such as HST would be minimal. Our methods are also limited by the range of age estimation in the reference data, as it does not include older stellar populations. To have a complete understanding of galaxy mechanisms, the fine age estimation as seen in Sánchez-Gil et al. (2019) should be expanded to stellar age in billions of years.

Machine learning (ML) models used in this study are limited by their ability to handle different types of data. Uncertainty is present in all astronomical data, and the random forest (RF) model is inhibited by its inability to account for this. RF models that can handle uncertainty, for example Reis et al. (2019), are not yet widely available in GIS programs and packages. Reis et al. (2019) develop the Probabilistic Random Forest model that determines a probability distribution function for the classes within the reference set, improving accuracy of the original RF model by up to 30%.

A pitfall of the SVM model is its failure to support outlier samples (Foody and Mathur 2006). For our training data, outlier samples are likely present due to the pixel variation within astronomical imagery and within each age class. These may have lowered the accuracy of the SVM model. A second limitation of the SVM model is its poor handling of overlapping classes (Fu et al. 2015). Many of our training samples for each class overlapped, so the SVM algorithm has difficulty defining the hyperplanes. This can explain why the RF model outperforms the SVM model for some classifications. A solution to this could be the use of a Two-Step Classification SVM (TSC-SVM) algorithm developed by Fu et al. (2015) that improves how the original SVM model handles overlapping data by emphasizing the small differences between classification groups. However, the TSC-SVM model is not available in widely used GIS programs and packages, which are used in our methods presented in this study. Since the RF model results in unusually high accuracies and creates unnatural patterns, we suggest further testing of the SVM model for stellar age classification.

The tracing of the spiral arms follows methods done by Shabani et al. (2018) where we follow the prominent dust lanes within the spiral arms of M83. For use of our machine learning model to automate age classification of high resolution galaxy imagery, the model would need to be improved through the use of spiral arm fitting programs that are being frequently developed (Davis and Hayes 2014; Bekki 2021). For our model, the dust lanes are a good approximation of the spiral arm locations, so can be manually traced for a single galaxy. We recommend the use of an automated spiral arm fitting model for application on many galaxies. Further, future work can help to decrease uncertainty due to arm center distance by improving identification of true center.

Age classification can contribute to better understanding the formation and evolution of galaxies and the universe. Age estimation is a well-studied topic within the astronomical community, with many reliable SED fitting models available. Opportunities for improvement include developing new models that increase both the pixel and age resolution.

### 3.9 Conclusions

In this paper, we develop a machine learning (ML) model to classify age of regions within spiral galaxy M83. We use high resolution Hubble Space Telescope (HST) optical imagery in blue (F475W), visual (F555W-F547M), H-alpha (F657N), and near infrared (F814W) emission. Texture features calculated from the HST imagery and Euclidean distance calculated from spiral arms and galaxy center features are also tested for age classification. Random forest (RF) and support vector machine (SVM) supervised ML classification models are compared. We train ML models using a reference age map of M83 (Sánchez-Gil et al. 2019) that identifies stellar ages of approximately 1 – 8 million years and identify five age classes: A1 (1-2 Myr), A2 (3 Myr), A3 (4 Myr), A4 (5 Myr), and A5 (6-8 Myr). The RF model outperforms the SVM model for all classifications according to the overall map accuracies. Age classes A1 and A5 have the highest individual class accuracies. We find that Euclidean distance from spiral arms and galaxy center parameters are crucial for obtaining high classification accuracies in the 90% range. This finding supports the notion of a spiral density wave age gradient within galaxy M83. Classifications with Euclidean distance parameters can replicate the age map patterns of younger stars dominating the stellar population on the inner portion of the spiral arms and of older stars dominating the stellar population on the outer portion of the spiral arms. To optimize classification with distance parameters, mean textures calculated on the HST imagery should be used in conjunction with Euclidean distance parameters. Classification with distance parameters and mean textures results in an overall map accuracy of 98.63%. Although the use of distance parameters increases accuracy, the resulting classified maps do not share strong visual similarities to the reference age map. Classification with mean textures results in a lower overall accuracy of 71.44%, and a map that appears quite similar to the age map. We recommend further testing of textural analysis for age classification. Our results contribute

to the development of an age classification model that requires less imagery and reduces processing time, while maintaining a high pixel resolution.

### 3.10 References

- Abdeen, S., Davis, B. L., Eufrazio, R., Kennefick, D., Kennefick, J., Miller, R., Shields, D., Monson, E. B., Bassett, C., & O'Mara, H. (2022). Evidence in favour of density wave theory through age gradients observed in star formation history maps and spatially resolved stellar clusters. *Monthly Notices of the Royal Astronomical Society*, *512*(1), 366–377. <https://doi.org/10.1093/mnras/stac459>
- Bekki, K. (2021). Quantifying the fine structures of disk galaxies with deep learning: Segmentation of spiral arms in different Hubble types. *Astronomy & Astrophysics*, *647*, A120. <https://doi.org/10.1051/0004-6361/202039797>
- Belgiu, M., & Drăguț, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, *114*, 24–31. <https://doi.org/10.1016/j.isprsjprs.2016.01.011>
- Blair, W. P., Chandar, R., Dopita, M. A., Ghavamian, P., Hammer, D., Kuntz, K. D., Long, K. S., Soria, R., Whitmore, B. C., & Frank Winkler, P. (2014). An expanded HST /WFC3 survey of M83: Project overview and targeted supernova remnant search. *The Astrophysical Journal*, *788*(1), 55. <https://doi.org/10.1088/0004-637X/788/1/55>
- Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, *20*(3), 273–297. <https://doi.org/10.1007/BF00994018>
- Côté, P., Abraham, B., Balogh, M., Capak, P., Carlberg, R., Cowan, N., Djazovski, O., Drissen, L., Drout, M., Dupuis, J., Evans, C., Fantin, N., Ferrarese, L., Fraser, W., Gallagher, S., Girard, T., Gleisinger, R., Grandmont, F., Hall, P., ... Venn, K. (2019). *CASTOR: A Flagship Canadian Space Telescope*. Zenodo. <https://doi.org/10.5281/ZENODO.3758463>

- Davis, D. R., & Hayes, W. B. (2014). SpArcFiRe: Scalable automated detection of spiral galaxy arm segments. *The Astrophysical Journal*, 790(2), 87.  
<https://doi.org/10.1088/0004-637X/790/2/87>
- Dressel, L. (2022). *Wide Field Camera 3 Instrument Handbook* (Version 14.0). Baltimore: STScI. <<https://hst-docs.stsci.edu/wfc3ihb>>.
- ESA/ESO/NASA. (2021). *FITS Liberator 3*.  
 <[https://esahubble.org/projects/fits\\_liberator/download\\_v30/](https://esahubble.org/projects/fits_liberator/download_v30/)>.
- Farhadian, R., & Clarke, B. R. (2022). A note on the Helmert transformation. *Communications in Statistics - Theory and Methods*, 51(15), 5258–5264.  
<https://doi.org/10.1080/03610926.2020.1836223>
- Foody, G. M., & Mathur, A. (2006). The use of small training sets containing mixed pixels for accurate hard image classification: Training on mixed spectral responses for classification by a SVM. *Remote Sensing of Environment*, 103(2), 179–189. <https://doi.org/10.1016/j.rse.2006.04.001>
- Fu, M., Tian, Y., & Wu, F. (2015). Step-wise support vector machines for classification of overlapping samples. *Neurocomputing*, 155, 159–166.  
<https://doi.org/10.1016/j.neucom.2014.12.035>
- Gardner, J. P., Mather, J. C., Clampin, M., Doyon, R., Greenhouse, M. A., Hammel, H. B., Hutchings, J. B., Jakobsen, P., Lilly, S. J., Long, K. S., Lunine, J. I., Mccaughrean, M. J., Mountain, M., Nella, J., Rieke, G. H., Rieke, M. J., Rix, H.-W., Smith, E. P., Sonneborn, G., ... Wright, G. S. (2006). The James Webb Space Telescope. *Space Science Reviews*, 123(4), 485–606.  
<https://doi.org/10.1007/s11214-006-8315-7>
- Goutte, C., & Gaussier, E. (2005). A probabilistic interpretation of Precision, Recall and F-Score, with implication for evaluation. In D. E. Losada & J. M. Fernández-Luna (Eds.), *Advances in Information Retrieval*, 3408, 345–359. Springer Berlin Heidelberg. [https://doi.org/10.1007/978-3-540-31865-1\\_25](https://doi.org/10.1007/978-3-540-31865-1_25)

- Hall-Beyer, M. (2017). Practical guidelines for choosing GLCM textures to use in landscape classification tasks over a range of moderate spatial scales. *International Journal of Remote Sensing*, 38(5), 1312–1338. <https://doi.org/10.1080/01431161.2016.1278314>
- Haralick, R. M., Shanmugam, K., & Dinstein, I. (1973). Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics, SMC-3*(6), 610–621. <https://doi.org/10.1109/TSMC.1973.4309314>
- Hunter, D. A., Elmegreen, B. G., Rubin, V. C., Ashburn, A., Wright, T., Józsa, G. I. G., & Struve, C. (2013). Star formation in two luminous spiral galaxies. *The Astronomical Journal*, 146(4), 92. <https://doi.org/10.1088/0004-6256/146/4/92>
- Jensen, E. B., Talbot, R. J., Jr., & Dufour, R. J. (1981). M83. III - Age and brightness of young and old stellar populations. *The Astrophysical Journal*, 243, 716. <https://doi.org/10.1086/158640>
- Joye, W. A., & Mandel, E. (2003). New features of SAOImage DS9. *Astronomical Data Analysis Software and Systems XII ASP Conference Series*, 295, 489–492.
- Kiar, A. K., Barmby, P., & Hidalgo, A. (2017). Deconstructing a galaxy: Colour distributions of point sources in Messier 83. *Monthly Notices of the Royal Astronomical Society*, 472(1), 1074–1087. <https://doi.org/10.1093/mnras/stx2037>
- Koo, B.-K., Baek, J.-W., & Chung, K.-Y. (2021). Weight feedback-based harmonic MDG-Ensemble model for prediction of traffic accident severity. *Applied Sciences*, 11(11), 5072. <https://doi.org/10.3390/app11115072>
- Kwik, R. J., Wang, J., Barmby, P., & Holwerda, B. W. (2022). Galactic component mapping of galaxy UGC 2885 by machine learning classification. *Advances in Space Research*, 70(1), 229–247. <https://doi.org/10.1016/j.asr.2022.04.032>
- Laureijs, R., Amiaux, J., Arduini, S., Auguères, J.-L., Brinchmann, J., Cole, R., Cropper, M., Dabin, C., Duvet, L., Ealet, A., Garilli, B., Gondoin, P., Guzzo, L., Hoar, J., Hoekstra, H., Holmes, R., Kitching, T., Maciaszek, T., Mellier, Y., ... Zucca, E.

(2011). *Euclid Definition Study Report*. ArXiv:1110.3193 [Astro-Ph].  
<http://arxiv.org/abs/1110.3193>

Lawrence, R. L., Wood, S. D., & Sheley, R. L. (2006). Mapping invasive plants using hyperspectral imagery and Breiman Cutler classifications (randomForest). *Remote Sensing of Environment*, 100(3), 356–362.  
<https://doi.org/10.1016/j.rse.2005.10.014>

Lee, J.H., Kim, S.C., Park, H.S., Ree, C.H., Kyeong, J., Chung, J. (2011). Hubble Space Telescope pixel analysis of the interacting face-on spiral galaxy NGC 5194 (M51A). *The Astrophysical Journal*, 740, 42.  
<https://doi.org/10.1088/0004-637X/740/1/42>

Lee, J. H., Kim, S. C., Ree, C. H., Kim, M., Jeong, H., Lee, J. C., & Kyeong, J. (2012). HUBBLE SPACE TELESCOPE PIXEL ANALYSIS OF THE INTERACTING S0 GALAXY NGC 5195 (M51B). *The Astrophysical Journal*, 754(2), 80.  
<https://doi.org/10.1088/0004-637X/754/2/80>

Lin, C. C., & Shu, F. H. (1964). On the Spiral Structure of Disk Galaxies. *The Astrophysical Journal*, 140, 646. <https://doi.org/10.1086/147955>

QGIS 3.16. (2021). *Geographic Information System Developers Manual*.  
[docs.qgis.org/3.16/en/docs/developers\\_guide/index.html](https://docs.qgis.org/3.16/en/docs/developers_guide/index.html).

Rajendran, G. B., Kumarasamy, U. M., Zarro, C., Divakarachari, P. B., & Ullo, S. L. (2020). Land-use and land-cover classification using a human group-based particle swarm optimization algorithm with an LSTM classifier on hybrid pre-processing remote-sensing images. *Remote Sensing*, 12(24), 4135.  
<https://doi.org/10.3390/rs12244135>

Reis, I., Baron, D., & Shahaf, S. (2019). Probabilistic Random Forest: A machine learning algorithm for noisy data sets. *The Astronomical Journal*, 157(1), 16.  
<https://doi.org/10.3847/1538-3881/aaf101>

- RStudio Team. (2020). *RStudio: Integrated Development for R*; RStudio: Boston, MA, USA. [www.rstudio.com/](http://www.rstudio.com/)
- Sánchez-Gil, M. C., Alfaro, E. J., Cerviño, M., Pérez, E., Bland-Hawthorn, J., & Jones, D. H. (2019). Hierarchical Bayesian approach for estimating physical properties in nearby galaxies: Age Maps (Paper II). *Monthly Notices of the Royal Astronomical Society*, 483(2), 2641–2670. <https://doi.org/10.1093/mnras/sty3106>
- Sen, S., Agarwal, S., Chakraborty, P., & Singh, K. P. (2022). Astronomical big data processing using machine learning: A comprehensive review. *Experimental Astronomy*, 53(1), 1–43. <https://doi.org/10.1007/s10686-021-09827-4>
- Shabani, F., Grebel, E. K., Pasquali, A., D’Onghia, E., Gallagher, J. S., Adamo, A., Messa, M., Elmegreen, B. G., Dobbs, C., Gouliermis, D. A., Calzetti, D., Grasha, K., Elmegreen, D. M., Cignoni, M., Dale, D. A., Aloisi, A., Smith, L. J., Tosi, M., Thilker, D. A., ... Pellerin, A. (2018). Search for star cluster age gradients across spiral arms of three LEGUS disc galaxies. *Monthly Notices of the Royal Astronomical Society*, 478(3), 3590–3604. <https://doi.org/10.1093/mnras/sty1277>
- Song, Z., Zhou, S., & Guan, J. (2014). A novel image registration algorithm for remote sensing under affine transformation. *IEEE Transactions on Geoscience and Remote Sensing*, 52(8), 4895–4912. <https://doi.org/10.1109/TGRS.2013.2285814>
- Spergel, D., Gehrels, N., Baltay, C., Bennett, D., Breckinridge, J., Donahue, M., Dressler, A., Gaudi, B. S., Greene, T., Guyon, O., Hirata, C., Kalirai, J., Kasdin, N. J., Macintosh, B., Moos, W., Perlmutter, S., Postman, M., Rauscher, B., Rhodes, J., ... Zhao, F. (2015). *Wide-Field InfrarRed Survey Telescope-Astrophysics Focused Telescope Assets WFIRST-AFTA 2015 Report*. ArXiv:1503.03757 [Astro-Ph]. <http://arxiv.org/abs/1503.03757>
- Story, M., & Congalton, R. G. (1986). Accuracy assessment: A user’s perspective. *Photogrammetric Engineering and Remote Sensing*, 52(3), 397–399.

- Walcher, J., Groves, B., Budavári, T., & Dale, D. (2011). Fitting the integrated spectral energy distributions of galaxies. *Astrophysics and Space Science*, *331*(1), 1–51. <https://doi.org/10.1007/s10509-010-0458-z>
- Wei, P., Zou, H., Lin, L., Zhou, X., Liu, X., Kong, X., Ma, L., & Ma, S.-G. (2021). Spatially-resolved stellar population properties of the M51–NGC 5195 system from multi-wavelength photometric data. *Research in Astronomy and Astrophysics*, *21*(1), 006. <https://doi.org/10.1088/1674-4527/21/1/6>
- Whitmore, B. C., Lee, J. C., Chandar, R., Thilker, D. A., Hannon, S., Wei, W., Huerta, E. A., Bigiel, F., Boquien, M., Chevance, M., Dale, D. A., Deger, S., Grasha, K., Klessen, R. S., Kruijssen, J. M. D., Larson, K. L., Mok, A., Rosolowsky, E., Schinnerer, E., ... Williams, T. (2021). Star cluster classification in the PHANGS-HST survey: Comparison between human and machine learning approaches. *Monthly Notices of the Royal Astronomical Society*, *506*(4), 5294–5317. <https://doi.org/10.1093/mnras/stab2087>
- Yang, B., Zhu, W., Rezaei, E. E., Li, J., Sun, Z., & Zhang, J. (2022). The optimal phenological phase of maize for yield prediction with high-frequency UAV remote sensing. *Remote Sensing*, *14*(7), 1559. <https://doi.org/10.3390/rs14071559>
- Zhou, Y., Yu, Q., Hua, S., Yin, W., & Li, Y. (2013). An automatic global-to-local image registration based on SIFT and thin-plate spline (TPS). *2013 IEEE International Geoscience and Remote Sensing Symposium - IGARSS*, 2535–2538. <https://doi.org/10.1109/IGARSS.2013.6723338>

## Chapter 4

### 4 Conclusion

#### 4.1 Summary

Studying components within galaxies can provide important clues to the formation and evolution of the universe, the Milky Way, and our Solar System. Distributions of galaxy components and stars of different ages are direct indicators of the overall age of the galaxy. Older galaxies have low rates of star formation and a large population of old stars, while younger galaxies have higher rates of star formation and many young and blue stars. Younger galaxies also have more dust and gas within them that fuels star formation. Because we cannot resolve galaxy components in very distant galaxies, we can study nearby galaxies and make inferences about the more distant ones. Over time, components in distant galaxies will be better resolved through advances in telescope capabilities. Machine learning methods can contribute to our understanding of galaxy mechanics while allowing for the rapid processing of data. Because remote sensing and astronomy are compatible fields of research, in this Thesis we apply remote sensing image processing methods to digital imagery of galaxies.

In this Thesis, we research two nearby spiral galaxies. Chapter 2 describes the development of maximum likelihood classifier (MLC), random forest (RF), and support vector machine (SVM) supervised machine learning models trained to classify each pixel within HST imagery of UGC 2885 into one of five galaxy component classes and one class representing the celestial background behind the galaxy: young stars, old stars, dust lanes, galaxy center, outer disc, and celestial background. The model is trained on different combinations of parameters: Hubble Space Telescope (HST) imagery, band ratios calculated from the HST imagery, textural features calculated from the HST imagery, and Euclidean distance calculated from the spiral arm and galaxy center features within UGC 2885. Distance and textural features are found to have the highest accuracy of galaxy component classification.

Chapter 3 discusses the development of RF and SVM supervised learning models that are trained to classify stellar age within galaxy M83 using an age map by Sánchez-Gil et al.

(2019) as reference data. The machine learning models are trained using HST imagery, textural features calculated from the HST imagery, and Euclidean distance from the spiral arms and galaxy center features within M83. The training data is manually selected from the age map and we identify five age classes: A1 (1 to 2 Myr), A2 (3 Myr), A3 (4 Myr), A4 (5 Myr), and A5 (6 to 8 Myr). Use of distance parameters results in the highest accuracies but exhibits artificial line features throughout the classified raster. Textural features have a lower accuracy but a stronger resemblance to the age map.

## 4.2 Conclusions

The research objectives chosen for this study were achieved. Below are the corresponding findings:

- (i) Machine learning is successful at classification of galaxy components and stellar age within nearby spiral galaxies. In Chapter 2, MLC, RF, and SVM models are developed to classify galaxy components within UGC 2885. The SVM model is found to be the most successful for galaxy component classification. The MLC model is consistently outperformed by the RF and SVM models, so is not as useful for galaxy component classification. The SVM model achieves an average overall accuracy (OA) of 84.9% whereas the MLC and RF models achieve an OA of 80.5% and 82.6% respectively. Average F1 score statistics are 80.4%, 80.9%, and 82.6% for MLC, RF, and SVM. Chapter 3 focused on classification of stellar age in M83 using RF and SVM models. Results indicate that the RF model is a better choice for stellar age classification. The RF model achieves an OA of 98.63% for the highest performing classification, where SVM achieves 95.09% OA. F1 score accuracies for the same classification are 98.65% and 95.12% for RF and SVM respectively, demonstrating that F1 score is in agreement with OA. For a majority of the classifications, RF outperforms the SVM models.

In Chapter 3, we make use of an age map produced by Sánchez-Gil et al. (2019) as a reference data for labeled data creation. The reference data is successful for producing high accuracy of classification, however, the highest accuracy classifications lack resemblance to the age map. The use of a higher resolution age

map for reference data would benefit our model and increase confidence of age classification.

- (ii) Chapters 2 and 3 have similar findings, indicating that Euclidean distance and textural features are the most useful parameters for classification of galaxy components and stellar age respectively. The combination of Euclidean distance, mean texture parameters for all three HST images, and HST visual image (F606W) results in the highest OA of 95.4% for both RF and SVM models, and 94.3% for MLC. For the same classification, F1 score statistics for MLC, RF, and SVM models are 92.9%, 94.4%, and 94.6% respectively. In Chapter 2, we found band ratios to be the least useful for galaxy component classification. In Chapter 3, from the classifications performed, the HST imagery and Hydrogen-alpha (F657N) texture features are found to be least useful for stellar age estimation. Using distance parameters in conjunction with HST imagery and texture features increases accuracy by up to 50%. The classification with the poorest performance is with HST imagery, having an OA of 48.17% for RF and 48.66% for SVM.
- (iii) Chapter 2 found the galaxy center and celestial background were the best classified. The galaxy center had high average individual class accuracies: 87% user's accuracy, 100% producer's accuracy, and 93% F1 score. The celestial background had classification accuracy but is not a part of the galaxy. The outer disc was also well classified. The dust lanes and old stellar populations are similar in appearance, so the ML models had difficulty differentiating them. Chapter 3 identified the youngest ages of 1 to 2 Myr (class A1) as being the best classified. This is expected due to their unique H-alpha emission. The remainder of the classes, A2 to A5, are similar in appearance so have lower classification accuracy. Addition of Euclidean distance parameters improves classification accuracy for all classes.

### 4.3 Limitations and Future Work

Study of celestial phenomena presents limitations. Galaxies are very distant and the only method of studying them is through passive sensing, making in-situ ground truth infeasible. Chapter 2 classifies galaxy components by analysis of galaxy component emission in

different wavelength bands. In Chapter 3, we make use of a reference age map of M83 (Sánchez-Gil et al., 2019). The age map resolution is lower than that of the Hubble Space Telescope imagery, so the use of a higher resolution age map would benefit our model. Lack of ground truthing makes training site selection difficult. When using the age map for interpretation, different age classes appear visually similar, so there is some uncertainty present with training site creation. A higher resolution of age classification would also be beneficial for understanding galaxy mechanics of formation and evolution. In Chapter 3, only young stars aged 1 to 8 million years are classified. Older stellar populations above 8 million years, such as those dominating the galaxy center and the inter-arm regions, are not considered in analysis. Along with improvement of reference data and age resolution, our models can be tested on digital imagery of galaxies with different physical properties. Spiral galaxies are all unique, so training on one galaxy and testing on many others would benefit the understanding of distributions of galaxy components and stellar ages. Applying our methods to other types of galaxies such as irregular and elliptical would also aid in understanding the differences in galaxy component formation and evolution between different galaxy types.

Astronomical data such as the Hubble Space Telescope imagery has uncertainty that comes in two forms: statistical and systematic. The statistical uncertainty is introduced through the presence of noise in the data. For instance, if there is a weak signal-to-noise ratio there are a limited number of photons being received and a high amount of noise. Systematic uncertainty occurs due to imperfect calibration of the telescope imagery. Our models from both Chapter 2 and 3 do not account for such uncertainties. The incorporation of machine learning models that are capable of handling such uncertainty (e.g., Reis et al., 2019) would benefit the results and confidence of our models. However, models capable of accounting for uncertainty are not available in widely used Geographic Information Systems or in programming packages. Another area of improvement to our models is the automation of the identification of spiral arm and galaxy center features for Euclidean distance calculation. In both Chapters 2 and 3 we manually trace spiral arms. Chapter 2 uses a piecewise fitting model while in Chapter 3 we trace the dust lanes as good approximations of the spiral arms. Models that automate the identification of spiral arms are in development (Bekki, 2021), and the addition of these models into our own would allow for more rapid

machine learning classification and testing of the model on images of other galaxies. Moreover, the development of models that better define the true center of the spiral arms and therefore decrease uncertainty of fitting is an important opportunity for future work.

Further research should be done to determine the full potential for textural features in processing of astronomical imagery on a pixel-basis. In both Chapters 2 and 3 we found textural features to be useful for classification. Elliptical galaxies are smooth in appearance and have a large population of old stars, so texture may not be as useful. However, irregular galaxies have many galaxy components and stars of different ages, so textural analysis may prove useful. Texture performs best when used on high-resolution imagery, so textural analysis is limited to digital imagery from telescopes with similar resolution such as the James Webb Space Telescope (Gardner et al., 2006). As more high-resolution imagery becomes available, textural analysis will become more relevant. For texture analysis and our machine learning models, testing on different wavelength bands should also be done. Our methods use visible and near-infrared wavelength bands; we do not consider the ultraviolet and infrared wavelengths, which may contribute useful galaxy component and stellar age information.

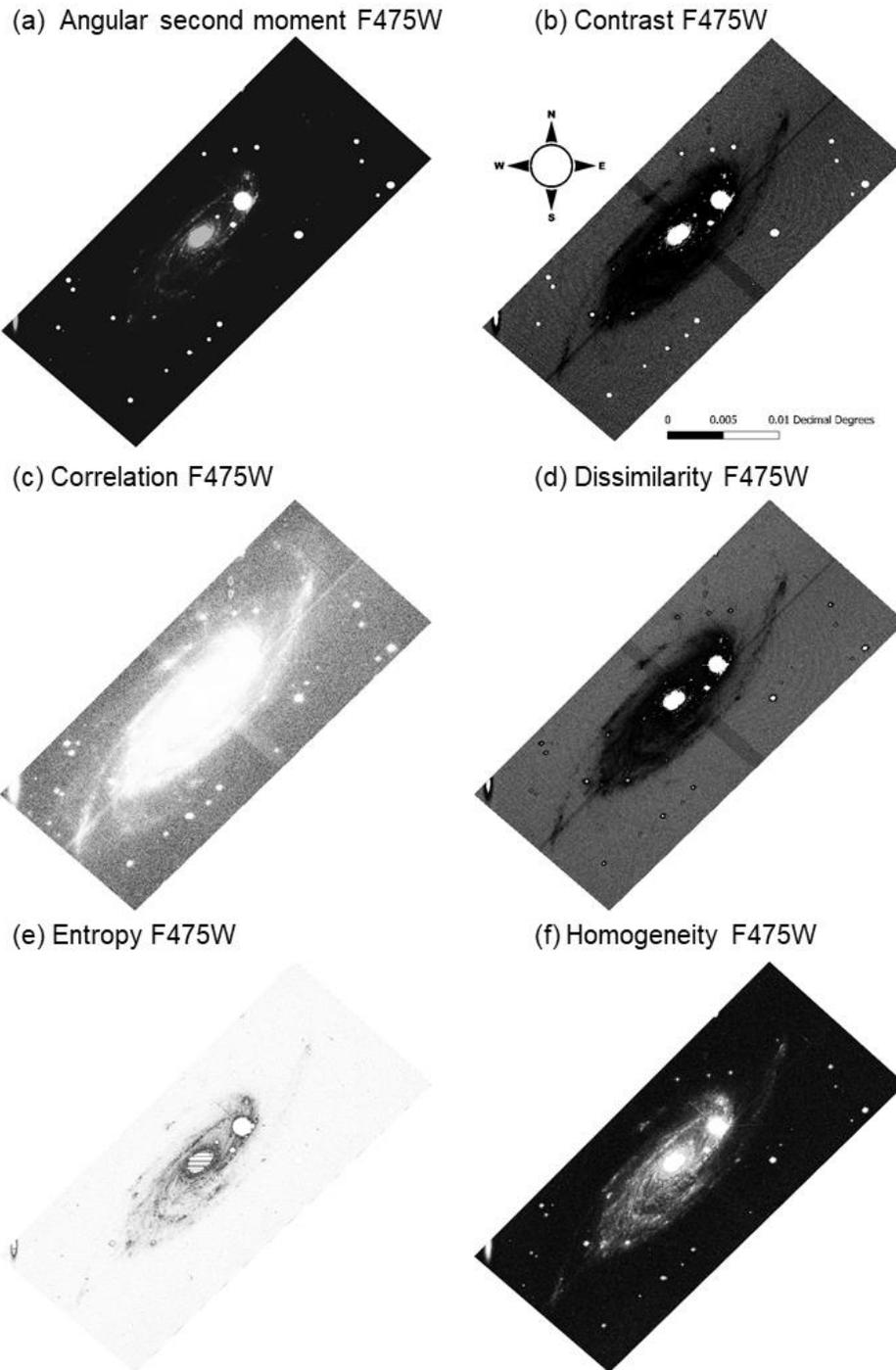
Our work contributes to automated classification of digital imagery of spiral galaxies. Classifying galaxy components and stellar populations of different ages helps to better understand their distributions within galaxies having different visual and physical properties. Overall, our machine learning models should be tested on imagery from different telescopes, in different wavelengths, and on imagery of different types of celestial phenomena.

## 4.4 References

- Bekki, K. (2021). Quantifying the fine structures of disk galaxies with deep learning: Segmentation of spiral arms in different Hubble types. *Astronomy & Astrophysics*, 647, A120. <https://doi.org/10.1051/0004-6361/202039797>
- Gardner, J. P., Mather, J. C., Clampin, M., Doyon, R., Greenhouse, M. A., Hammel, H. B., Hutchings, J. B., Jakobsen, P., Lilly, S. J., Long, K. S., Lunine, J. I., Mccaughrean, M. J., Mountain, M., Nella, J., Rieke, G. H., Rieke, M. J., Rix, H.-W., Smith, E. P., Sonneborn, G., ... Wright, G. S. (2006). The James Webb Space Telescope. *Space Science Reviews*, 123(4), 485–606. <https://doi.org/10.1007/s11214-006-8315-7>
- Reis, I., Baron, D., & Shahaf, S. (2019). Probabilistic Random Forest: A machine learning algorithm for noisy data sets. *The Astronomical Journal*, 157(1), 16. <https://doi.org/10.3847/1538-3881/aaf101>
- Sánchez-Gil, M. C., Alfaro, E. J., Cerviño, M., Pérez, E., Bland-Hawthorn, J., & Jones, D. H. (2019). Hierarchical Bayesian approach for estimating physical properties in nearby galaxies: Age Maps (Paper II). *Monthly Notices of the Royal Astronomical Society*, 483(2), 2641–2670. <https://doi.org/10.1093/mnras/sty3106>

## 5 Appendices

### Appendix A: Textures

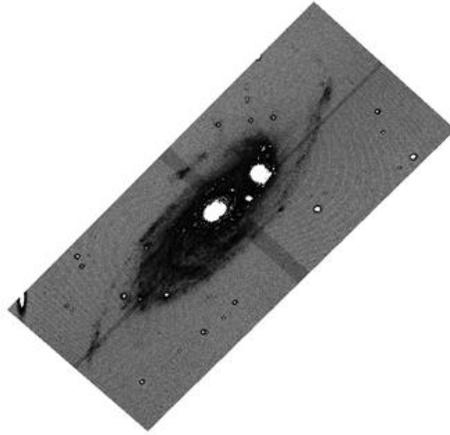


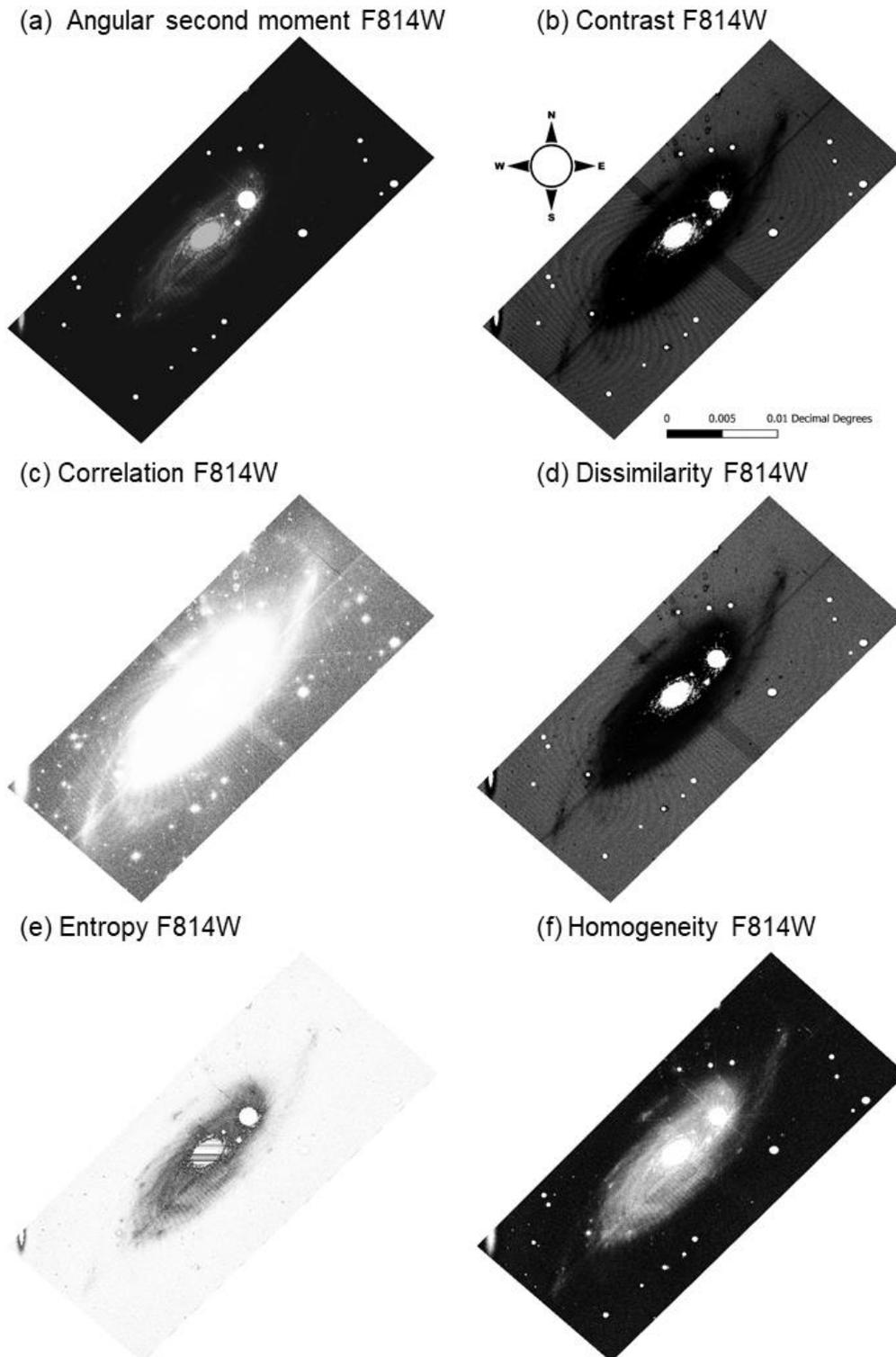
**Figure A 1. Haralick textures calculated on the blue-green HST image (F475W) of UGC 2885.**

(g) Mean F475W



(h) Standard deviation F475W

**Figure A 1. Continued.**

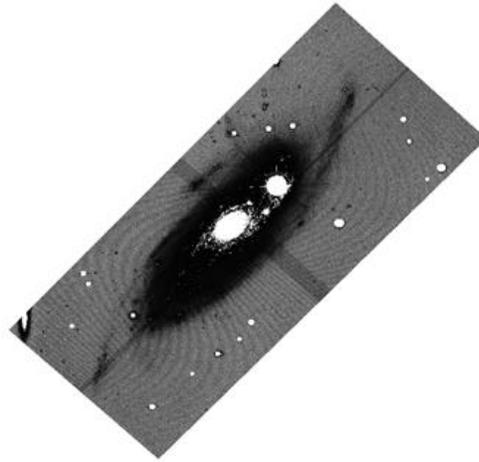


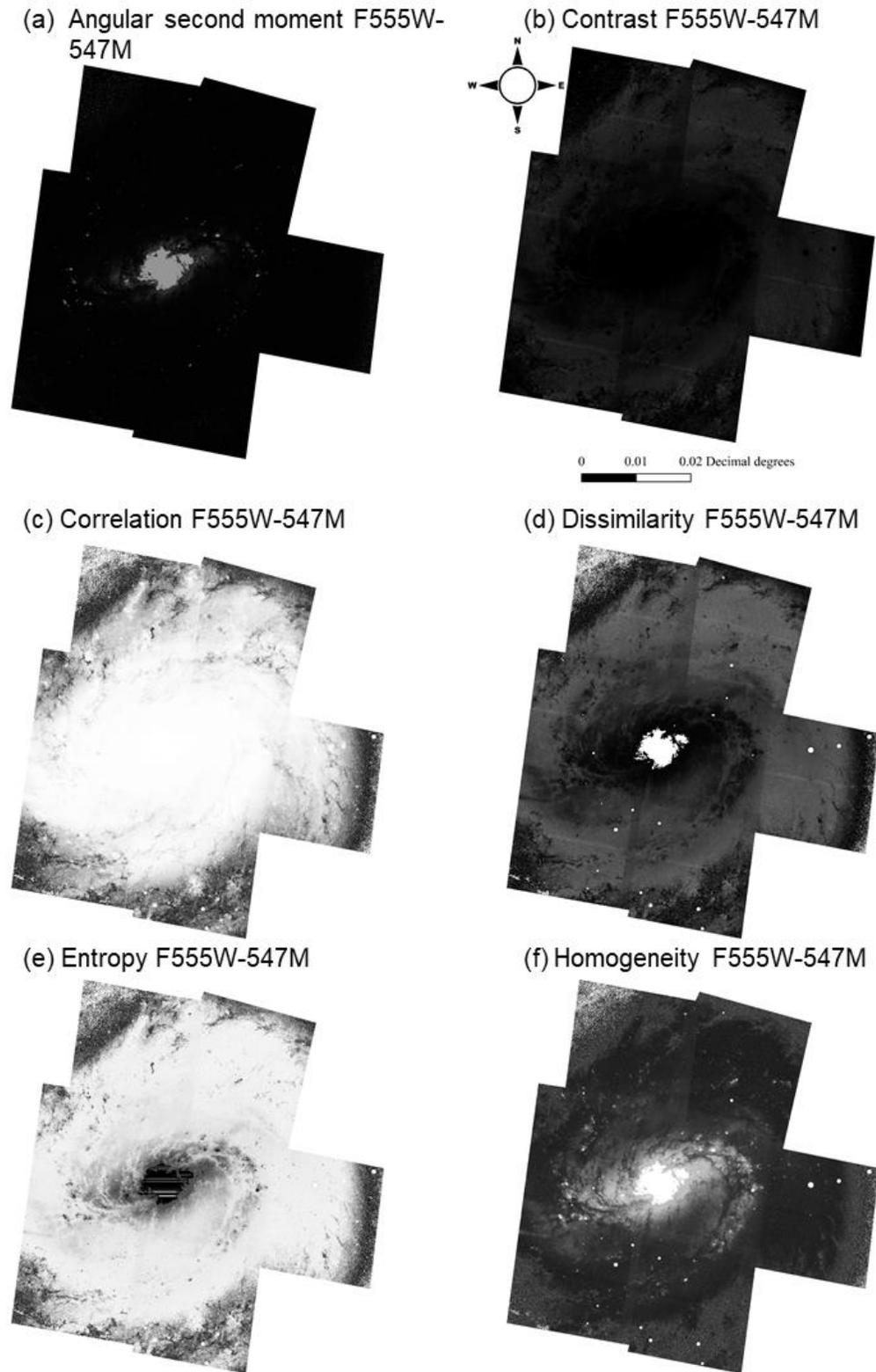
**Figure A 2. Haralick textures calculated on the near-infrared HST image (F814W) of UGC 2885.**

(g) Mean F814W



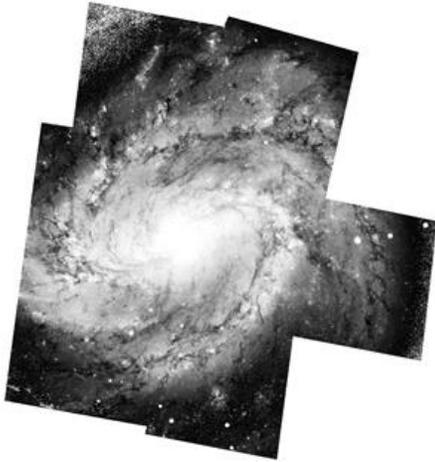
(h) Standard deviation F814W

**Figure A 2. Continued.**

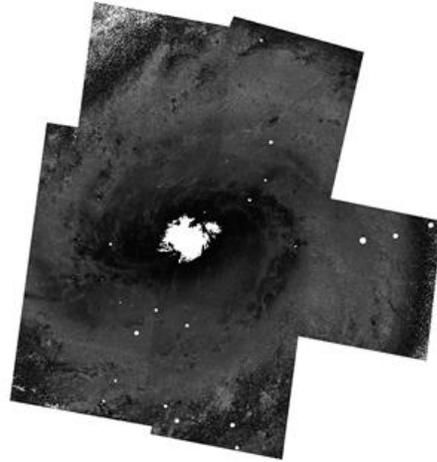


**Figure A 3. Haralick textures calculated on the visual HST image (F555W-547M) of M83.**

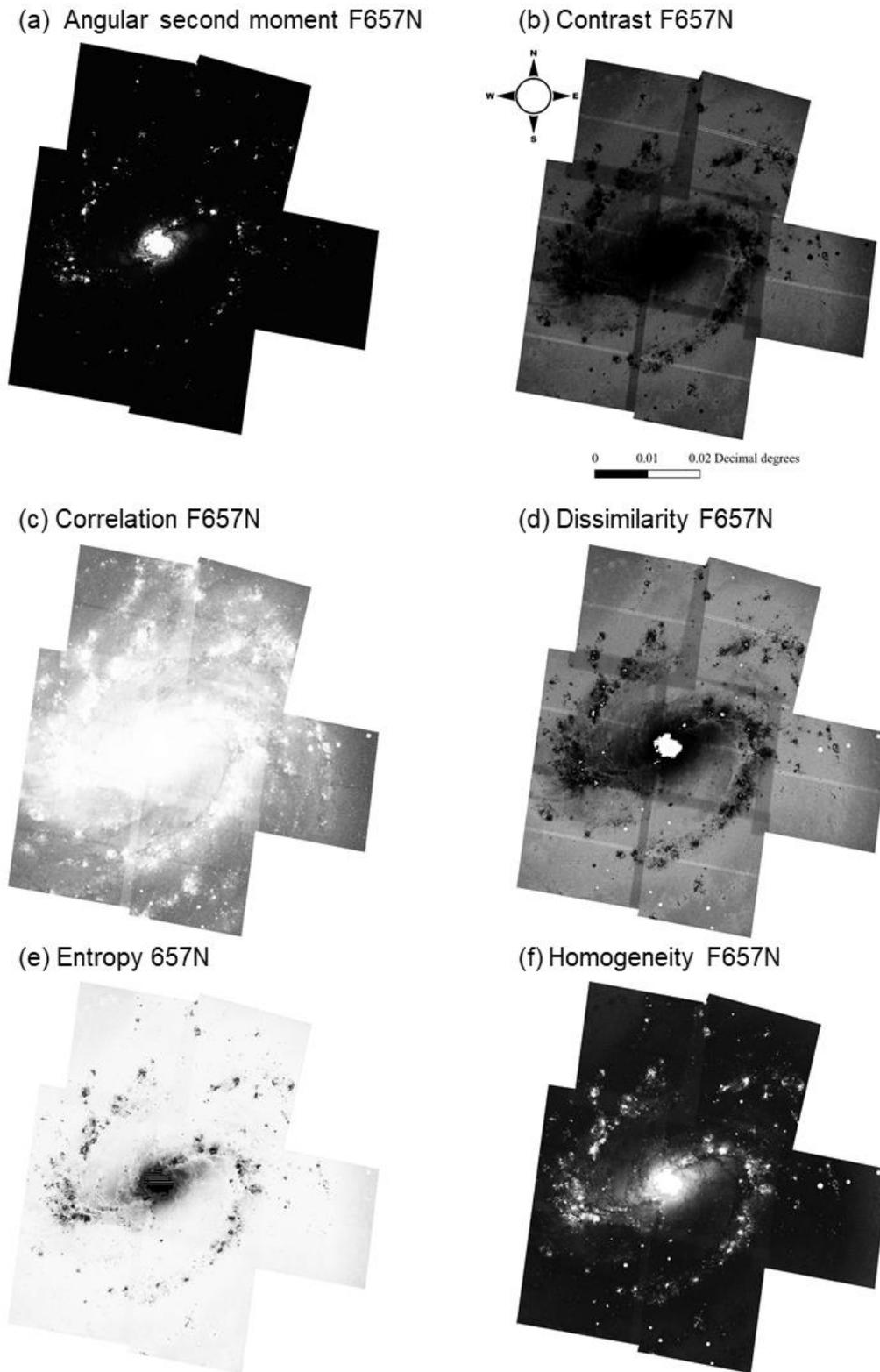
(g) Mean F555W-547M



(h) Standard deviation F555W-547M



**Figure A 3. Continued.**

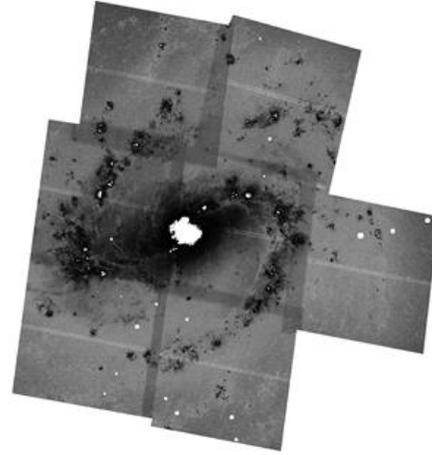


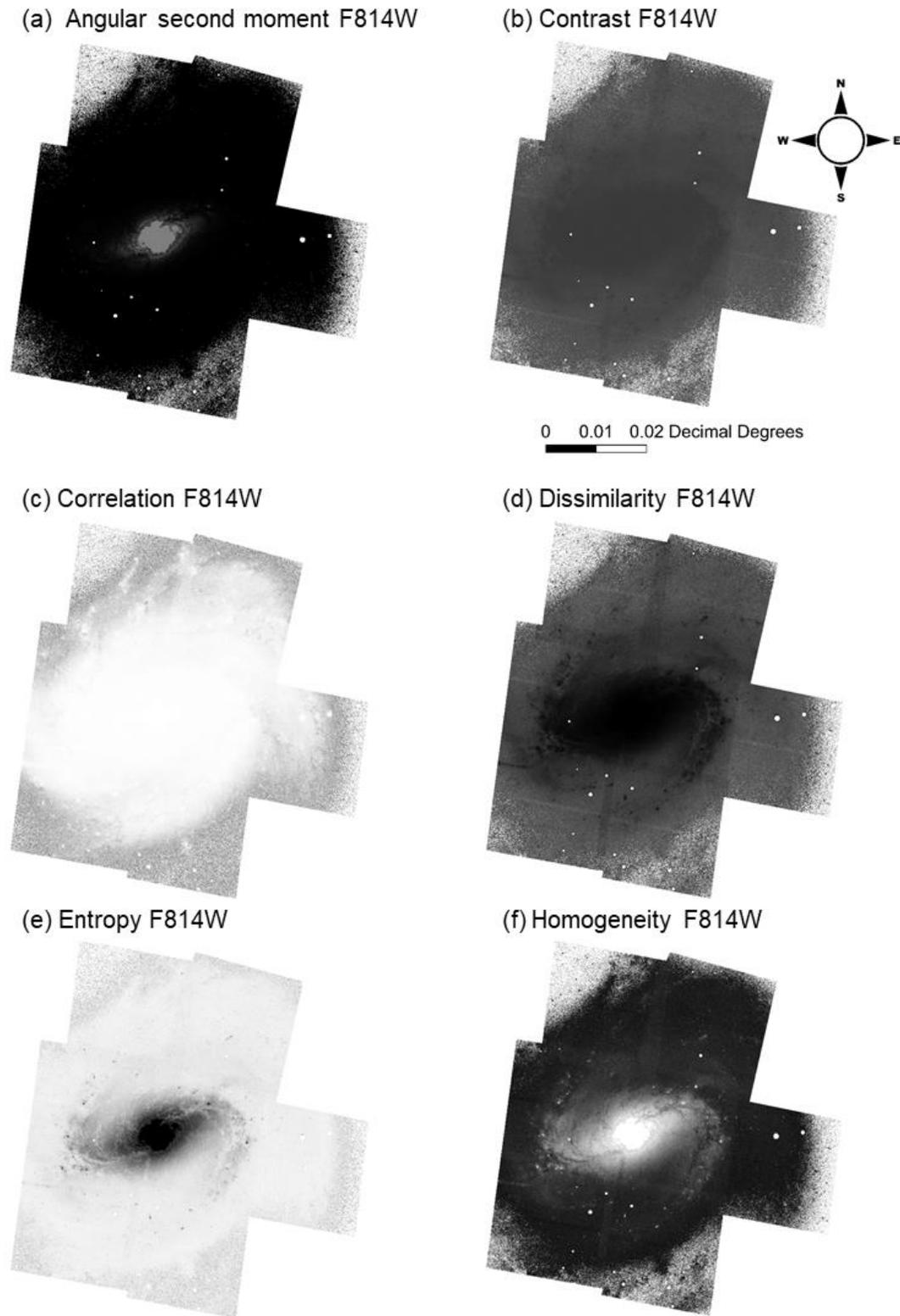
**Figure A 4. Haralick textures calculated on the Hydrogen-alpha HST image (F657N) of M83.**

(g) Mean F657N



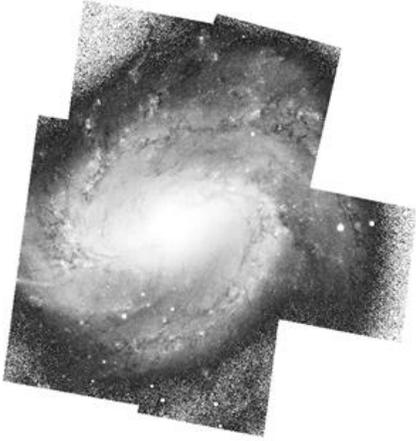
(h) Standard deviation F657N

**Figure A 4. Continued.**

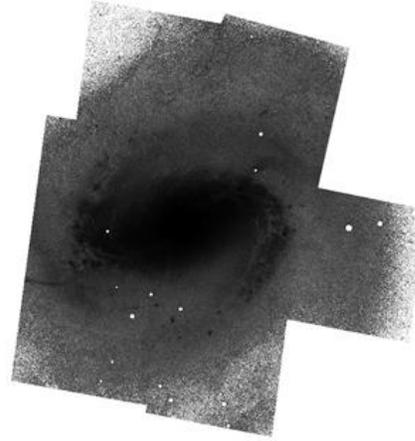


**Figure A 5. Haralick textures calculated on the near-infrared HST image (F814W) of M83.**

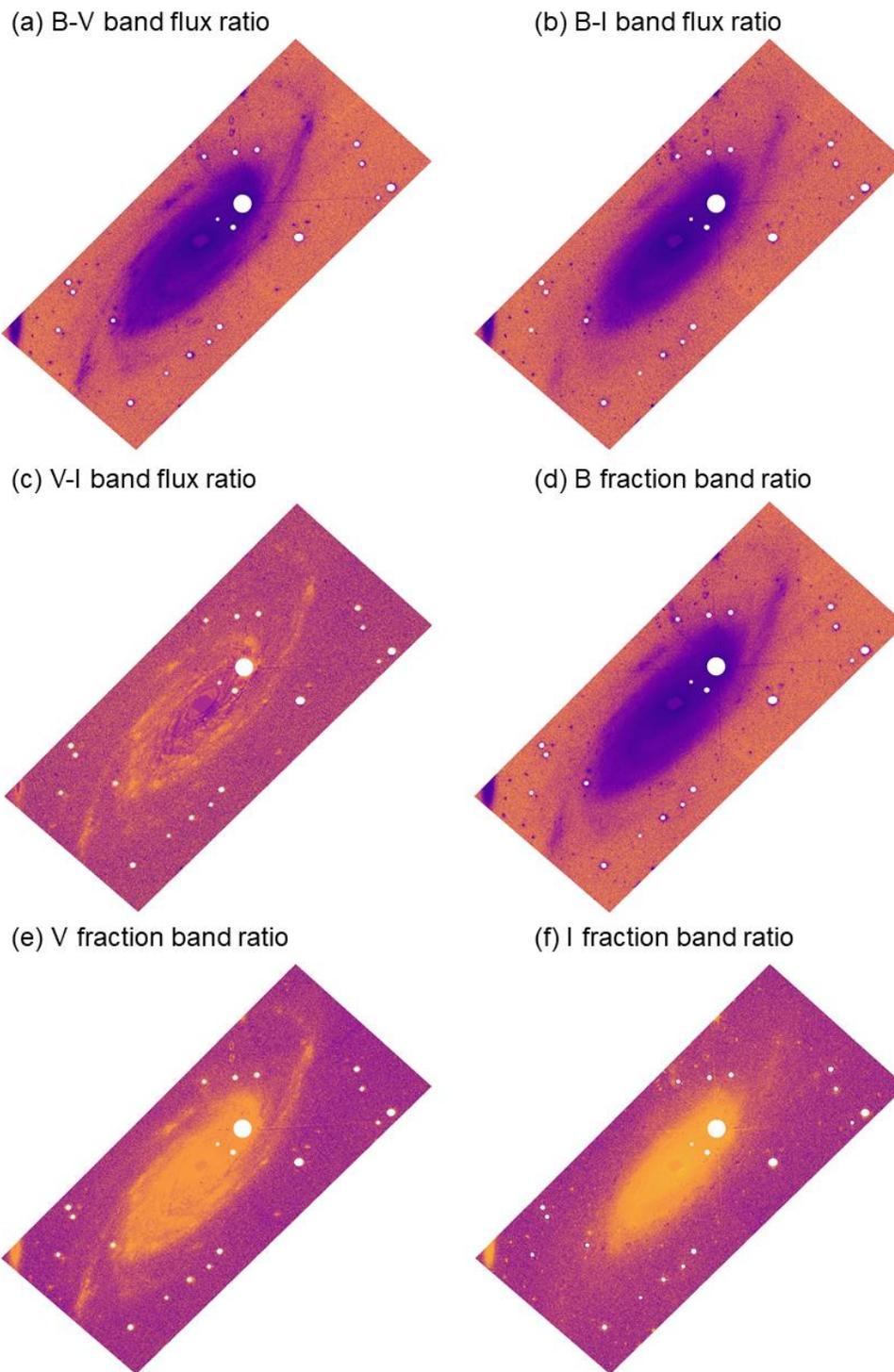
(g) Mean F814W



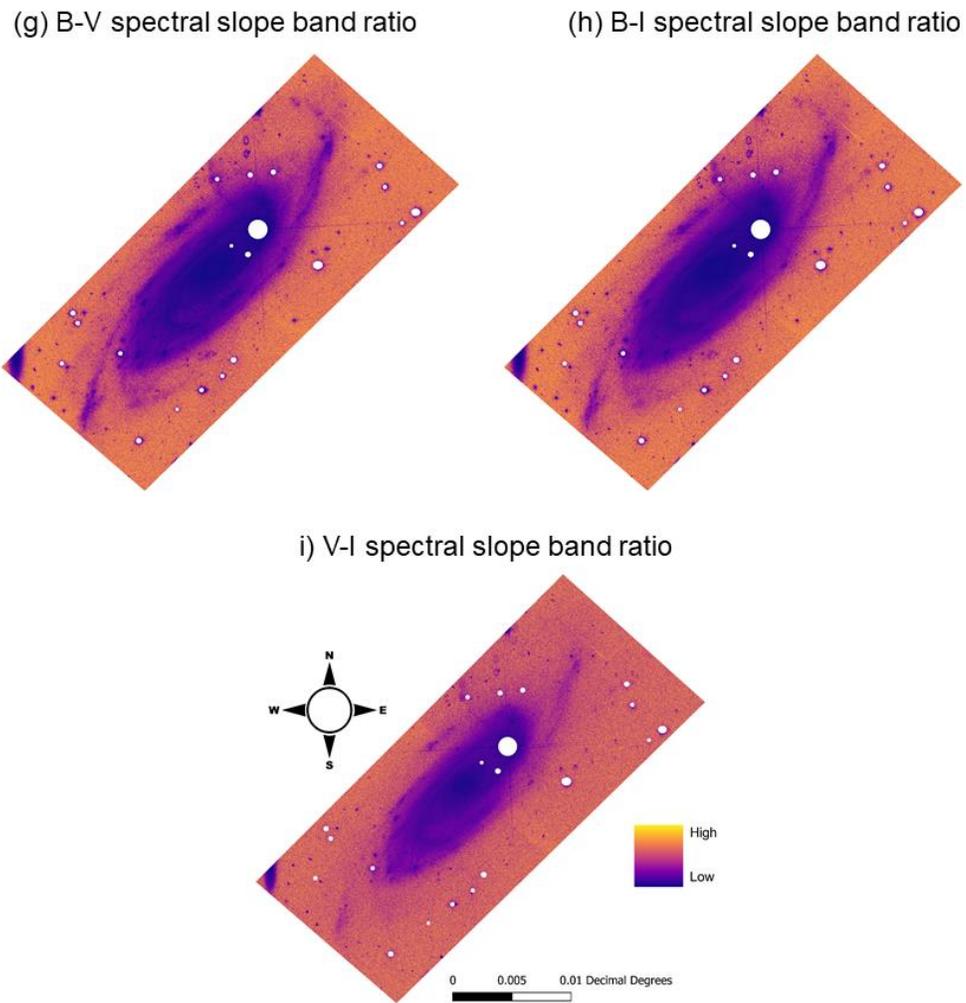
(h) Standard deviation F814W

**Figure A 5. Continued.**

## Appendix B: Band Ratios

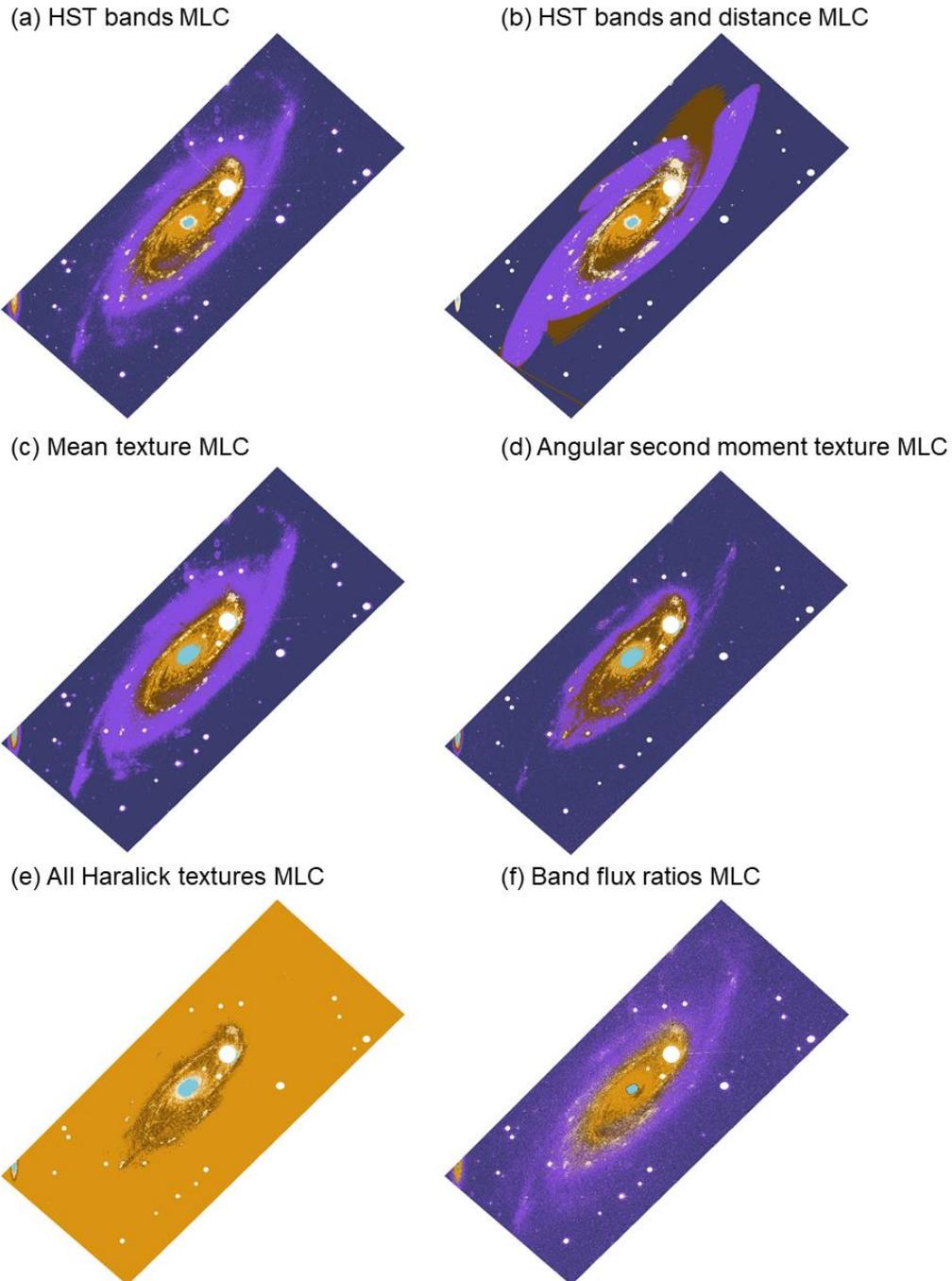


**Figure B 1.** Band ratios calculated from HST imagery of galaxy UGC 2885.



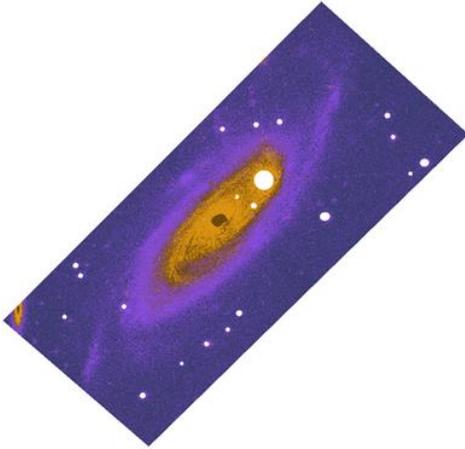
**Figure B 1. Continued.**

## Appendix C: Classified Rasters

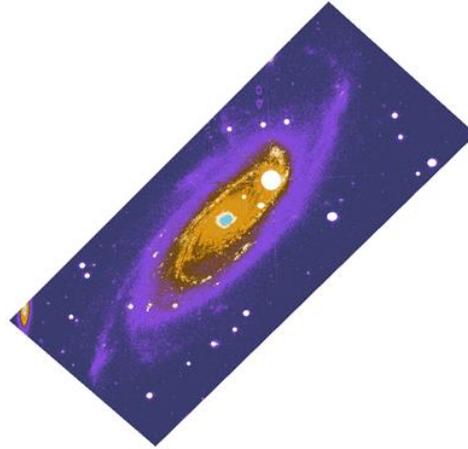


**Figure C 1. Maximum likelihood classifier classifications of components within galaxy UGC 2885 not shown in Chapter 2. Classifications use 70% of the reference set for training and 30% of the reference set for testing.**

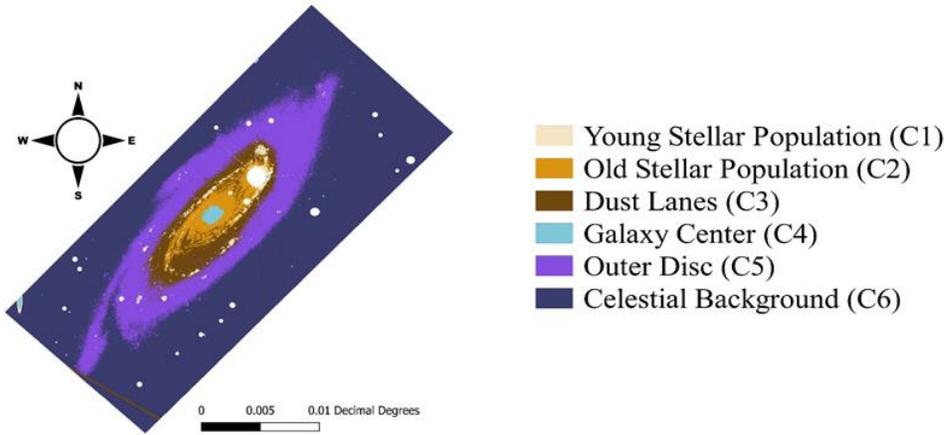
(g) Fraction band ratios MLC

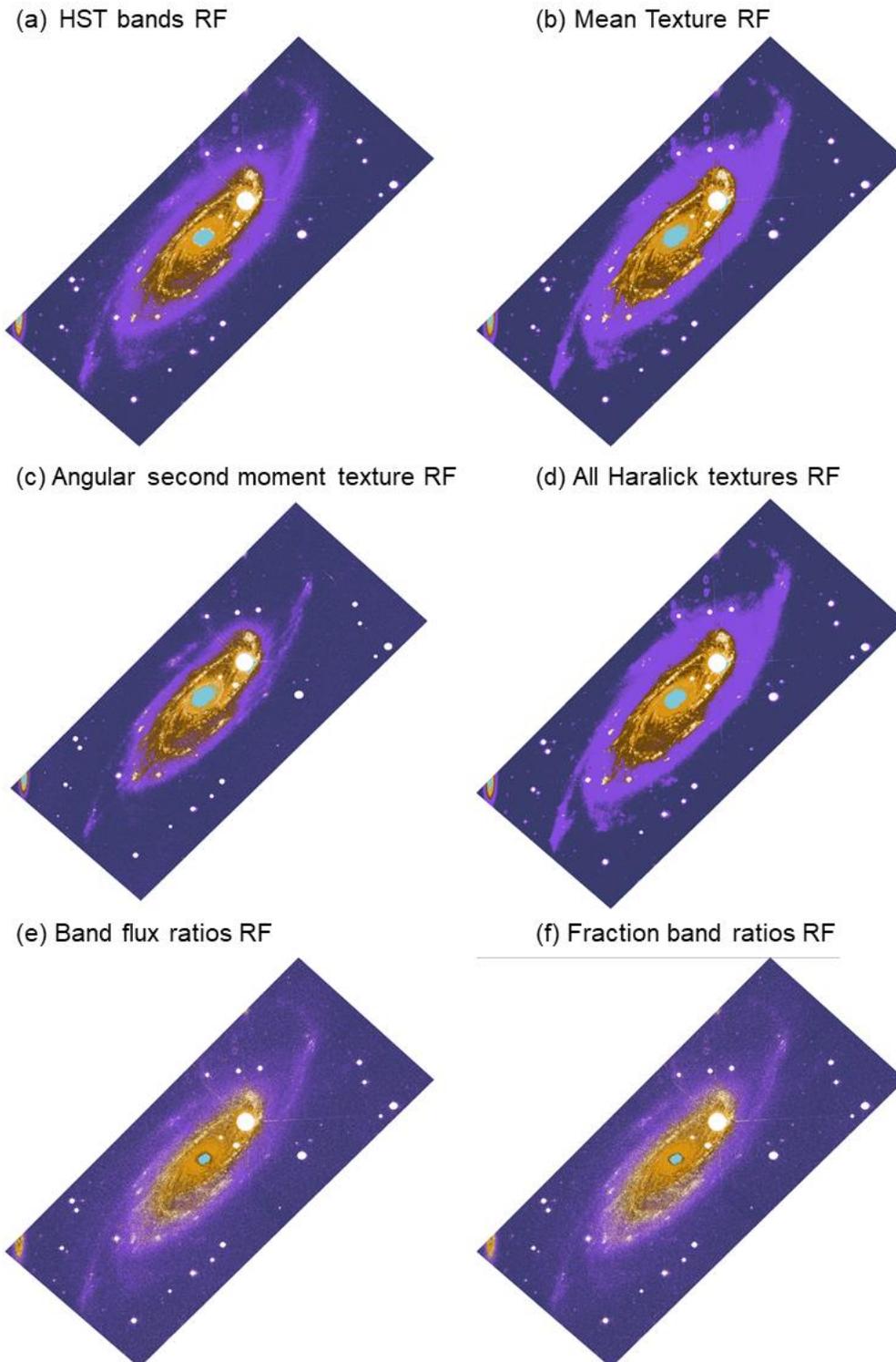


(h) Spectral slope band ratios MLC



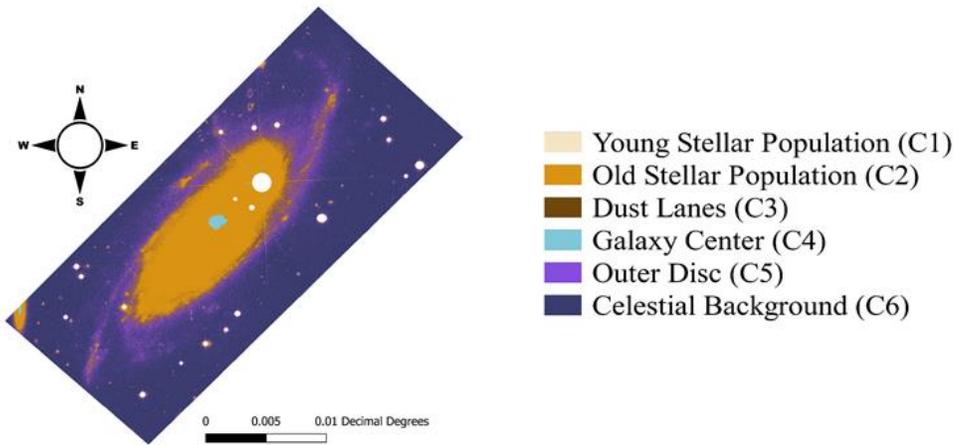
i) Top 7 MDG parameters MLC

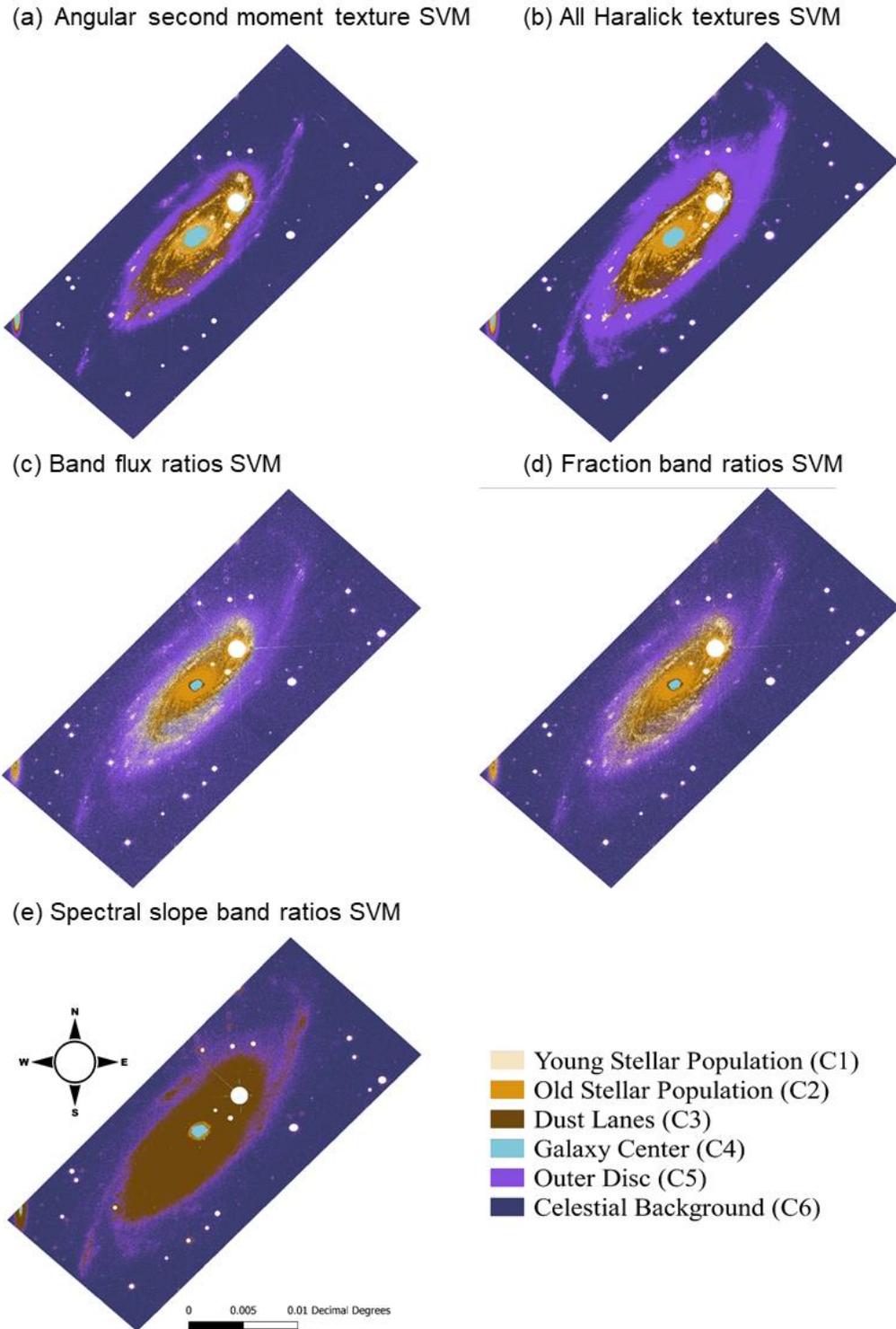
**Figure C 1. Continued.**



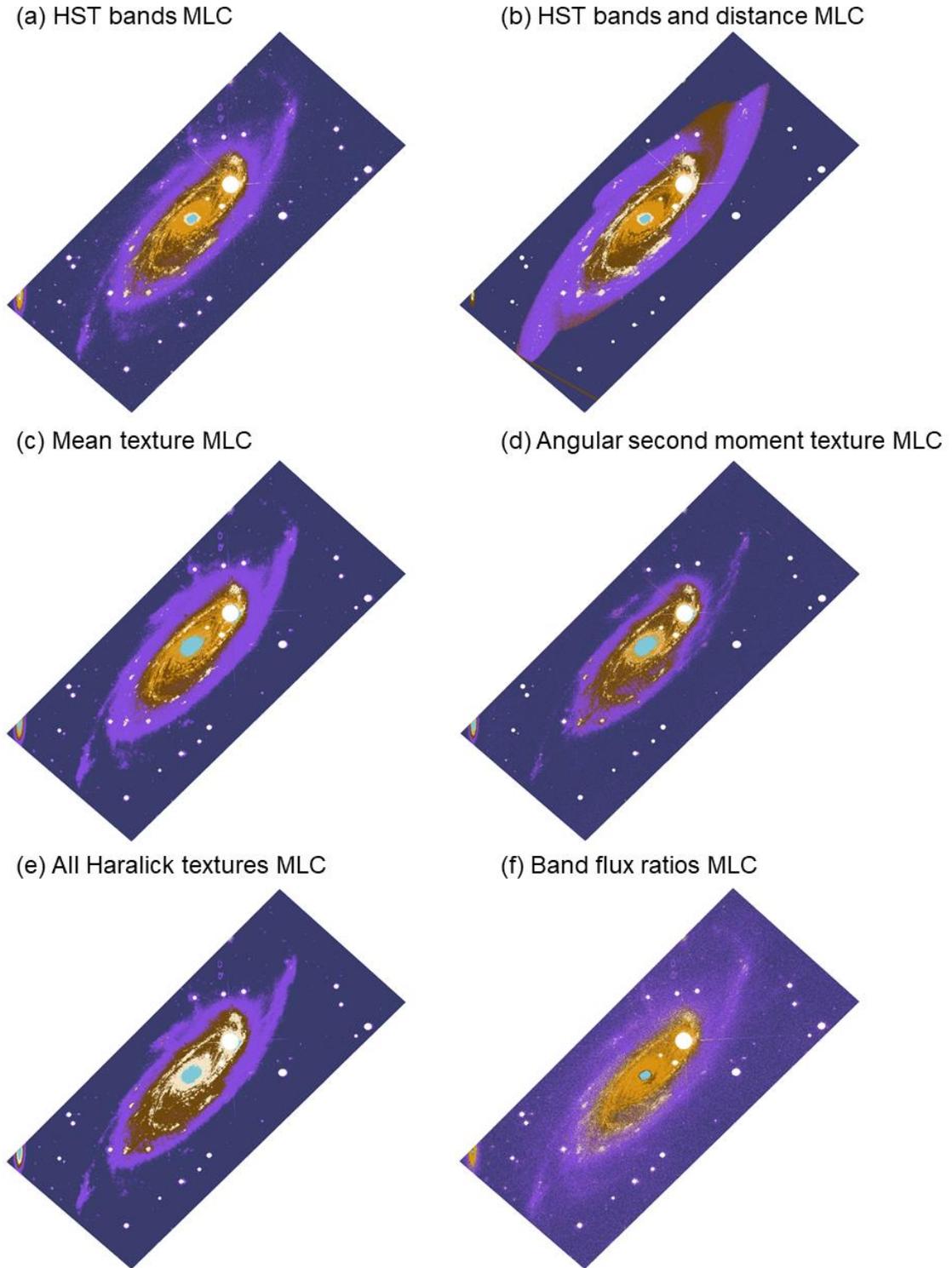
**Figure C 2. Random forest classifications of components within galaxy UGC 2885 not shown in Chapter 2. Classifications use 70% of the reference set for training and 30% of the reference set for testing.**

(g) Spectral slope band ratios RF

**Figure C 2. Continued.**

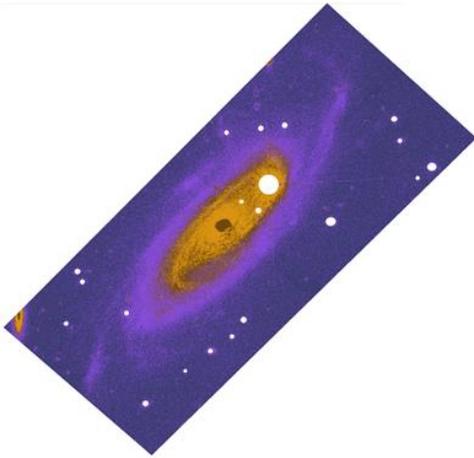


**Figure C 3. Support vector machine classifications of components within galaxy UGC 2885 not shown in Chapter 2. Classifications use 70% of the reference set for training and 30% of the reference set for testing.**

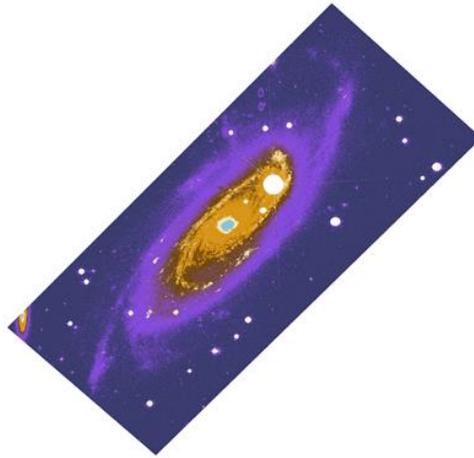


**Figure C 4. Maximum likelihood classifier classifications of components within galaxy UGC 2885 not shown in Chapter 2. Classifications use 30% of the reference set for training and 70% of the reference set for testing.**

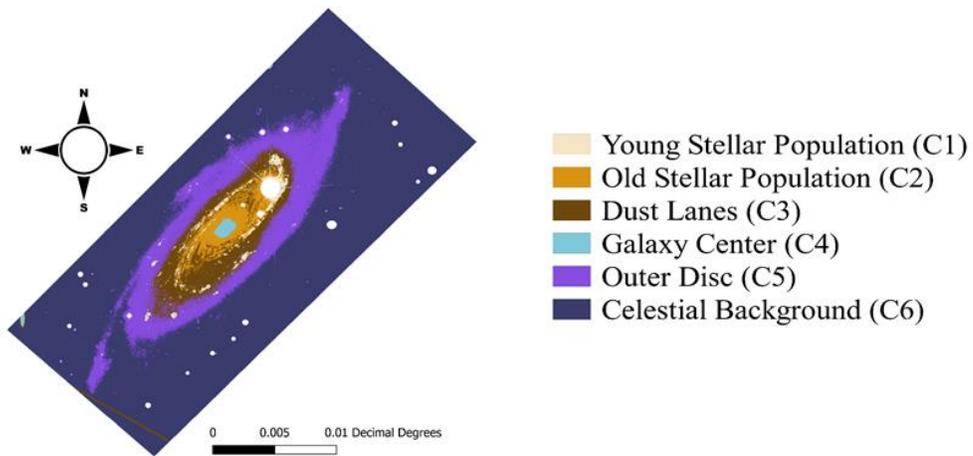
(g) Fraction band ratios MLC

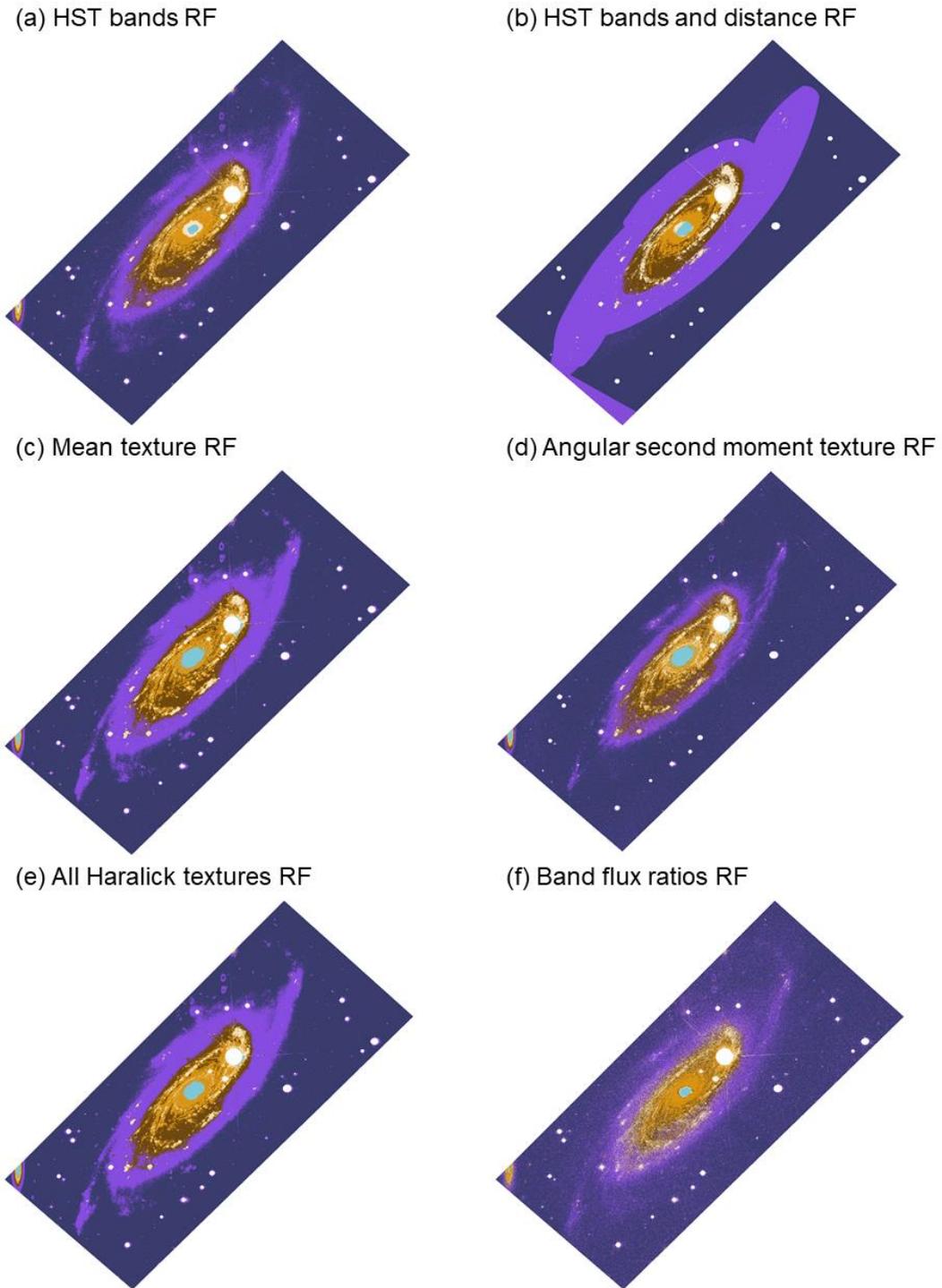


(h) Spectral slope band ratios MLC



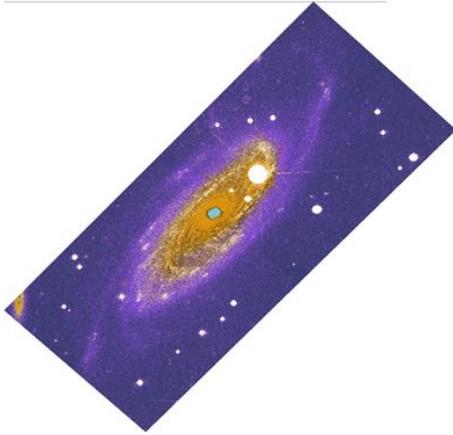
i) Top 7 MDG parameters MLC

**Figure C 4. Continued.**

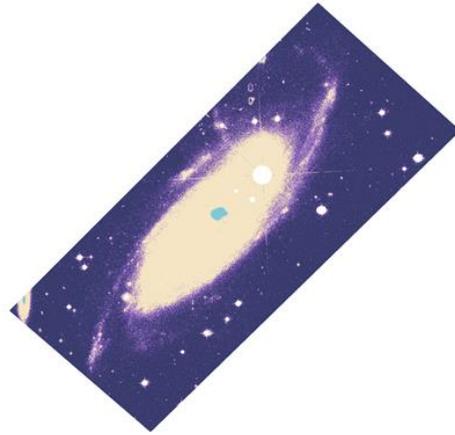


**Figure C 5. Random forest classifications of components within galaxy UGC 2885 not shown in Chapter 2. Classifications use 30% of the reference set for training and 70% of the reference set for testing.**

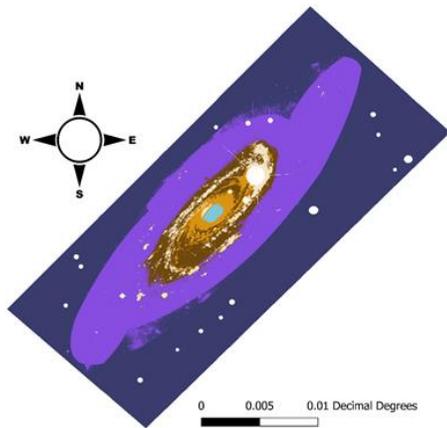
(g) Fraction band ratios RF



(h) Spectral slope band ratios RF

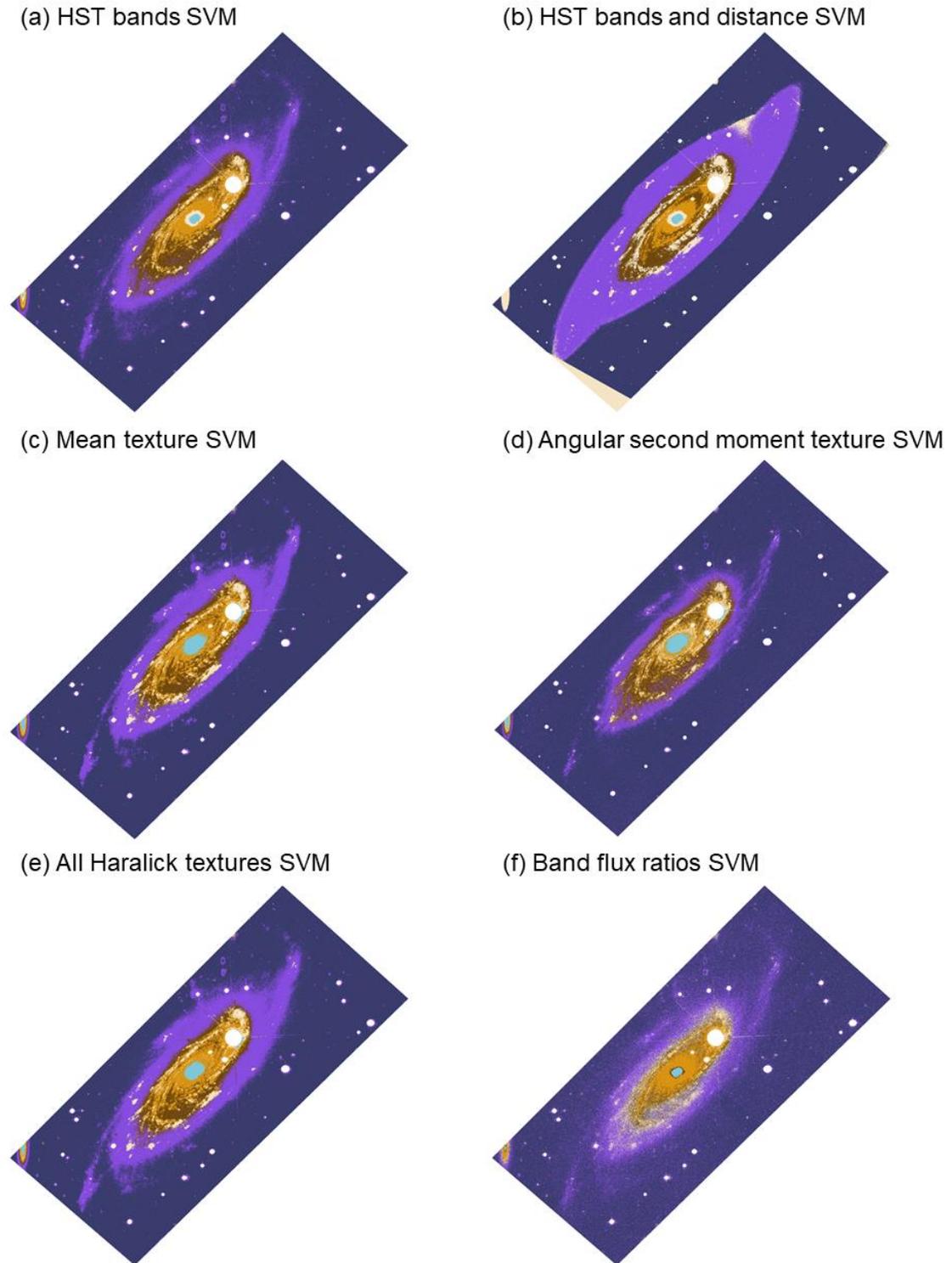


i) Top 7 MDG parameters RF



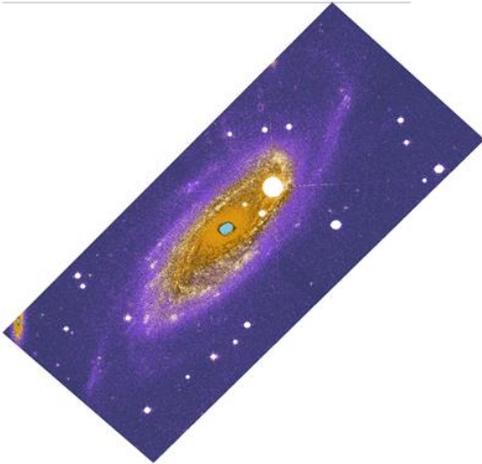
- Young Stellar Population (C1)
- Old Stellar Population (C2)
- Dust Lanes (C3)
- Galaxy Center (C4)
- Outer Disc (C5)
- Celestial Background (C6)

**Figure C 5. Continued.**

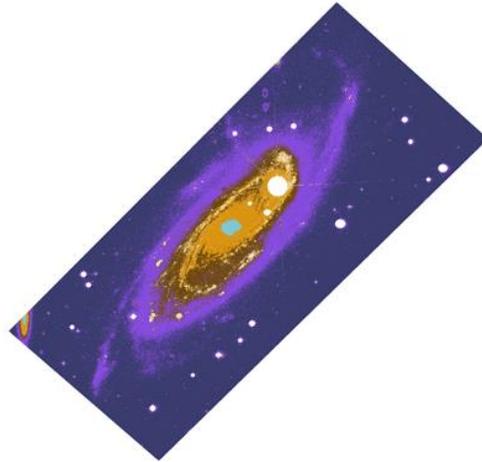


**Figure C 6. Support vector machine classifications of components within galaxy UGC 2885 not shown in Chapter 2. Classifications use 30% of the reference set for training and 70% of the reference set for testing.**

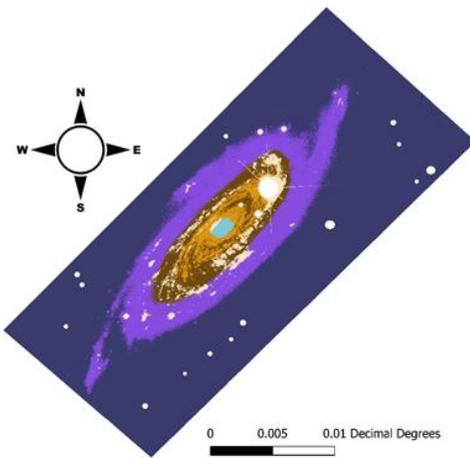
(g) Fraction band ratios SVM



(h) Spectral slope band ratios SVM

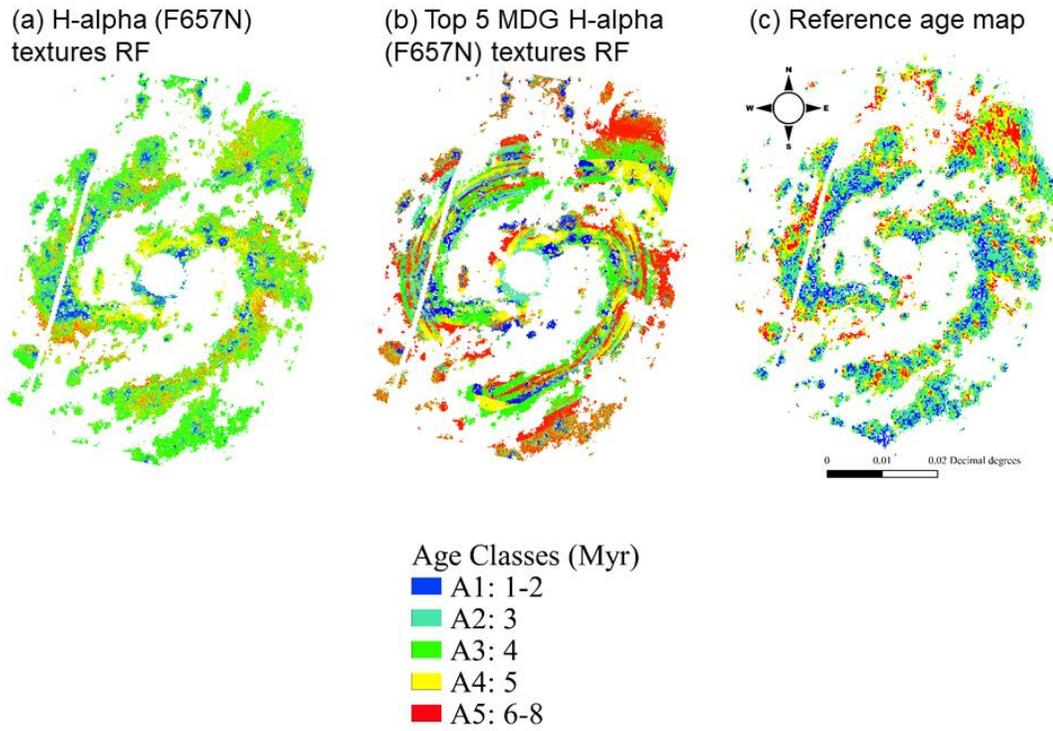


i) Top 7 MDG parameters SVM

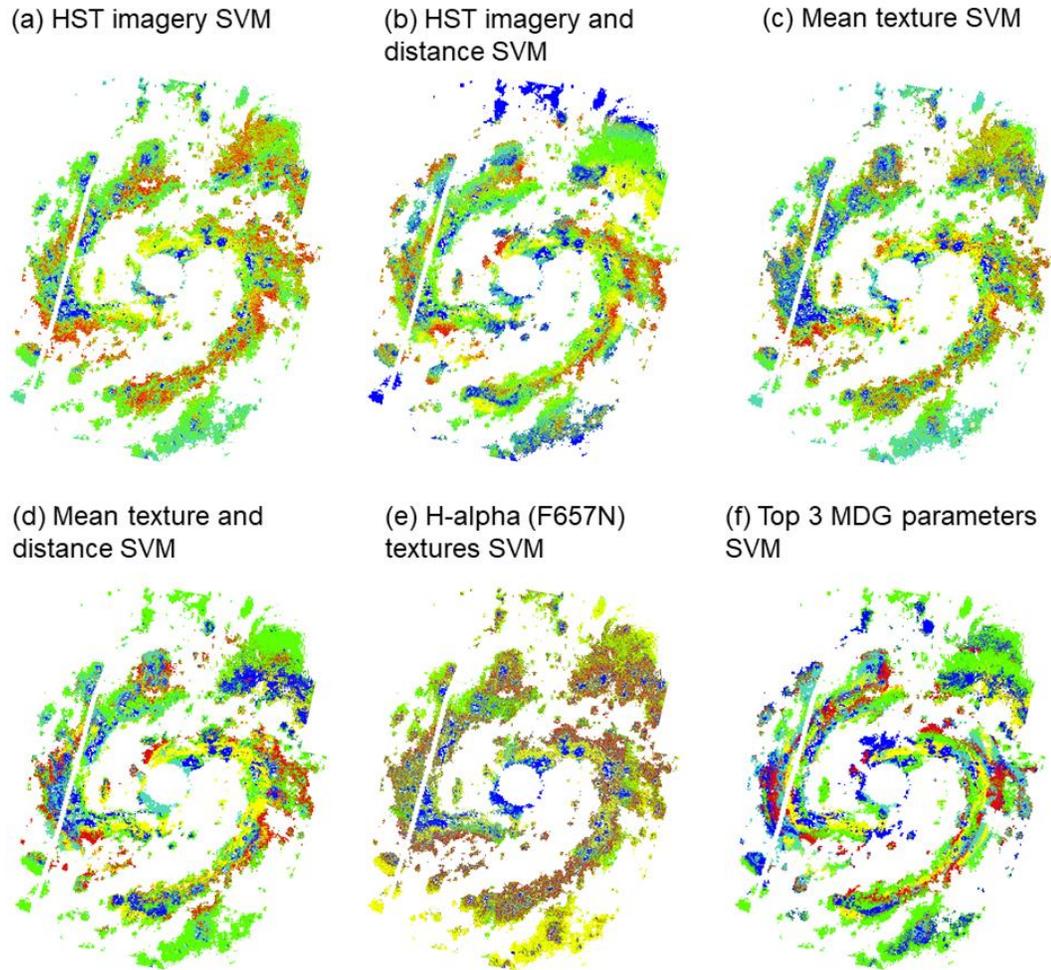


- Young Stellar Population (C1)
- Old Stellar Population (C2)
- Dust Lanes (C3)
- Galaxy Center (C4)
- Outer Disc (C5)
- Celestial Background (C6)

**Figure C 6. Continued.**

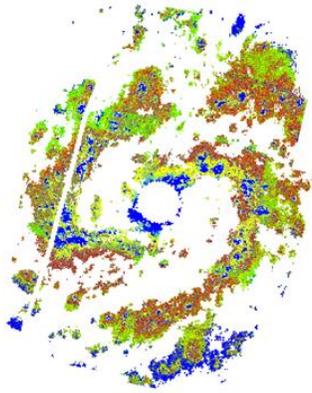


**Figure C 7. Random forest classifications of stellar age within galaxy M83 not shown in Chapter 3. Classifications use 80% of the reference set for training and 20% of the reference set for testing.**

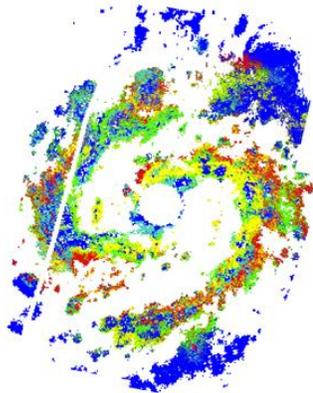


**Figure C 8. Support vector machine classifications of stellar age within galaxy M83 not shown in Chapter 3. Classifications use 80% of the reference set for training and 20% of the reference set for testing.**

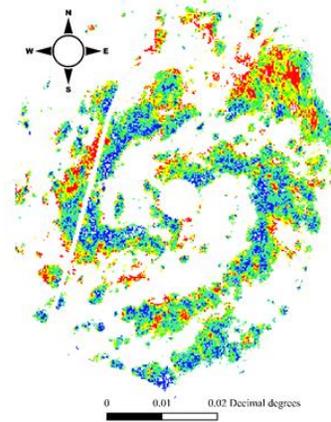
(e) Top 5 MDG H-alpha textures SVM



(f) Top 16 MDG parameters SVM



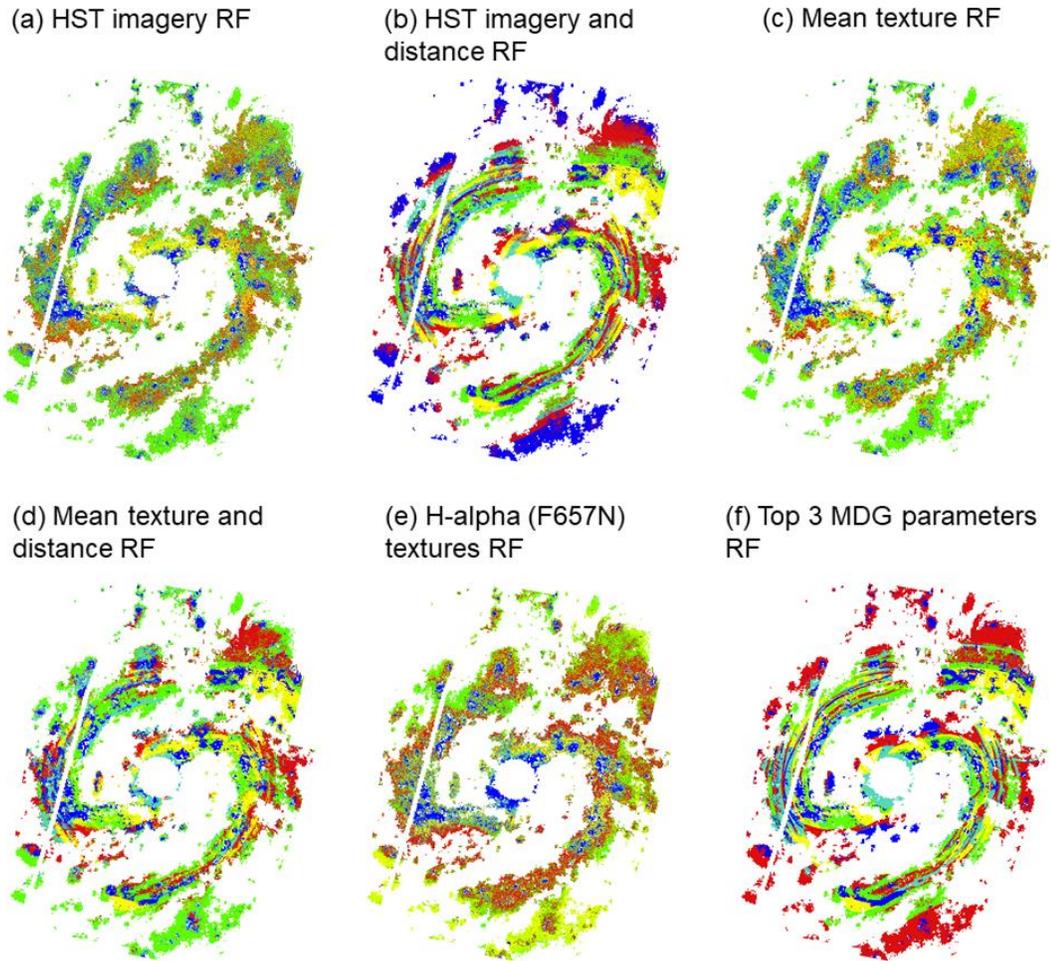
(g) Reference age map



Age Classes (Myr)

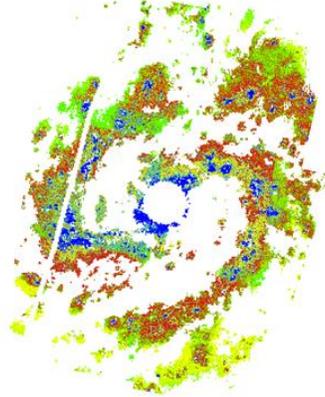
- A1: 1-2
- A2: 3
- A3: 4
- A4: 5
- A5: 6-8

**Figure C 8. Continued.**

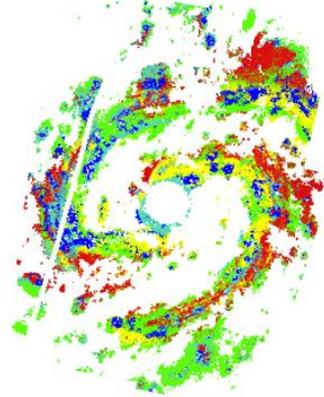


**Figure C 9. Random forest classifications of stellar age within galaxy M83 not shown in Chapter 3. Classifications use 20% of the reference set for training and 80% of the reference set for testing.**

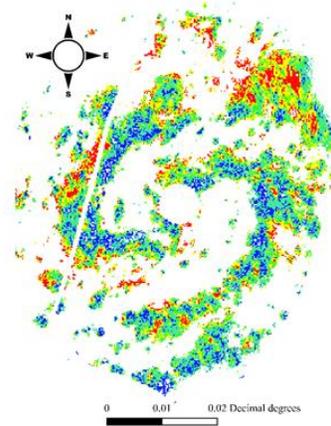
(e) Top 5 MDG H-alpha textures RF



(f) Top 16 MDG parameters RF



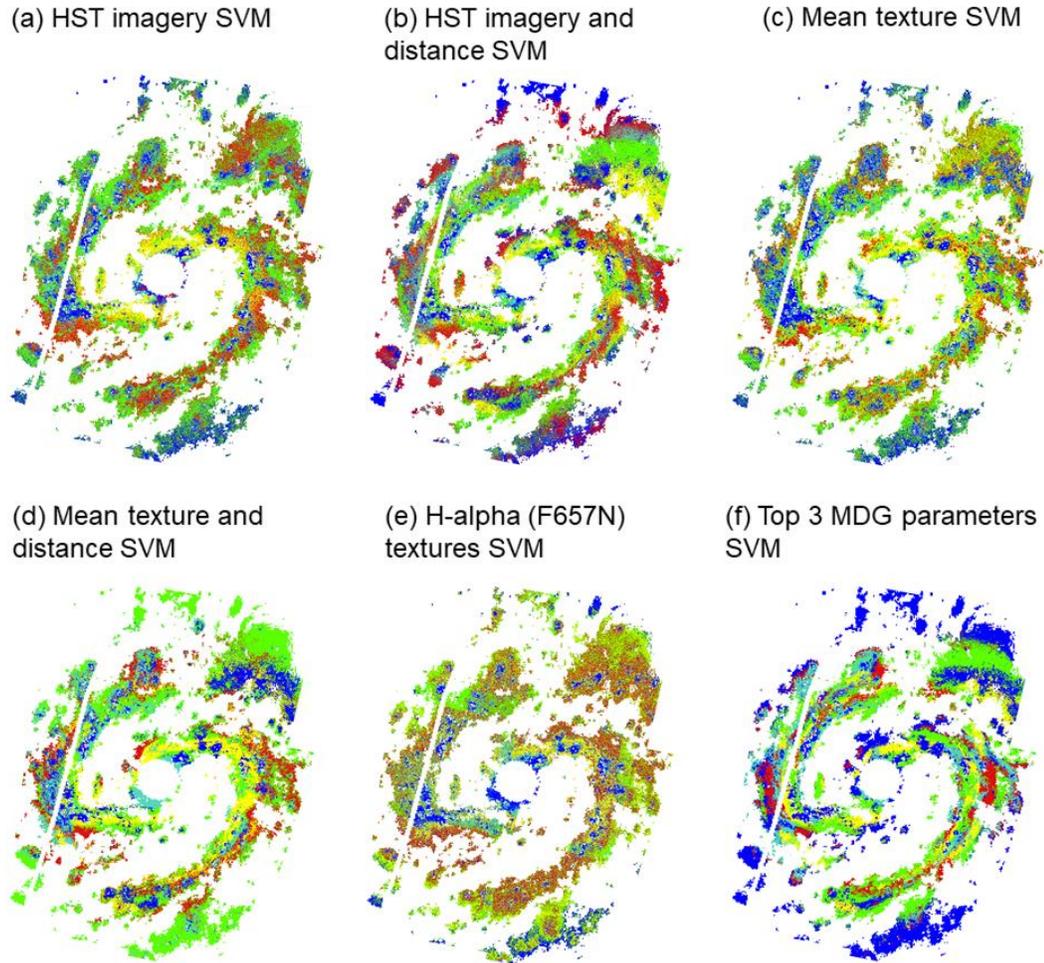
(g) Reference age map



Age Classes (Myr)

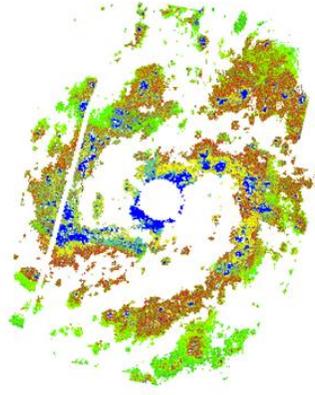
- A1: 1-2
- A2: 3
- A3: 4
- A4: 5
- A5: 6-8

**Figure C 9. Continued.**

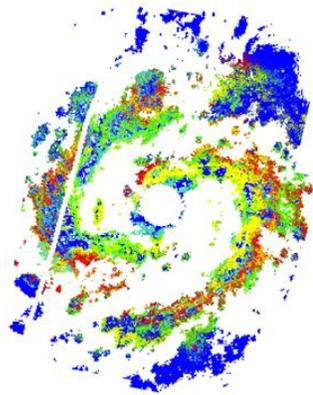


**Figure C 10. Support vector machine classifications of stellar age within galaxy M83 not shown in Chapter 3. Classifications use 20% of the reference set for training and 80% of the reference set for testing.**

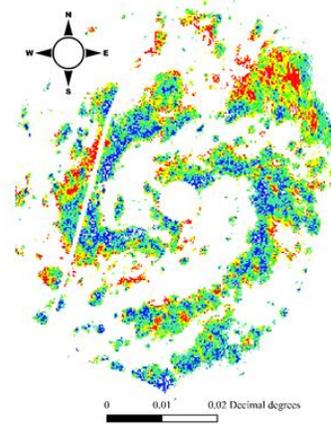
(e) Top 5 MDG H-alpha textures SVM



(f) Top 16 MDG parameters SVM



(g) Reference age map



Age Classes (Myr)

- A1: 1-2
- A2: 3
- A3: 4
- A4: 5
- A5: 6-8

**Figure C 10. Continued.**

## Appendix D: Copyright Releases from Publications

### Thesis Permissions

ASR Editor-in-Chief

22 July 2022 at 3:30

To: Robin Kwik

Dear author,

thanks for your email

first of all, congratulation for your recently published paper in ASR (Galactic component mapping of galaxy UGC 2885 by machine learning classification)

yes, this is correct. You can publish this article as part of your thesis. There is no copyright involved in such a case

Wishing you a long and fruitful scientific career

Pascal Willis, Ph.D.-habil.

Editor-in-Chief

Advances in Space Research

## Curriculum Vitae

**Name:** Robin Kwik

**Post-secondary Education and Degrees:** The University of Western Ontario  
London, Ontario, Canada  
2016-2020 B.Sc. (Hons) Geographic Information Science

The University of Western Ontario  
London, Ontario, Canada  
2020-2022 M.Sc. Geography and Environment, Collaborative Specialization in Planetary Science and Exploration

**Honours and Awards:** Dean's Honor List  
The University of Western Ontario  
2020

King's University College Entrance Scholarship  
King's University College  
2016

**Related Work Experience** Teaching Assistant  
The University of Western Ontario  
2020-2022

Research Assistant  
GITA Lab, The University of Western Ontario  
2019-2022

Precision Agriculture Field Technician  
A&L Canada Laboratories Inc.  
2020

### Publications:

**Kwik, R.J.**, Wang, J., Barmby, P., and Holwerda, B.W. (2022). Galactic component mapping of galaxy UGC 2885 by machine learning classification. *Advances in Space Research*, 70, 229–247. <https://doi.org/10.1016/j.asr.2022.04.032>

**Kwik, R.J.**, Wang, J., Barmby, P., and Sánchez-Gil, M.C. Age classification of spiral galaxy M83 by supervised machine learning. Submitted to *Experimental Astronomy*, July, 2022. Under review.

**Conference Presentations:**

**Kwik, R.J.**, Wang, J., & Barmby, P. (2022, July 11-14). *Machine learning classification of Hubble Space Telescope imagery by remote sensing methods* [Conference presentation]. 10<sup>th</sup> International Conference on Agro-Geoinformatics 43<sup>rd</sup> Canadian Symposium on Remote Sensing, Québec City, QC, Canada.