LÓGICA DIFUSA E ANÁLISE DE IMAGEM APLICADA À IDENTIFICAÇÃO DE ESPÉCIES NO COMPLEXO SIDERASTREA DE

BLAINVILLE, 1830 (CNIDARIA, SCLERACTINIA).

CURITIBA

## FLAVIO ALBERTO PÉREZ

# LÓGICA DIFUSA E ANÁLISE DE IMAGEM APLICADA À IDENTIFICAÇÃO DE ESPÉCIES NO COMPLEXO SIDERASTREA DE BLAINVILLE, 1830 (CNIDARIA, SCLERACTINIA). 

Dissertação apresentada como requisito parcial à obtenção do grau de Mestre em Zoologia, no Programa de Pós-Graduação em Zoologia, Setor de Ciências Biológicas, Universidade Federal do Paraná.

Orientador: Prof. Dr. Marcos Barbeitos

CURITIBA

Universidade Federal do Paraná. Sistema de Bibliotecas.
Biblioteca de Ciências Biológicas.
(Telma Terezinha Stresser de Assis -CRB/9-944)

Pérez, Flavio Alberto
Lógica difusa e análise de imagem aplicada à identificação de espécies no complexo Siderastrea de Blainville, 1830 (Cnidaria, Scleractinia). / Flavio Alberto Pérez. - Curitiba, 2018.

48 p. : il. ; 30cm.

Orientador: Marcos Barbeitos
Dissertação(Mestrado) - Universidade Federal do Paraná, Setor de Ciências Biológicas. Programa de Pós-Graduação em Zoologia.

1. Microscopia. 2. Cnidários. I. Título. II. Barbeitos, Marcos. III.

Universidade Federal do Paraná. Setor de Ciências Biológicas. Programa de Pós-Graduação em Zoologia.

MINISTÉRIO DA EDUCAÇÃO
SETOR SETOR DE CIENCIAS BIOLOGICAS

## TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em ZOOLOGIA da Universidade Federal do Paraná foram convocados para realizar a arguição da Dissertação de Mestrado de FLAVIO ALBERTO PÉREZ intitulada: LÓGICA DIFUSA E ANÁLISE DE IMAGEM APLICADA À IDENTIFICAÇÃO DE ESPÉCIES NO COMPLEXO SIDERASTREA DE BLAINVILLE, 1830 (CNIDARIA, SCLERACTINIA), após terem inquirido o aluno e realizado a avaliação do trabalho, são de parecer pela sua $\qquad$ no rito de defesa.

A outorga do título de mestre está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

Curitiba, 07 de Março de 2018.


## ACKNOWLEDGMENTS

I am grateful to the examining board for evaluating my work, to my adviser, Professor Dr. Marcos Barbeitos, responsible for the Laboratory of Evolution of Marine Organisms (LEOM), to the collaborators of the Universidade Estadual de Campinas (UNICAMP), Dr. Estevão Esmi and Dr. João Florindo for the participation and help in this work, to my laboratory colleagues and to all those people and departments of this institution that have collaborated during my studies. This work was supported by PIBIC/CNPq/UFPR (scholarship 40001016008P4) and Fundação Boticário de Proteção à Natureza (no. 1040_2015).

## RESUMO

O gênero Siderastrea representa um desafio em termos de identificação dada a grande sobreposição existente na distribuição de caracteres diagnósticos quantitativos entre espécies. O objetivo deste trabalho é testar a utilidade de métodos de última geração em descrição e classificação de imagens para a identificação das espécies que compõem o chamado "Complexo Siderastrea do Atlântico", formado por S. radians, S. siderea e S. stellata. As imagens foram obtidas utilizando-se microscopia eletrônica de varredura da superfície do esqueleto de colônias pertencentes às três espécies, coletadas desde Búzios (RJ) a Maraxanguape (RN). Foi utilizado o método de Padrões Binários Locais Completos (CLBP), uma técnica simples porém eficiente, para caracterizar imagem com grande variação localizada em padrões, que tem alcançado resultados excelentes em classificação de textura em bases de dados representativos. Para a classificação supervisionada das imagens, foi utilizada uma rede neural artificial difusa, que simula uma memória associativa ( $\Theta$-FAM) e estabelece a correspondência entre os descritores da imagem e uma das três espécies de corais. A abordagem foi testada usando 370 imagens, sendo 92 imagens de S. radians, 72 de S. siderea e 206 de $S$. stellata. O sucesso médio da classificação individual das imagens obtido em experimentos de validação cruzada foi de $91 \pm 5 \%$, superando todos os outros classificadores testados. Os resultados sugerem que as novas tecnologias de análise e classificação de imagens, conhecidas como visão computadorizada, podem ser ferramentas valiosas para serem aplicadas nas ciências biológicas.

Palavras-chave: inteligência artificial, classificação supervisionada, microscopia, morfometria, computer vision.


#### Abstract

Species identification in the genus Siderastrea is challenging due to the large overlap in diagnostic characters among species. The goal of this thesis is to test the utility of state of the art methods in image analysis and classification to species identification in the so called "Atlantic Siderastrea complex", composed of S. radians, S. siderea and S. stellata. Images were obtained from coralla using a scanning electron microscope. Individuals were collected along the Brazilian coast, from Búzios (RJ) to Maraxanguape (RN). Image analysis was accomplished using Complete Local Binary Patterns (CLBP), a simple but efficient technique to characterize images with localized pattern variation and that has had excellent performance in texture recognition of images from representative databases. Supervised image classification was aided by a fuzzy artificial network that simulates associative memory ( $\Theta$-FAM), establishing the correspondence between image descriptors and putative species. The approach was tested using 370 images, being 92 images of $S$. radians, 72 of $S$. siderea and 206 of S. stellata. Average classification success in crossvalidation experiments was $91 \pm 5 \%$, outperforming all other tested classifiers. The results suggest that the so called computer vision may be a useful tool to be applied in the biological sciences.


Key words: artificial intelligence, supervised classification, microscopy, morphometrics, computer vision.

## SUMMARY

1 INTRODUCTION ..... 6
2 OBJECTIVES ..... 13
3 MATERIAL AND METHODS ..... 13
3. 1 SPECIMEN COLLECTION AND PREPARATION ..... 13
3.2 PHOTOGRAPHIC SAMPLING AND CHARACTER QUANTIFICATION ..... 16
3.3 IMAGE ANALYSIS ..... 19
4 RESULTS ..... 24
5 DISCUSSION ..... 27
6 FUTURE WORK ..... 29
REFERENCES ..... 30
APPENDIX A - IMAGE ANALYSIS ..... 34
APPENDIX B - FUZZY SUPERVISED CLASSIFICATION ..... 36
APPENDIX C - MEASUREMENTS ON THE COLONIES ..... 39
APPENDIX D - BOXPLOT GRAPH OF MEASURED VARIABLES ..... 49
APPENDIX E - SIDERASTREA COLONIES COLLECTION ..... 50

## 1. INTRODUCTION

The genus Siderastrea is composed of five species: S. radians (Pallas, 1766); S. siderea (Ellis and Solander, 1786); S. savignyana Milne Edwards and Haime, 1850, S. stellata Verrill, 1868 and S. glynni Budd and Guzmán, 1994, with controversy about the existence of this last species, given that $S$. glynni it would not be a species, but the product of an unintentional transplant of $S$. siderea colonies from the Caribbean to the eastern Pacific (GLYNN et al., 2016).

Siderastrea savignyana is deeply divergent of from all other Siderastrea species. Siderastrea stellata and S. radians have fixed differences, however, S. glynni and S. siderea share identical sequence types, and together form a monophyletic clade (FORSMAN et al., 2005). Siderastrea stellata and $S$. radians have similar reproductive strategies, high intrapopulational variability and moderate genetic structuring showing sympatric distribution (NEVES et al., 2008). Within regions, dispersal ability appears to be influenced by aspects of reproduction and life history. The broadcasting species $S$. siderea is able to maintain gene flow among populations separated by long distances along the coast of Brazil but in contrast, brooding species, such as $S$. radians, have more restricted gene flow (NUNES et al., 2011).

Siderastrea stellata, S. radians and S. siderea form the so-called "Atlantic Siderastrea Complex" (MENEZES et al., 2014). In this complex, morphological variation poses challenges to classification due to the overlapping of quantitative diagnostic traits, such as, number of septa, corallite diameter, columella diameter, number of papillae and number of synapticular rings, (NEVES et al., 2010; MENEZES et al., 2013; MENEZES et al., 2014; GARCÍA et al., 2017). Thus the interspecific morphological limits in this group remain controversial (MENEZES et al., 2013). Siderastrea has been revised over time by different authors. First descriptions were mainly qualitative, evolving towards a more quantitative approach in time, (PALLAS, 1766; EDWARDS and HAIME, 1850; MILNE EDWARDS 1857; VERRIL, 1868; GREGORY, 1895; DUERDEN, 1902; VAUGHAN, 1919).

The genus was first described in the eighteenth century by (PALLAS, 1766). It was called Madrepora and the publication makes reference to Madrepora radians for the region of the "Mare americanum" and Madrepora stellaris, located in the Indian Ocean. EDWARDS and HAIME (1850), through S. galaxea erected the genus Siderastrea as a encrusting corals, of convex form and dense tissue, fine corallites joined by their walls, deep pit and papilose, little developed columella, with thin, closely spaced, denticulate septa. MILNE EDWARDS
(1857, p. 507, 509), described Astrea radians (transferred from Madrepora) as having "the great diagonal of the chalices of 3 or 4 mm ; and its depth, 2 or a little more", "Columella formed by one or two very compact, visible tubers, 3 complete cycles and, in general, a variable number of fourth cycle partitions that are unusual in many systems". He also describe Astrea siderea, as having "undeveloped columella, reduced to two or three small papillae" and a "large diagonal of the chalices of 4 to 5 mm , its depth 2 ".

Later, VERRIL (1868), in a description of Siderastrea stellata and S. radians mentioned that $S$. stellata differed from $S$. radians in having "larger cells", which appear more open, with thinner septa and, consequently, wider intervening spaces, and four complete cycles of septa. In the beginning of the $20^{\text {th }}$ century, DUERDEN (1902), in his work on Siderastrea radians disagreed with MILNE EDWARDS (1857) and VERRIL (1868), who argued that it is usual to consider that all internal cycles are completed in hexamerical arrangement and all missing septa belong to the last cycle. On the other hand, referring to the development of the septa, DUERDEN (1902) states that "it is clear that such cyclic plans do not express the true ordinal or morphological relationships of the septa", considering that the last and the penultimate cycle vary in the same degree i.e., septa may be missing from either cycle.

GREGORY (1895), who considered S. stellata as a synonym of S. siderea, was contradicted by VAUGHAN (1919), who affirmed that $S$. stellata a very distinct species and not a synonym of $S$. siderea, and reports $S$. siderea owning larger "masses" (colonies) and calices larger on average than those found in $S$. radians. VAUGHAN also described $S$. radians as having corallite diameters between 2-3 mm to 6 mm , whereas in Stellata corallite diameters range from $2.5-5 \mathrm{~mm}$ and in $S$. siderea, from 6 mm to 8 mm , demonstrating that a clear overlap exists. On the other hand, BUDD and GUZMÁN (1994) and REYES et al. (2010) reported calyx diameters for $S$. radians from 2.5-3.5 mm, while in $S$. siderea diameters ranges between 3.0-5.0 mm.

In her work, NEVES et al. (2010) highlights the identification of a colony in particular, as Siderastrea siderea, following the diagnostic morphological skeletal traits to distinguish Atlantic siderastreids with information based on specimens analyzed at the National Museum of Natural History (NMNH) and additional references provided by different authors. The characteristics considered by the author were the diameter of the corallite (3.5 mm to 4 mm ), number of septa (40-64 with fourth cycle ranging from 16-24 septa), arrangement of the septa (regularly distributed in four and sometimes five cycles), type and number of dentition of the primary septum (thin, 0.1 mm in width and 2.0 mm in length, with

10-13 dentations), diameter ( 0.8 mm ) and depth of the columella ( 2.0 mm ), number of papillae (3-6) and number of synapticular rings (4). It was also considered the appearance of the specimen examined with a small dry skeleton $(4.0 \times 3.0 \mathrm{~cm})$, hexagonal, with massive growth form development, encrusting base and distinct oval surface.

MENEZES et al. (2013) analyzed the corallite characters and evaluated the magnitude of variation within and among colonies of S. stellata and S. radians. For this, it is described the use of three different vertical zones of the colony surface (top, middle and edge), selecting six quantitative parameters: corallite and collumela diameter (based on the mean of two greater diameters), columellar diameter, septal number, thecal thickness (which includes the external wall and all the synapticulae), columellar depth, and the average distance between adjacent sampled calices, finding that the septal number was the most important for differentiating the two species.

On the other hand, SANTOS et al. (2004) analyzes the skeletal morphology of $S$. stellata, comparing it with the congeners $S$. siderea and $S$. radians, in order to evaluate the possible latitudinal differences in the populations of Siderastrea. To do this, the shape of the colony and the corallites is described and a descriptive statistics of the morphometric characters in populations of Pernambuco, Paraíba and Panama is performed by using four descriptive parameters: corallite diameter, columellar distance, total number of septa per corallite and number of columella per $\mathrm{cm}^{2}$. The authors found significant differences in corallite diameter, columellar distance and number of septa between $S$. stellata of Pernambuco and Paraíba. The $S$. stellata colonies of geographically close reef environments exhibit marked variation. The morphometric similarities between the populations of $S$. stellata of Paraíba and S. siderea, as well as between S. stellata of Pernambuco and S. radians corroborate that in great geographic distances, the differentiations of the scleractinian species based on the skeletal morphology, become difficulties.

As seen above diagnostic characters are often expressed in terms of ranges or distributions. This complicates species identification because these distributions overlap considerably among species, as evidenced by data compiled from more recent literature (SANTOS et al., 2004; MENEZES et al., 2013; GARCÍA et al., 2017). For instance, columella diameters in $S$. radians from the Gulf of Mexico differ from the measurements made by MENEZES et al. (2013) for the same species in the state of Bahia (Fig. 1a). Likewise, results obtained for the distance among corallites by SANTOS et al. (2004) show that there is more overlap between S. radians from Panama and S. stellata from Pernambuco than within these same species collected from the state of Paraíba (Fig. 1b). In contrast,
distance among columella in $S$. radians from the Gulf of Mexico differ from the measurements made by SANTOS et al. (2004) for S. radians from Panama (Fig. 1c). Number of septa in S. radians from the Gulf of Mexico by GARCÍA et al. (2017) differs from the results obtained by SANTOS et al. (2004) for $S$. radians from Panama and again, there is extreme geographic variation between S. stellata from Paraíba and Pernambuco for this variable (Fig. 1e). The corallite diameter distribution obtained by MENEZES et al. (2013) for S. stellata in the state of Bahia is nearly identical to what was found by SANTOS et al. (2004) for $S$. radians in Panama and S. stellata in Pernambuco (Fig. 1g). An extreme geographic variation between the states of Pernambuco and Paraíba is present in the number of columela per $\mathrm{cm}^{2}$ obtained by SANTOS et al. (2004) for $S$. stellata (Fig. 1h). Theca thickness (Fig. 1d) and columella depth distributions (Fig. 1f) were completely disparate for the same species collected from those two geographic areas.


Figure1. Comparison of results obtained for measured quantitatives variables reported by different authors. Data obtained from Santos et al. (2004) Tab. 4, Menezes et al. (2013) Tab. 4 and García et al. (2017) Tab. 2. The symbols represent the mean, and the bars the confidence intervals calculated as standard deviation x 1,96. Locations are color coded and symbols identify each species (see legend). a. columella diameter (mm), b. distance among corallite (mm), c. distance among columella (mm), d. theca thickness (mm), e. number of septa, f. columella depth (mm), g. corallite diameter (mm), h. number of columella per $\mathrm{cm}^{2}$. SOURCE: The authors (2018).


Figure 2. Overlap observed between S. radians, S. siderea and S. stellata for different variables belonging to García et al. (2017), (not available in fig. 1). The characters represented in this figure were not used by other authors, so they could not be compared to those distributions if Fig. 1. The author do not specify how species delimitation was accomplished before running the discriminant analysis, that requires a priori assignment of specimens to species. a. length of first cycle. b. corallite longest diameter. c. corallite shortest diameter, d. distance among calcification center. e. space between septa. f. synapticular ring width. g. septum thickness. h. number of synapticular ring. i.number of denticules. j. length of last septum. SOURCE: The authors (2018).

Beyond the difficulties in clearly delimiting species due to overlapping distributions, there seems to be author specific disagreements in terms what ranges characterize each species. We also suspect that there may be lack of standardization in terms of how to measure certain variables. For the variable theca thickness (Fig. 1f) details of how the measurements were made by GARCÍA et al. (2017), are not defined, while the measurement of this variable by MENEZES et al. (2013) includes the external wall and all the synapticulae. Details of the columella depth measurement (Fig. 1g) by GARCÍA et al. (2017) are not defined, while MENEZES et al. (2013) considers the columella depth as the vertical distance from the bottom of the columella to the edge of the corallite. However, details about the methodology used are not reported. Although MENEZES et al. (2013) and GARCÍA et al. (2017) worked with specimens from very different geographic regions, it is unlikely that the extreme discrepancy among measurements obtained by these two authors is the product of differences in environmental conditions and/or genetic differences between populations from the Gulf of Mexico and Bahia, given that if the differences were due to environmental variation, they would hardly be so uniform among the three species, so that the distributions found in Brazil and the Caribbean were so different, being more likely the existence of a lack of standardization in the measurements between authors .

The use of software for image processing might be a valuable tool in the recognition and classification of biological structures because it required minimal human input, eliminating error due to subjectivity in variable interpretation. Digital image processing allows for classification of images, feature extraction, multi-scale signal analysis, pattern recognition and projection. It has been applied to remote sensing, feature extraction, face detection, microscope and medical image processing (BASAVAPRASAD; RAVI, 2014).

Texture is an important visual attribute in automated image analysis. Some computational approaches to identify species based on leaf morphology have achieved great degree of success. Such methods are able to differentiate the species based on leaf image properties, where texture is the main analyzed feature. One example of texture analysis is the identification of plant species in the absence of reproductive structures such as flowers and fruits, which are the main source of diagnostic characters but are present year around (DA SILVA et al., 2016).

Another example of this are the identificaction of green plants (Viridiplantae) or species and/or cultivars of the genus Brachiaria given its fundamental importance for practical issues of biology, agronomy and animal science (FLORINDO et al., 2014). Texture analysis from images was also used in the taxonomic identification by genus and species of
arbuscular mycorrhizal fungi (AMF), soil microorganisms and obligate symbionts of plants. The method is based on details of structure, such as color, texture, dimensions and shape. The techniques of taxonomic identification require specialized knowledge, especially experience and training in identification of structures and this is usually tedious and requires time. The processes of image analysis to classify objects have proven to be a useful tool in saving time within scientific research processes (MELÉNDEZ ACOSTA ; RIOS FIGUEROA ; MARÍN HERNÁNDEZ, 2013).

We employed a methodology based on texture due to the very nature of images of coral surfaces obtained via scanning electron microscopy.

Here, we implement an innovative approach to in surpervised classification of images skeletal surfaces. Image descriptors were extracted from each image using a wellknown method of image analysis called Completed Local Binary Patterns (CLBP), (GUO; ZHANG; ZHANG, 2010). Each feature vector assigned to the corresponding (putative) coral species using a fuzzy associative memory strategy, designed to accomodate uncertainty in species identification due to overlapping in diagnostic characters.

## 2. OBJECTIVES

To assess the procedure of classification through image analysis and fuzzy supervised learning to species discrimination in the genus Siderastrea as test of the method's applicability to other taxonomic problems in the order Scleractinia.

## 3. MATERIAL AND METHODS

### 3.1. SPECIMEN COLLECTION AND PREPARATION

The Brazilian coast is divided in four main regions, northern, northeastern, eastern and oceanic islands (LEÃO et al., 2016). Coral reefs are found in three major sectors of the tropical coast of Brazil: the northern region conformed by the states of Amapá, Pará, Maranhão, Piauí and Ceará, the northeastern by Rocas, Fernando de Noronha Island, Rio

Grande do Norte and Paraíba, and the eastern coasts by Bahia, (north, central, bay, south). The southern region which covers the states of Espírito Santo, Trindade Island, Rio de Janeiro, São Paulo, Paraná, Santa Catarina and Rio Grande do Sul, has only isolated coral populations, but no real reefs (LEÃO; KIKUCHI; TESTA, 2003).

Sampling was carried out in the states of Rio de Janeiro (Búzios), Espírito Santo (Aracruz), Bahia (Abrolhos, Boipeba), Pernambuco (Tamandaré), Rio Grande do Norte (Maracajaú) and Ceará (Fortaleza), (Fig. 3), (TABLE1). The specimens collected were $S$. stellata (endemic to Brazil), S. radians, and S. siderea. Siderastrea stellata is reported to include the Caribbean, the Gulf of Mexico and the states of Piauí, Ceará, Rio Grande do Norte, Paraíba, Pernambuco, Alagoas, Sergipe, Bahia, Espirito Santo, Rio de Janeiro, Atol das Rocas, Fernando de Noronha and Trindade islands in Brazil. Siderastrea radians with occurrence in the Caribbean, the southern Gulf of Mexico, Florida, Bahamas, Bermuda and the eastern Atlantic, while S. siderea is present in the Caribbean, Gulf of Mexico, Florida, Bahamas, Bermuda, and it has been recorded as far as the northern coast of North Carolina. For the Brazilian coast there are records of $S$. radians for the states of Rio Grande do Norte, Paraíba, Pernambuco, Alagoas, Bahia and Espirito Santo, and in the Bahia state for S. siderea.


Figure 3. Map of Brazil indicating place of sampling and species collected in each locality along the coast. Colors: Red (Siderastrea stellata), green (S. radians) and yellow (S. siderea). SOURCE: Adapted from www.pixabay.com (2018)

Samples were obtained from shallow-waters of 0.5-1.0 m deep by snorkeling, or SCUBA diving in deeper waters (from 1 m to 30 m ).

| STATE | LOC. | SITE | LAT. | LONG. | DEPTH | S.rad. | S.stel. | S.sid. |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Bahia | Abrolhos | Chapeirão | -17.90 | -38.83 | 14m-15m | X | X | X |
|  |  | Chapeirinhos da Sudeste | -18.20 | -38.79 | 8m-12m | X | X | X |
|  |  | Caldeiras | -18.20 | -38.87 | 4m-5.5m | X | X | X |
|  |  | Chapeirões do sul | -18.07 | -38.73 | 21m-23m | X | X | X |
|  |  | Redonda | -18.00 | -38.89 | 5m-6m | X | X | X |
|  | Ilha de Boipeba | Tassimirim | -13.58 | -38.91 | 0.5m-4m | X | X | X |
|  |  | Ponta dos Castelhanos | -13.67 | -38.89 | 0.5m-2m | X | X | X |
| Espírito santo | Aracruz | Praia do Pichado | -20.01 | -40.16 | $\begin{gathered} \hline 0.5 \mathrm{~m}- \\ 1.5 \mathrm{~m} \end{gathered}$ | X | X | - |
| Rio de janeiro | Búzios | Praia da Tartaruga | -22.80 | -41.91 | 0.5m-2m | - | X | - |
| Rio grande do norte | Maracajaú | Parrachos de Maracajaú | -5.39 | -35.25 | 0.5m-2m | X | X | - |
| Pernambuco | Tamandaré | Praia de Campas | -8,74 | -35.08 | 0.5m-2m | X | X | - |
| Ceará | Fortaleza | Cabeço do Arrastado | -3.60 | -38.39 | 28m-30m | - | X | - |
|  |  | Cabeço do Balanço | -3.60 | -38.39 | 20m-22m | - | X | - |
|  |  | Praia do Aterro | -3.72 | -38.51 | - | - | X | - |

TABLE 1. Table indicating sampling sites, geographic coordinates, sampling depth and species collected. Loc. (locality), lat. (latitude), long. (longitude). SOURCE: The authors (2018).

Colony fragments were removed using a hammer and chisel, and placed in plastic bags for transportation. The method used to eliminate colonial tissues basically consisted in allowing the colony to decompose for more than a week, which lead to the loss of most of the tissue. Subsequently, each sample was washed with tap water, placed in a new bag, and immersed in a solution of sodium hypochlorite in order to remove any remaining traces of tissue. The samples were then washed again with tap water to permanently clean the colony. The colonial skeletons were kept in an electric plant dryer for 15 hours before being deposited in the Collection of the Department of Zoology in the Federal University of Paraná (Appendix E).

### 3.2. PHOTOGRAPHIC SAMPLING AND CHARACTER QUANTIFICATION

For the quantification of characteristics were taken two types of photographs, optical microscopy, for measurement of diagnostic characteristics and electronic microscopy for texture analysis. Ninety-three colonies were used for the collection of data and photographs of mature corallites (i.e., those with, at least a complete third septal cycle, excluding those on the edges of the colonies) were taken. A total of 930 optical photographs (five of individual corallites and five of their columellas), considering 10 photos per colony, were taken with a Leica stereomicroscope corresponding to $2.5 \times$ magnification for corallites and $6.3 \times$ for the respective columellas.

Five quantitative characters were selected for morphological analysis: corallite (Fig. 4a) and collumela diameter (Fig. 4e), (both based on the mean of the largest diameter and the smallest diameters); number of septa (Fig. 4c); first cycle septum thickness (Fig. 4b) and length (Fig. 4d). The number of sinapticular rings and papillae of the columella were also determined. The number of synapticular rings corresponds to the calculation of the number of rings formed by the set of synapticules around the corallite. Septa arrangement was determined through the ImageJ program by marking on the optical photographs, the septa corresponding to each cycle, thus defining cycle completeness. The qualitative characters budding type and septa continuity were also considered. The continuity of the septum was determined by observing the existence or not of a septum as a division between adjacent corallites and colonies. The qualitative analysis corresponded to the observation of the budding type from a general photograph of the colony with $10 \times$ magnification.

The number of sinapticular rings and papillae of the columella were later discarded since they were considered unreliable data given the quality of the photograph for that particular type of measurement. Corallite diameter, number of septa, septa arrangement, thickness and length of the first cycle septum, diameter of the columella, budding type and continuity of the septum were retained for species identification, (TABLE 2).

## MORPHOLOGICAL QUANTITATIVE CHARACTERS MEASURED

Corallite diameter
Number of septa
Septa arrangement
Thickness of the first cycle septum
Length of the first cycle septum
Diameter of the columella
Budding type
Continuity of the septum

TABLE 2. Quantitative variables measured for morphological analysis. SOURCE: The authors (2018).

Putative identifications were made through optical photographs based on ordinary taxonomic characters following the work of NEVES et al. (2010). This work was used as a guide for the identification of species since it has complete and tabulated information prepared from a compilation of descriptions made by different authors. As NEVES states, the table was elaborated by compiling data obtained through a modification by BUDD and GUZMÁN (1994) and with additional information from MILNE EDWRADS and HAIME (1849, 1857), VERRILL $(1868,1901)$, VAUGHAN (1919), YOUNGE (1935), ALMY JR. and CARRIÓN-TORRES (1963), LABOREL (1969, 1970), FOSTER (1979, 1980), DEBROT et al. (1998). Furthermore, the table also contains information collected by the author herself.

For the identification of species in the present work, it was considered the number of septa, the existence, or not, of a complete fourth cycle, the type of budding (intra-tentacular in S. radians and S. stellata and extra-tentacular in S. siderea) and the continuity of the septum (continuous in S. radians and S. stellata and discontinuous in S. siderea). It was not possible to measure the variable depth of the columellar fossa since the work was performed on optical photographs that did not allow this type of measurement.

In some colonies the number of septa allowed to identify a priori the species, for example in the identification of $S$. radians with respect to the other two species. Those colonies with less than 34 septa and absence of a complete fourth cycle were identified as $S$. radians. Since the number of septa is a characteristic with a high degree of overlap, for most colonies it was necessary to take into account, together with the number of septa, the presence or absence of a complete fourth cycle. The presence of a fourth complete cycle allows discarding the colony as belonging to $S$. radians (in which a fourth complete cycle never
occurs), (NEVES et al., 2010; MENEZES et al., 2013). Colonies with more than 34 septa, incomplete or complete fourth cycle, continuous septum and intra-tentacular budding were identified as $S$. stellata. Colonies with more than 40 septa, presence of a complete fourth cycle, discontinuous septa and extra-tentacular budding were identified as $S$. siderea, (TABLE 3). In this species the fourth cycle is rarely incomplete, meaning that it is almost always complete.

|  | SPECIES IDENTIFICATION |  |  |
| :--- | :---: | :---: | :---: |
| Number of septa | S. stellate | S. radians | S. siderea |
| Fourth cycle | $>34$ | $<34$ | $>40$ |
| Continuity of the septum | Conplete or Incomplete | Absence or Incomplete | Complete |
| Budding type | Intratentacular | Continuous | Discontinuous |

TABLE 3. Diagnostic characteristics used in the identification of the species of the genus Siderastrea. SOURCE: The authors (2018).

Nevertheless, the methodology does not allow a clear identification due to the great overlap in most of the variables. Because only S. stellata is reported for Buzios, we considered all specimens from this site to belog to this particular species.

For texture analysis, another series of electronic microscopy photographs were taken from a subset of 63 colonies with a scanning electron microscopy with magnification of 11 x (covering more than one corallite) and three other pictures of individual corallites, with variable magnification in the 20 x range. Images were replicated 3 times per colony and scale (6 images per colony) totalling 378 micrographs. Finally 370 images were analyzed given that 8 of the photographs belonging to the original data had poor image quality to be used.


Figure 4. Quantitative measured variables. (a) corallite diameter ( $\mathrm{a}=$ larger diameter, $\mathrm{a}^{\prime}=$ small diameter), (b) first cycle septum thickness, (c) number of septa, (d) first cycle septum length, (e) columella diameter (b=larger diameter, $\mathrm{b}^{\prime}=$ small diameter). SOURCE: The authors (2018).

### 3.3. IMAGE ANALYSIS

We conducted a pilot evaluation of a number of different texture quantification strategies (data not shown). Completed Local Binary Patterns (CLBP), GUO, ZHANG AND ZHANG (2010) descriptors were the most efficient, as measured by the success in preliminary image classification using simple cross-validation analysis. CLBP is an extension of LBP (Local Binary Pattern - OJALA, PIETIKÄINEN and MÄENPÄÄ, 2002), a type of visual descriptor used for classification in texture-oriented "computer vision" (i.e. pattern recognition by computer software). LBP is based on "feature vectors", or $n$-dimensional numerical vectors that represent some object. Under this approach, feature vectors are computed according to the following steps, which make up the LBP code:

1. The image is divided in "cells", which are sections made up by a fixed number of pixels.
2. One pixel is selected and its shade of gray is compared to each of its closest neighbors. For simplicity, we will assume that 8 neighbors are selected. Because the electronic micrographs are black and white, the shades of gray are internally coded as integer numbers ranging from 0-255. If the center pixel value is lower than the value of its neighbor, the stored value is 1 , otherwise, it is 0 , (HASSANIEN et al., 2014). Comparisons are always made in the same direction, either clockwise or counter clockwise. Thus, the end results is a 8 -digit binary, which is normally converted to its decimal representation.
3. The procedure is repeated for every pixel in the cell;
4. A histogram is generated by counting the frequency of each number in the grid. By concatenating the histograms obtained for all the cells, one arrives at the feature vector of the whole window. This numeric representation of the image can now be used in image classification.

LBP has important properties, such as being insensitive to image rotation and scale, OJALA et al. (2000), thus allowing the combination of images taken from different viewpoints and at different scales into a single analysis. Differences in grayscale values are first computed as a vector, and these values are decomposed into their sign and magnitude components. Sign records if the difference between the center pixel and its neighbor is positive or negative. Magnitude is the absolute value of this difference. Sign vectors can be converted into feature vectors according to the steps outlined above. Finally, a third feature vector can be computed by employing the image's mean magnitude, instead of the average magnitude of the cell. Thus, CLBP generate three feature vectors per cell, which are joined in a single three-dimensional histogram (Fig.5). Because it is normalized with respect to the average image magnitude, this method generates descriptors which are also robust to variation in image luminosity. A formal description of this method is provided in Appendix A.


Figure 5. Flowchart indicating operation of the local binary pattern method. LDSMT: Local difference signmagnitude transform, S: sign component, M: magnitude component, CLBP_S, CLBP_M and CLBP_C: sign, magnitude and central operators respectively. SOURCE: adapted from (Guo et al., 2010 $\overline{0}$ ).

Feature vectors are numeric representations of the image, so they may be used as variables in multivariate statistics, such as discriminant analysis (DA). DA seeks to build linear functions that minimize the variance among data points associated with a user-defined class (in our case the class would be one of the three species) while maximizing the variance among the classes. These functions are normally built using part of the input variables, or the training data. The accuracy of these functions can then be evaluated by measuring its success in the assignment of the remaining data to the correct class. This is called cross-validation of the discriminant functions. The success in classification may be affected by the particular choice of the data used as training set, so cross validation is normally bootstrapped by generating a large number of discriminant functions from randomly sampled training data and then estimating the average classification success from the cross validation results obtained for each set of functions.

There are alternatives to this technique such as support vector machines (SVM), knearest neighbor algorithms (k-NN), CORTES and VAPNIK (1995) and Naive-Bayes, DUDA and HART (1973) classifiers. SVM follow a logic similar to DA in the sense that it uses part of the data to create the training set. K-NN uses a similar approach, but the algorithm assigns a data point to a labeled cluster based on the number of closest neighbors found in the multivariate space that already belong to that cluster. Finally, in naive Bayesian
classification, the posterior probability that image belong to a species is estimated in classical Bayesian fashion, by multiplying the prior probability of observing a feature vector by the likelihood function. Both the prior and the likelihood are estimated from the training set.

We evaluated the classification accuracy of all these strategies (see results) against a newly proposed method, based on artificial neural networks (ANN). ANNs seeks to emulate the workings of the human brain in pattern recognition, such as assigning a given image to its associated class. An ANN is made up of artificial neurons, simple elements that receive an input, change their internal state based on that input (i.e. are activated or not) and then generate an output based on input and activation. This output has a certain weight and it may or may not be generated depending on the value returned by the activation function. Artificial neurons are normally divided into input and output neurons and one or more intermediate layers of neurons, called hidden layers.

Differently from aforementioned classification strategies, the functions are not fixed, their outputs are modulated by the weights and the very emission of an output is decided by the activation function. These weights and the activation functions may be modified in time so the network is capable of "learning". For instance, a ANN may be trained to recognize images of cats by being fed a stream of "cat" and "not-cat" images from which learns to extracts features that are very particular to these animals (fur, whiskers, slit-like pupils, etc.) as opposed to other features that are not (e.g. the texture of the couch a cat may be lying on or the feathers of a bird from a "not-cat" image). In this context, the input of the network are the image descriptors (such as feature vectors) and the output is the class the image belongs to. In supervised training, the network is "told", at the end of the process, what is the correct class for the image. This process is called backpropagation. In one form of ANN the hidden layer is competitive, because neurons compete for the "right" to respond to a subset of input data, leading to neuron specialization, which is well suited to find clusters within data. The backpropagation approach is called multilayer perceptron, HAYKIN (1998) and it was also evaluated in this study with respect to its classification accuracy.

The learning process of such a network is similar to the process of recalling a memory when confronted with a certain visual pattern. For instance, when the brain is presented with a picture of a cat, it will retrieve the conceptual memory of the animal. On the other hand, when confronted with an unknown animal, humans can be sure that it is unknown precisely because they cannot associate any memory with that image. Hence, learning in pattern recognition can be defined as the process of building associative memory (AM) or the mapping of certain features to a concept or object. A more complicated version of associative
memory can be build when the association between image and concept is fuzzy. For instance, a person may be unsure if an image of a wild cat corresponds to a leopard or jaguar because it does not recall exactly the coat coloration patterns associated with each species.

Fuzzy set theory was introduced in 1965 by Lofti Zadeh, ZADEH (1965) and aims to provide a mathematical treatment for objects or concepts with uncertain or imprecise boundaries. This theory has been applied in several problems of many fields of knowledge such as chemistry, biology, engineering, etc. (PEDRYCZ; GOMIDE, 2007; YAGER; ZADEH, 2012). One of the reasons for its popularization is due the fact that it has been shown as an appropriate theory to model uncertain concepts described in natural language, PEDRYCZ and GOMIDE (2007) and GRAÑA, (2009. Since one or more parameters and variables involving in phenomenon and tasks may be uncertain or imprecise, several mathematical models and methods based on fuzzy set theory have been proposed by many researchers in the literature, (PEDRYCZ; GOMIDE, 2007; GRAÑA, 2009; KABURLASOS; KEHAGIAS, 2013). In classification problems, the identification of an appropriate label class may be uncertain and, in this case, each class can be represented by fuzzy sets. Thus, the classification task boils down to determining the membership degree of a given input in each class. The membership degree is called pertinence. It ranges between 0 and 1 but, unlike probabilities, it does not sum up to 1 . This means that the degree of pertinence of an element to a certain set (in our case, of a image to a certain species) may be 1.00 to $S$. radians, 0.60 to S. siderea and 0.55 to $S$. stellata .

The equivalent measure fuzzy associative memory (E-FAM), ESMI et al. (2014), ESMI, SUSSNER and SANDRI (2015) is a mathematical model that can be used to solve classification and regression problems. An E-FAM is an associative memory (AM), a model geared to associate input-output pairs called fundamental memories and it is also able to deal with noise and corrupted inputs (BUCKLEY; HAYASHI, 1994; VALLE; SUSSNER, 2008). In our case, these pairs are the feature vectors and the species the images they describe should belong to. E-FAMs belong to a general class of associative memories called $\Theta$-FAMs that can be viewed as fuzzy neural network with a competitive hidden layer (ESMI et al., 2014). In a fuzzy neural network, weights of the connections or synapses represent fuzzy sets (BUCKLEY; HAYASHI, 1994). In particular, the calculations of the hidden neurons of an EFAM are given in terms of equivalent measures that are nothing more than functions which evaluate the degree of similarity between two fuzzy sets, ZENG and LI (2006), BUSTINCE, PAGOLA and BARRENECHEA (2007), i.e., the feature vectors and the species associated with its corresponding image. Thus, $\Theta$-FAMs are highly suited classifiers for the type of
problem under consideration. They have been successfully applied in several benchmarks datasets available on the internet, SUSSNER et al (2012), ESMI et al. 2012, ESMI et al (2014), ESMI, SUSSNER and SANDRI (2015) and other real problems such as vision based self-localization of mobile robot, SUSSNER et al.(2012), and identification of speakers (ESMI et al., 2012). A formal description of $\Theta$-FAM is found in Appendix B.

Pertinence values for AM were obtained using 5 -fold cross-validation experiments. This is a non-exhaustive strategy in which the dataset (370 images) is randomly partitioned into subsamples. One subsample is retained as validation data and the remaining subsamples are used as the training set. The cross-validation is then repeated 4 more times, so that each subsample is used as validation data at least once.

In order to assess differences in pertinence values across species and image scales, we fit a log-linear model AGRESTI (1990) to the experiment data. Pertinence was treated as a binary response variable, being set to 0 if smaller than 1 . Results were non significant (loglikelihood ratio test $=1.64$, d.f. $=7, \mathrm{p}<0.98$ ), indicating that images at both scales were useful in this experiment and that no species had pertinence values significantly lower than the others.

## 4. RESULTS

From the analysis of images it was determined that 92 of the images belonged to specimens putatively identified as $S$. radians, 72 as $S$. siderea, and 206 S. stellata. In this first step, for CLBP, we employed grids of $24 \times 24$ pixels since these values produced the best experimental results. The 5 -fold cross-validation using $\Theta$-FAM produced a mean classification accuracy of $91 \%( \pm 5)$, which was higher than any other method, but only marginally so if compared do K-NN (TABLE 4).

| METHOD | ACCURACY (\%) |
| :--- | :---: |
| Multi-layer Perceptron | 78.6 |
| Naive-Bayes | 55.1 |
| K-Nearest-Neighbors (KNN) | 89.5 |
| Support Vector Machine (SVM) | 87.0 |
| Proposed method ( $\Theta$-FAM) | 91.0 |

TABLE 4. Comparative values of accuracy obtained through different methods. SOURCE: The authors (2018)

In all colonies, maximum pertinence matched the putative species for all images, with the exception of colony 284 , collected in Boipeba (BA). This colony was identified as $S$. stellata, but one image had the highest pertinence $S$. siderea. Results also reflect accurately the expected geographic distribution in Brazil, LEÃO, KIKUCHI and TESTA (2003) with the exception of the occurrence of 4 S. siderea colonies in Maraxanguape (RN).


Figure 6, (previous page). Maximum pertinence grouped by colony, putative species and geographic location. Results of $\Theta$-FAM cross-validation classification analysis of CLBP feature vectors. Pertinence values are color coded according to the legend. Images (rectangles) were grouped by colony, then by putative species and geographic location. Field numbers in the x -axis identify columns along and images are numbered according to the y-axis. Geographic locations (Armação dos Búzios, RJ; Aracruz, ES; Arquipélago dos Abrolhos BA; Ilha de Boipeba, BA; Tamandaré, PE and Maraxanguape, RN) are arranged in order of decreasing latitude, from left to right. A few micrographs were accidentally shot from areas of the skeleton that had been scraped during tissue collection for DNA analysis and hence were discarded (gray shaded rectangles). All images matched the colony's putative species, with the single exception of colony 284 , highlighted by the black box (see text). SOURCE: The authors (2018).

## 5. DISCUSSION

In this work, we demonstrated the usefulness of texture analysis combined with fuzzy neural network classifiers in species discrimination in the Atlantic Siderastrea complex. To our knowledge, this also is the first time that such state of the art image analysis technology has been applied to Scleractinian corals. This is also the first study of this problematic genus that has explicitly considered uncertainty in species identification by integrating it into the classification analysis through the use of fuzzy logic.

Contemporary authors (e.g Santos et. al, 2004; Menezes et al. 2013, 2014 and Garcia et al., 2017) have overtly or tacitly acknowledged the imprecise taxonomic limits by remarking on the overlap in quantitative diagnostic characters among species. We were unable to find other works on Siderastrea taxonomy in which the usefulness of the method to species discrimination was properly evaluated. Although all the aforementioned authors conducted canonical discriminant analysis (CDA) on their data, none ever attempted to assess the accuracy of the discriminant functions as classifiers through cross-validation experiments. Instead, they attempted to validate their results using some form of post-hoc multivariate analysis of the CDA scores. Besides having assumptions that are often hard to meet (such as multivariate normality, homogeneity of covariance matrices, sphericity, etc.) multivariate statistics such as MANOVA are designed to evaluate the size and significance of the factor's effect Izenman, IZENMAN (2008) and not classification accuracy. Additionally, authors often report significance and not the size of the effect i.e., how much the means corresponding to each species are apart in the multivariate space. It is well known in statistics that, if the
effect is small, identification will require large sample size, MONTGOMERY (2008) which increases the power of the test. A large $n$ is easily achieved in most morphometric studies hence the common situation (e.g. Menezes et al. 2014 - Fig. 2) of considerable overlap among data points belonging to different species/localities and yet the authors' claim that the analysis was successful because the p-values which are lower than the classical type I error probability cutoff (0.05).

Classification accuracy by $\Theta$-FAM was outstanding, particularly in face of the variation in luminosity and scale across different images and the uncertainty in image membership. This is remarkable given the large variation in diagnostic characters in the specimens that we considered to belong to $S$. stellata, that many times exceeded the overall range observed for other species and even the previously published ranges (see Appendix D). Log-linear analysis of the top cross-validation experiment suggests that this species does not significantly differs from others in level of image pertinence. These results suggest that CLBP has probably captured aspects from the coral images which are intrinsic to each of the three species, but are not well expressed by the quantification of traditional taxonomic characters. The approach outlined in this work seems promising for coral taxonomy, which has been strongly based on skeletal characters.

As classifiers, artificial neural networks have a strong advantage over other strategies which is their ability to learn. MITCHELL (1997) defined machine learning in the following terms "A computer program is said to learn from experience $E$ with respect to some class of tasks $T$ and performance measure $P$ if its performance at tasks in $T$, as measured by $P$, improves with experience $E . "$ In other words, the more images the network is exposed to, the more efficient it should become as a classifier. This opens new possibilities which are radically different from most self-contained, often irreproducible morphometric studies on corals published so far. The nearly ubiquity of electronic microscopy in most research institutes and the frequent use that coral taxonomists make of this technology suggests that there are terabytes of images stored in labs around the world that could be assembled in a large on-line database. Neural networks front-ended by web portal could thus be used to classify new images provided by interested users, expediting taxonomic work, especially in the case of untrained researchers. Although this may seem far-fetched, it is gradually becoming a reality in other areas such as medical imaging (e.g SHI et al., 2009; MITCHELL 1997). The goal is to provide the means to efficiently compare a newly-acquired image to millions of previously classified images of the same type, greatly improving the accuracy of
the diagnostics. The large availability of biological images in databases such as Morphobank (https://morphobank.org/) suggests that this reality may be closer than we suspect.

## 6. FUTURE WORK

It will be interesting to evaluate if $\Theta$-FAM provides accuracy superior to other supervised classification methods, including DCA, when using with the classical diagnostic characters compiled from this study. This should also put in perspective the accuracy gain afforded by image analysis algorithms, such as CLBP, as opposed to traditional morphometrics. As far as we now, we have the most extensive collection effort in a single project along the Brazilian coast, especially in the case of the endemic species S. stellata. Thus, we may conduct more extensive image sampling of our own collection, focusing on further detailing the geographic variation found in that particular species. By using collection localities as classes, we may actually detect clusters along the Brazilian coast that may or may not have share environmental conditions. These patterns could be back-correlated to variation in biologically meaningful characters, such as corallite size and spacing, which are well known to respond to environmental variation (e.g. BUDD, 1980; MILLER, 1994; MUKO et al., 2000).

## REFERENCES

AGRESTI, A. Categorical data analysis. New York: Wiley, 1990.
ALMY JR., C.; C. CARRIÓN-TORRES. Shallow-water stony corals of Puerto Rico. Caribbean Journal of Science 3(2-3):133-163, 1963.

BASAVAPRASAD, B.; RAVI, M. Study on the importance of image processing and its applications. International Journal of Research in Engineering and Technology,Vol.3, 155-160, 2014.

BIRKHOFF, G. Lattice Theory, 3rd edn. (American Mathematical Society, Prov- idence, 1993).

BLAINVILLE, H.M. Zoophytes. Dictionnaire des Sciences Naturelles, Paris 60:310-358, 1830.

BUCKLEY, J.; HAYASHI, Y. Fuzzy sets and systems 66, 1 (1994).
BUDD, A. F. Environmental variation in skeletal morphology within the Caribbean Reef Corals Montastraea annularis and Siderastrea siderea. Bull. Mar. Sci. 30: 678-790, 1980.
BUDD, A.; GUZMÁN, H. Siderastrea glynni, a new species of scleractinian coral (Cnidaria, Anthozoa) from the Eastern Pacific. Proc. Biol. Soc. Wash. 107: 591-599, 1994.

BUSTINCE, H.; BARRENECHEA, E.; PAGOLA, M. Fuzzy Sets and Systems 157, 2333 (2006).

BUSTINCE, H.; PAGOLA, M.; BARRENECHEA, E. Information Sciences 177, 906 (2007).

COLÍN GARCÍA, N.; CAMPOS, J.; TELLO MUSI, J.; FORSMAN, Z.; MONTERO MUÑOZ, J.; MONSALVO REYES, A.; ARIAS GONZÁLEZ, J. Comparative molecular and morphological variation analysis of Siderastrea (Anthozoa, Scleractinia) reveals the presence of Siderastrea stellata in the Gulf of Mexico. The Biological Bulletin, 232: 58-70, 2017.

CORTES, C.; VAPNIK, V. Support-Vector Networks. Machine Learning (1995) 20: 273.
DA SILVA, N.; DA SILVA OLIVEIRA, M.; DE ALMEIDA FILHO, H.; SOUZA PINHEIRO, L.; ROSSATTO, D.; KOLB, R., BRUNO, O. Leaf epidermis images for robust identification of plants. Sci. Rep. 6, 25994; doi: 10.1038/srep25994, 2016.
DEBROT, A.O.; KUENEN, M.M.C.E.; DEKKER, K. Recent declines in the coral fauna of the Spaanse Water, Curaçao, Netherlands Antilles. Bulletin of Marine Science 63(3): 571580, 1998.

DUDA, R.O.; HART, P.E. Pattern classification and scene analysis, Wiley, NewYork, 1973.

DUERDEN, J.E. West Indian madreporarian polyps. National Academy of Sciences 8: 403599, 1902.
ESMI, E.; SUSSNER, P.; BUSTINCE, H.; FERNÁNDEZ, J. IEEE Transactions on Fuzzy Systems (2014), accepted for publication, to appear.

ESMI, E.; SUSSNER, P.; SANDRI, S. Fuzzy Sets and Systems (2015).

ESMI, E., SUSSNER, P.;VALLE, M.E.; SAKURAY, F.; BARROS, L.C. Fuzzy associative memories based on subsethood and similarity measures with applications to speaker identification, in Hybrid Artificial Intelligent Systems, eds. E. Corchado, V. Snasel, A. Abraham, M. Wozniak, M. Graña and S.-B. Cho, Lecture Notes in Computer Science, Vol. 7209 (Springer Berlin / Heidelberg, 2012) pp. 479-490.
ESMI, E.; SUSSNER, P.; BUSTINCE, H.; FERNÁNDEZ, J. IEEE Transactions on Fuzzy Systems 23, 313 (2015).

ESMI, E.; SUSSNER, P.; SANDRI, S. Fuzzy Sets and Systems 292, 242 (2016).
FAN, J.; XIE, W.; PEI, J. Fuzzy Sets and Systems 106, 201(Sep. 1999).
FLORINDO, J. B.; DA SILVA, N.; DA SILVA OLIVEIRA, M.; ROMUALDO, L.; DA SILVA, F.; DE CERQUEIRA LUZ, P.; HERLING, V.; BRUNO, O. Brachiaria species identification using imaging techniques based on fractal descriptors. Comput. Electron. Agric. 103, 48-54, 2014.
FODOR, J.; ROUBENS, R. Fuzzy Preference Modelling and Multicriteria Decision Support, Fundamental Theories of Physics (Springer, 1994).
FORSMAN, Z.; GUZMAN, H.; CHEN, C.; FOX, G.; WELLINGTON, G. "An ITS Region Phylogeny of Siderastrea (Cnidaria: Anthozoa): Is S. Glynni Endangered or Introduced?" Coral Reefs 24 (2): 343-47, 2005.
FOSTER, A.B. Phenotypic plasticity in the reef corals Montastrea annularis (Ellis and Solander) and Siderastrea siderea (Ellis and Solander). Journal of Experimental Marine Biology and Ecology 39: 25-54, 1979.

FOSTER, A.B. Environmental variation in morphology within the Caribbean reef corals Montastrea annularis and Siderastrea siderea. Bulletin of Marine Science 30: 678-709, 1980.

GLYNN, P.; GRASSIAN, B.; KLEEMANN, K.; MATÉ, J. The true identity of Siderastrea glynni Budd \& Guzmán, 1994, a highly endangered eastern Pacific scleractinian coral. Coral Reefs (2016) 35:1399-1404.
GRAÑA, M. Neurocomputing 72, 2065 (2009).
GREGORY, J. Contributions to the Geology and Physical Geography of the West Indies. Quart. Jour. Geol. Soc. Loud., vol. LI, pp. 255 -310, PLXI, 1895.
GUO, Z.; ZHANG, L.; ZHANG, D. IEEE Transactions on Image Processing 19, 1657 (2010).

HASSANIEN, A.; TOLBA, M.; AZAR, A. Advanced machine learning technologies and applications, Springer, Cairo, Egypt, 2014.

IZENMAN, A.J. Modern Multivariate Statistical Techniques: Regression, Classification, and Manifold Learning Series. Springer Texts in Statistics. 2008, approx. 760 p..
KABURLASOS, V.; KEHAGIAS, A. Fuzzy Systems, IEEE Transactions on (2013), to appear, available online.
LABOREL, J. Madréporaires et hydrocoralliaires récifaux des côtes brésiliennes. Systématique, écologie, répartition verticale et géographique. Annales de l'Institut Océanographique 47: 15-229, 1969/70.

LEÃO, Z.; KIKUCHI, R.; TESTA, V. Corals and coral reefs of Brazil. Latin American Coral Reefs, Elsevier Science, 9-52, 2003.
LEÃO, Z.; KIKUCHI, R.; FERREIRA B.; NEVES, E.; SOVIERZOSKI, H.; OLIVEIRA M.; MAIDA, M.; CORREIA, M.; JOHNSSON, R. Brazilian coral reefs in a period of global change: A synthesis. Brazilian journal of oceanography, 64(sp2):97-116; 2016.
MARTINS-BEDÉ, F.; GODO, L.; SANDRI, S.; DUTRA, L.; FREITAS, C.; CARVALHO, O.; GUIMARÃES, R.; AMARAL, R. Classification of Schistoso-miasis prevalence using fuzzy case-based reasoning, in Proceedings of the 10th International Work-Conference on Artificial Neural Networks: Part I: Bio-Inspired Systems: Computational and Ambient Intelligence, IWANN '09 (Springer-Verlag, Berlin, Heidelberg, 2009).
MELÉNDEZ ACOSTA, J.; RIOS FIGUEROA, H.; MARÍN HERNÁNDEZ, A. Recuperador de imágenes en base a contenido: una aplicación con esporas. Departamento de Física e Inteligencia Artificial, 1-10, 2013

MENEZES N.; NEVES, E.; BARROS, F.; PAPA DE KIKUCHI, R.; JOHNSSON, R. Intracolonial variation in Siderastrea de Blainville, 1830 (Anthozoa, Scleractinia): taxonomy under challenging morphological constraints. Biota Neotrop, vol. 13, no.1, 2013.

MENEZES N.; NEVES, E.; PAPA DE KIKUCHI, R.; JOHNSSON, R. Morphological variation in the atlantic genus Siderastrea (Anthozoa, Scleractinia). Papéis Avulsos de Zoologia, 54(16):199-208, 2014.

MILLER, K. J. Morphological variation in the coral genus Platygyra: environmental influences and taxonomic implications. Mar. Ecol. Prog. Ser. 110: 19-28, 1994.

MILNE EDWARDS, H. Histoire naturelle des Coralliaires ou polypes proprement dits. Paris: Librairie Encyclopédique de Roret, II. 505-513, 1857.
MILNE EDWARDS, H.; HAIME, J. A monograph of the British fossil corals. Palaeontolographical Society of London 3(7): 1-71, 1850.

MITCHELL, T. Machine Learning. McGraw Hill. p. 2. ISBN 0-07-042807-7, 1997.
MONTGOMERY, D. C. Design and Analysis of Experiments. 7 edition. Hoboken, NJ: Wiley, 2008.
MUKO, S.; KAWASAKI, k.; SAKAI, k.; TAKASU, F.; SHIGESADA, N. Morphological plasticity in the coral Porites sillimaniani and its adaptive significance. Bull. Mar. Sci. 66: 225-239, 2000.

NEVES, E.; ANDRADE S.; DA SILVEIRA, F.; SOLFERINI, V. "Genetic Variation and Population Structuring in Two Brooding Coral Species (Siderastrea Stellata and Siderastrea Radians) from Brazil". Genetica 132 (3): 243-54, 2007. doi:10.1007/s10709-007-9168-z.

NEVES, E.; SILVEIRA, F; PICHON, M.; JOHNSSON, R. Cnidaria, Scleractinia, Siderastreidae, Siderastrea siderea (Ellis and Solander, 1786): Hartt Expedition and the first record of a Caribbean siderastreid in tropical Southwestern Atlantic. Check List 6: 505-510, 2010.

NUNES, F.; NORRIS, R.; KNOWLTON, N. "Long Distance Dispersal and Connectivity in Amphi-Atlantic Corals at Regional and Basin Scales." PLoS ONE 6 (7): e22298. doi:10.1371/journal.pone.0022298, 2011.

OJALA, T.; PIETIKÄINEN, M.; MÄENPÄÄ, T. Gray Scale and Rotation Invariant Texture Classification with Local Binary Patterns. In: Computer Vision - ECCV 2000. ECCV 2000. Lecture Notes in Computer Science, vol 1842. Springer, Berlin, Heidelberg, 2000.
OJALA, T.; PIETIKÄINEN, M.; MÄENPÄÄ, T. IEEE Transactions on Pattern Analysis and Machine Intelligence 24, 971 (2002).

PALLAS, P.S. Elenchus Zoophytorum. Hagae-Comitum: Varrentrapp F. 451 p., 1766.
PEDRYCZ, W.; GOMIDE, F. Fuzzy Systems Engineering: Towards Human-Centric Computing (Wiley, IEEE Press, New York, 2007).

REYES, J.; SANTODOMINGO, N.; FLÓREZ, P. Corales Escleractinios de Colombia. Invemar Serie de Publicaciones Especiales, No. 14. Santa Marta, 246 p., 2010.
SANTOS, M.; AMARAL, F.; BEZERRA DE SÁ, F.; ABAGE DE LIMA, M. 2004. Morphological variation of Montastrea cavernosa and Siderastrea stellata (Cnidaria: Scleractinia) from the states of Maranhão, Paraíba and Pernambuco, Brasil. Biologia Geral e Experimetal, São Cristóvão, SE 5 (1): 5-11.

SUSSNER, P.; ESMI, E.; VILLAVERDE, I.; GRAÑA, M. Journal of Mathematical Imaging and Vision 42, 134 (2012).

VALLE, M.E.; SUSSNER, P. Fuzzy Sets and Systems 159, 747 (2008).
VAUGHAN, T.W. 1919. Fossil corals from Central America, Cuba, and Porto Rico, with an account of the American Tertiary, Pleistocene, and recent coral reefs. Bull. Smith. Inst. 103: 189-524.

VERRILL, A.E. Notice of the corals and echinoderms collected by Prof. C.F. Hartt, at the Abrolhos Reefs, province of Bahia, Brazil, 1867. Transactions of Connecticut Academy of Arts and Science 1(2): 351-371, 1868.

YAGER, R.R.; ZADEH, L.A. An introduction to fuzzy logic applications in intelligent systems (Springer Science \& Business Media, 2012).

YONGE, C.M. Studies on biology of Tortugas corals. II. Variation in the genus Siderastrea. Publications of the Carnegie Institute of Washington 452: 201-208, 1935.

ZADEH, L.A. Information and Control 8, 338 (1965).
ZENG, W.; LI, H. International Journal of Intelligent Systems 21, 639 (2006).

## APPENDIX A - IMAGE ANALYSIS

Completed Local Binary Patterns (CLBP) descriptors essentially are a combination of different strategies to compute local binary features. The first and most well-known of such strategies are the classical Local Binary Patterns (LBP), (OJALA; PIETIKÄINEN; MÄENPÄÄ, 2002). Let $g$ be a binary image of $\mathrm{W} \times \mathrm{H}$ dimensions, the LBP code of the reference pixel $g_{c}:=g\left(i_{c}, j_{c}\right)$ is defined in terms of its $P$ neighbors in the radius $R$ :
where $g_{p}=g\left(i_{c}+R \cos (2 \pi p / P), j_{c}+R \sin (2 \pi p / P)\right), p=1, \ldots, P$, and $s(x)=0$ if $x \geq 0$, otherwise, $\mathrm{s}(\mathrm{x})=1$ if $\mathrm{x}<0$. The values of those points that fall outside the grid of pixels in the discrete domain of the image are obtained by interpolation. The LBP descriptors are given by the histogram of LBP codes:

$$
H(k)=\sum_{i=1}^{W} \sum_{J=1}^{H} \delta\left(L B P_{P, R}(i, j), k\right), \quad k \in[0, k],
$$

where $\delta(\mathrm{x}, \mathrm{y}) \in\{0,1\}$ with $\delta(\mathrm{x}, \mathrm{y})=1 \Leftrightarrow \mathrm{x}=\mathrm{y}$ and K is the maximum value assigned to an LBP code. Another important definition in LBP theory is the $U$ value, corresponding to transitions between bits 0 and 1 in the LBP code:

$$
U\left(L B P_{P, R}\right)=\mathrm{I} s\left(g_{p-1}-g_{c}\right)-s\left(g_{0}-g_{c}\right) \mathrm{I}+\sum_{P=1}^{P-1} \mathrm{I} s\left(g_{p}-g_{c}\right)-s\left(g_{p-1}-g_{c}\right) \mathrm{I} .
$$

Given this, the locally rotation invariant binary pattern is defined for each point $\left(\mathrm{i}_{\mathrm{c}}\right.$, jc) by

$$
L B P_{P, R}^{r u m_{2} 2}\left(i_{c}, j_{c}\right)=\left\{\begin{array}{c}
\sum_{P^{-1}} s\left(g_{p}-g_{c}\right), \text { if } U\left(L B P_{P, R}\right) \leq 2  \tag{2}\\
P=0 \\
P+\text { otherwise }
\end{array}\right.
$$

The same circular neighborhood can also provide other interesting features. An important example is the local difference vector $\left[d_{0}, \cdots, d_{p}-1\right]$, where $d_{p}=g_{p}-g_{c}$, for $0 \leq p$
$\leq \mathrm{P}-1$. An important property of $\mathrm{d}_{\mathrm{p}}$ is its robustness to illumination changes. Besides, it can be decomposed into a sign $\left(\mathrm{s}_{\mathrm{p}}\right)$ and a magnitude $\left(\mathrm{m}_{\mathrm{p}}\right)$ component $\mathrm{d}_{\mathrm{p}}=\mathrm{s}_{\mathrm{p}} * \mathrm{~m}_{\mathrm{p}}$, where $\mathrm{s}_{\mathrm{p}}=$ $\operatorname{sign}\left(d_{p}\right)$ and $m_{p}=\left|d_{p}\right|$. This operation provides us with the sign vector $\left[\mathrm{s}_{0}, \cdots, s_{p}-1\right]$ and the magnitude vector $\left[m_{0}, \cdots, m_{P}-1\right]$. These vectors can also give rise to local codes in a similar manner to that employed in the classical LBP method. Those local codes are named completed LBP (CLBP), (GUO; ZHANG; ZHANG, 2010). The magnitude vector gives rise to the CLBP M code:

$$
C L B P_{-} C P, R=\sum_{P=0}^{P-1} T\left(t\left(m_{p}, c\right) 2^{p},\right.
$$

where $\mathrm{t}(\mathrm{x}, \mathrm{c})$ is a threshold function: $\mathrm{t}(\mathrm{x}, \mathrm{c})=1$ is $\mathrm{x} \geq \mathrm{c}$ and $\mathrm{t}(\mathrm{x}, \mathrm{c})=0$ otherwise, where $c$ is the mean value of $m_{p}$ over the whole image. Similarly, one can define CLBP S over the $s_{p}$ vector, but this coincide with the classical LBP code defined in (1). Finally, we have the CLBP C code, generated by the gray value of the reference pixel $\left(\mathrm{g}_{\mathrm{c}}\right)$ :

CLBP_C $_{P, R}=\mathrm{t}\left(\mathrm{g}_{\mathrm{c}}, \mathrm{c}_{1}\right)$,
where $c_{I}$ is the average gray level of the entire image.
CLBP M, CLBP S and CLBP C can be summarized by histograms like in the classical LBP descriptors and those histograms can be combined in two ways: by concatenation or in a three-dimensional joint histogram. Here we adopt the second strategy, which yields the best results in most scenarios.

## APPENDIX B - FUZZY SUPERVISED CLASSIFICATION

An associative memory (AM) is a mapping $\Phi: \mathrm{X} \rightarrow \mathrm{Y}$ designed to store a set of pairs of data $\mathrm{M}=\left\{\left(\mathrm{X}^{\xi}, \mathrm{y}^{\xi}\right) \in \mathrm{X} \times \mathrm{Y} \mid \xi=1, \ldots, \mathrm{p}\right\}$ called fundamental memory set or the set of fundamental memories. Ideally, an AM $\Phi$ satisfies $\Phi\left(\mathrm{x}^{\xi}\right)=\mathrm{y}^{\xi}$ for $\mathrm{i}=1, \ldots, \mathrm{p}$ and additionally is endowed with a certain type of tolerance: $\Phi\left(\mathrm{x}^{\sim}\right)=\mathrm{y}^{\xi}$ if $\mathrm{x}^{\sim}$ stands for a corrupted or noisy versions of $x \xi$. In practice, many AM models are not able to store all fundamental memories and present limited correction capacity, retrieving approximately y ${ }^{\xi}$ for corrupted or noisy versions of $\mathrm{x}^{\xi}$. A fuzzy associative memory (FAM) is AM that is also a fuzzy neural network, BUCKLEY and HAYASHI (1994) i.e., an artificial neural network whose inputs or weights are fuzzy. $\Theta$-fuzzy associative memories ( $\Theta$-FAMs) consist of a subclass of fuzzy associative memories having a competitive hidden layer whose the calculation of the $\xi$ th hidden neuron is given by a function $\Theta^{\xi}: L \rightarrow[0,1]$, where the symbol L denotes a bounded lattice, that is, a partial ordered set with maximal and minimal elements such that the infimum and supremum of any two elements of $L$ exist and belong to $L$ (BIRKHOFF, 1993). Particular cases of $\Theta^{\xi}$ functions are given by fuzzy subsethood or equivalent measures, leading to (weighted) subsethood, dual subsethood, and equivalent measure FAMs (SUSSNER et al., 2012; ESMI et al, 2016).

Given a finite set $\left\{\left(\mathrm{x}^{\xi}, \mathrm{B}^{\xi}\right) \in \mathrm{L} \times \mathrm{F}(\mathrm{Y}): \xi=1, \ldots, \mathrm{p}\right\}$ where L is a bounded lattice and $F(Y)$ denotes the class of fuzzy sets of an arbitrary universe $Y$. Let $\Theta^{\xi} \mathrm{L} \rightarrow[0,1]$ be functions such that $\Theta^{\xi}\left(\mathrm{x}^{\xi}\right)=1$ for $\xi=1, \ldots, \mathrm{p}$ and $\mathrm{v} \in \mathrm{R}^{\mathrm{p}}$, the $\Theta$-FAM based on $\Theta^{\xi}$ and v , for short $\Theta-F A M$, is a function $\mathrm{O}: \mathrm{L} \rightarrow \mathrm{F}(\mathrm{Y})$ defined for each $\mathrm{x} \in \mathrm{L}$ in (ESMI et al., 2015; ESMI et al., 2016).

$$
\begin{equation*}
O(x)=\bigcup_{j \in i v(x)} B^{j}, \tag{3}
\end{equation*}
$$

where $I_{v}(x)=j \in\{1, \ldots, p\}: v j \Theta^{j}(x)=\max _{\xi=1, \ldots, p} \mathrm{v}_{\xi} \Theta^{\xi}(x)$. Sufficient conditions for $\mathrm{O}\left(\mathrm{x}^{\xi}\right)=\mathrm{B}^{\xi}, \xi=1, \ldots, \mathrm{p}$, and a characterization of the basins of attraction around each $\mathrm{x}^{\xi}$ can be found in (ESMI et al., 2015).

In this paper, the focus is on equivalent measure FAMs (E-FAMs). Recall that a equivalent measure on a bounded lattice $L$ is a function $E: L^{2} \rightarrow[0,1]$ that satisfies the following conditions: (FODOR; ROUBENS, 1994; BUSTINCE; PAGOLA; BARRENECHEA, 2007; ESMI; SUSSNER; SANDRI, 2016).

E1) $E(x, y)=E(y, x)$ for all $x, y \in L$;
E2) $\mathrm{E}\left(0_{\mathrm{L}}, 1_{\mathrm{L}}\right)=0$;
E3) $E(x, x)=1$ for all $x \in L$;
E4) if $x \leq y \leq z$, then $E(x, z) \leq E(x, y)$ and $E(x, z) \leq E(y, z)$.

Let $\mathrm{E}^{\xi}$ be equivalent functions on $L$, the corresponding E-FAM is obtained by taking $\Theta^{\xi}(x)=E^{\xi}\left(x^{\xi}, x\right)$ for each $\xi=1, \ldots, p$.

Let $L=\left[a_{1}, b_{1}\right] \times \ldots \times\left[a_{n}, b_{n}\right], a_{i}, b_{i} \in R$ with $a_{i} \leq b_{i}$ for $i=1, \ldots, n$, and let $\lambda \in(0$, $1]^{\mathrm{n}}$ and $\mathrm{w} \in[0,1]^{\mathrm{n}}$ such that $\sum_{i=1}^{n} w_{i}=1$

The function $\mathrm{E}_{\lambda, \mathrm{w}}: \mathrm{L}^{2} \rightarrow[0,1]$ given by

$$
\begin{equation*}
E \lambda, w(x, y)=\sum_{i=1}^{n} w i \max \left(0,1-\frac{|x i-y i|}{\lambda i|b i-a i|}\right), \forall x, y \in L \tag{4}
\end{equation*}
$$

is a equivalent measure on L. (MARTINS-BEDÉ et al., 2009; ESMI; SUSSNER; SANDRI, 2016). Other examples of equivalent measures can be found in FAN, XIE and PEI (1999), BUSTINCE, BARRENECHEA and PAGOLA (2006), BUSTINCE, PAGOLA and BARRENECHEA (2007), ESMI et al. (2015), ESMI, SUSSNER and SANDRI (2016) and references therein.

In order to obtain a method to recognize the species that a given coral image it is assumed that each input image has a fixed size and is a photo of one of the species Siderastrea siderea (SD), Siderastrea stellata (SS), and Siderastrea radians (SR).

Moreover, it is supposed that we have at hand a set of p images of coral that were labelled in one of these three species by an expert. Under this hypotheses, the strategy to identify the coral species for a given coral image is composed of two steps. The first one consists of applying a method of image analysis that is scale and translation invariant, such as CLBP, since the photos may be taken from different positions and distances. Thus, the first step consists of associating each image with a feature vector or image descriptors, that is, with a point of $L=\left[a_{1}, b_{1}\right] \times \ldots \times\left[a_{n}, b_{n}\right], a_{i}, b_{i} \in R$ with $a_{i}<b_{i}$ for $i=1, \ldots, n$, where $n$ is the number of descriptors or features extracted using CLBP method. The second step consists of designing a E-FAM geared to associate each feature vector to one of these three species. To
this end, is considered $Y=\{S D, S S, S R\}$ such that each class label $y \in Y$ is associated with the fuzzy number of $\mathrm{F}(\mathrm{Y})$ given by the characteristic function of $\{\mathrm{y}\}$. Thus, the given labelled coral images leads us to a set of p fundamental memories that we use to obtain an E-FAM.

## APPENDIX C－MEASUREMENTS ON THE COLONIES

| COLONY |  |  | CORALLITE |  |  | SEPTUM |  |  | CYCLE |  |  |  |  |  |  |  |  | COLUMELLA |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| $\begin{aligned} & \stackrel{\rightharpoonup}{0} \\ & \text { Z } \\ & \text { Z } \end{aligned}$ |  |  | $\begin{aligned} & \dot{0} \\ & \text { 号 } \\ & \bar{Z} \end{aligned}$ |  |  | $\begin{aligned} & \text { च } \\ & \text { 号 } \\ & \vdots \end{aligned}$ |  |  | $\begin{aligned} & 0 \\ & \frac{0}{0} \\ & 0 . \\ & 0 . \\ & 0 . \\ & 0 \\ & 0 \\ & 0 \\ & 0 \\ & 0 \\ & 0 \end{aligned}$ | 0 0 0 0 0 0 0 0 0 0 0 0 0 0 | $\begin{aligned} & \stackrel{\rightharpoonup}{0} \\ & \stackrel{0}{6} \\ & \stackrel{\text { an }}{2} \end{aligned}$ | 0 0 0 0 0 0 0 0 0 0 0 0 |  | $\begin{aligned} & \overrightarrow{\ddot{U}} \\ & \stackrel{0}{4} \\ & \stackrel{y}{2} \end{aligned}$ |  |  | $\begin{aligned} & \stackrel{\ddot{訁}}{\ddot{0}} \\ & \stackrel{\rightharpoonup}{4} \end{aligned}$ |  |  |
| 3 | N | S | 1 | 4.172 | 3 | 60 | 0.116 | 1.651 |  | X |  |  | X |  |  | X |  | 0.3315 | 2 |
|  |  |  | 2 | 4.281 | 4 | 60 | 0.139 | 1.575 |  | X |  |  | X |  |  | X |  | 0.321 | 2 |
|  |  |  | 3 | 4.187 | 4 | 58 | 0.13 | 1.539 |  | X |  |  | X |  |  |  | X | 0.4395 | 4 |
|  |  |  | 4 | 4.143 | 4 | 58 | 0.154 | 1.693 |  | X |  |  | X |  |  |  | X | 0.3525 | 1 |
|  |  |  | 5 | 3.972 | 3 | 52 | 0.164 | 1.752 |  | X |  |  | X |  |  |  | X | 0.307 | 2 |
| 4 | N | S | 1 | 2.618 | 6 | 32 | 0.16 | 1.368 |  | X |  |  |  | X |  |  | X | 0.3655 | 3 |
|  |  |  | 2 | 3.008 | 5 | 36 | 0.169 | 1.021 |  | X |  |  |  | X |  |  | X | 0.37 | 5 |
|  |  |  | 3 | 3.153 | 4 | 38 | 0.181 | 1.094 |  | X |  |  |  | X |  |  | X | 0.3765 | 2 |
|  |  |  | 4 | 3.489 | 4 | 42 | 0.188 | 1.458 |  | X |  |  |  | X |  |  | X | 0.4245 | 4 |
|  |  |  | 5 | 2.944 | 4 | 38 | 0.172 | 1.267 |  | X |  |  |  | X |  |  | X | 0.461 | 6 |
| 5 | N | N | 1 | 3.98 | 4 | 44 | 0.196 | 1.235 |  | X |  |  |  | X |  |  | X | 0.539 | 6 |
|  |  |  | 2 | 3.341 | 4 | 38 | 0.194 | 1.318 |  | X |  |  |  | X |  |  | X | 0.4395 | 2 |
|  |  |  | 3 | 3.63 | 5 | 38 | 0.217 | 1.479 |  | X |  |  |  | X |  |  | X | 0.435 | 2 |
|  |  |  | 4 | 3.341 | 3 | 44 | 0.195 | 1.289 |  | X |  |  |  | X |  |  | X | 0.498 | 4 |
|  |  |  | 5 | 3.692 | 5 | 44 | 0.185 | 1.363 |  | X |  |  | X |  |  |  | X | 0.532 | 5 |
| 10 | N | S | 1 | 3.805 | 4 | 42 | 0.213 | 1.304 |  | X |  |  |  | X |  |  | X | 0.541 | 3 |
|  |  |  | 2 | 3.701 | 4 | 40 | 0.209 | 1.07 |  | X |  |  |  | X |  |  | X | 0.551 | 5 |
|  |  |  | 3 | 3.655 | 4 | 38 | 0.215 | 1.346 |  | X |  |  |  | X |  |  | X | 0.5095 | 5 |
|  |  |  | 4 | 3.567 | 4 | 36 | 0.183 | 1.336 |  | X |  |  |  | X |  |  | X | 0.523 | 2 |
|  |  |  | 5 | 3.095 | 4 | 38 | 0.19 | 1.034 |  | X |  |  |  | X |  |  | X | 0.4065 | 2 |
| 12 | N | S | 1 | 3.452 | 6 | 48 | 0.205 | 1.634 |  | X |  |  | X |  |  |  | X | 0.307 | 2 |
|  |  |  | 2 | 3.187 | 5 | 48 | 0.145 | 1.327 |  | X |  |  | X |  |  |  | X | 0.326 | 1 |
|  |  |  | 3 | 3.989 | 6 | 42 | 0.178 | 1.269 |  | X |  |  | X |  |  |  | X | 0.3245 | 2 |
|  |  |  | 4 | 3.539 | 6 | 50 | 0.142 | 1.748 | X |  |  |  | X |  |  |  | X | 0.254 | 1 |
|  |  |  | 5 | 3.774 | 6 | 48 | 0.143 | 1.731 | X |  |  |  |  | X |  |  | X | 0.3135 | 1 |
| 17 | N | S | 1 | 2.651 | 3 | 38 | 0.192 | 1.025 |  | X |  |  |  | X |  |  | X | 0.409 | 1 |
|  |  |  | 2 | 4.562 | 4 | 46 | 0.228 | 1.924 |  | X |  |  | X |  |  |  | X | 0.428 | 1 |
|  |  |  | 3 | 3.498 | 6 | 44 | 0.177 | 1.404 |  | X |  |  |  | X |  |  | X | 0.4565 | 1 |
|  |  |  | 4 | 3.653 | 5 | 48 | 0.149 | 1.333 |  | X |  |  | X |  |  |  | X | 0.6215 | 2 |
|  |  |  | 5 | 3.681 | 5 | 46 | 0.178 | 1.35 |  | X |  |  |  | X |  |  | X | 0.5955 | 2 |
| 18 | N | S | 1 | 3.417 | 3 | 40 | 0.179 | 1.14 |  | X |  |  |  | X |  |  | X | 0.5515 | 4 |
|  |  |  | 2 | 3.454 | 4 | 41 | 0.173 | 1.059 |  | X |  |  |  | X |  |  | X | 0.5385 | 2 |
|  |  |  | 3 | 4.239 | 6 | 46 | 0.121 | 1.092 |  | X |  |  | X |  |  |  | X | 0.558 | 5 |
|  |  |  | 4 | 2.919 | 4 | 41 | 0.178 | 1.394 |  | X |  |  |  | X |  |  | X | 0.5265 | 3 |
|  |  |  | 5 | 3.465 | 5 | 42 | 0.184 | 1.097 |  | X |  |  |  | X |  |  | X | 0.5165 | 2 |
| 20 | N | S | 1 | 3.633 | 4 | 50 | 0.154 | 1.454 | X |  |  |  | X |  |  |  | X | 0.48 | 2 |
|  |  |  | 2 | 3.459 | 6 | 48 | 0.158 | 1.145 | X |  |  |  |  | X |  |  | X | 0.5745 | 2 |




|  |  |  | 5 | 3.293 | 4 | 32 | 0.227 | 1.294 |  | X |  |  | X | X | 0.6035 | 3 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 108 | N | S | 1 | 3.405 | 4 | 30 | 0.267 | 1.053 |  | X |  |  | X | X | 0.5145 | 3 |
|  |  |  | 2 | 3.601 | 4 | 32 | 0.253 | 1.072 |  | X |  |  | X | X | 0.584 | 3 |
|  |  |  | 3 | 3.279 | 3 | 34 | 0.226 | 1.075 |  | X |  | X |  | X | 0.5955 | 5 |
|  |  |  | 4 | 3.786 | 5 | 30 | 0.258 | 1.226 |  | X |  |  | X | X | 0.6835 | 4 |
|  |  |  | 5 | 3.114 | 4 | 26 | 0.27 | 0.981 |  | X |  |  | X | X | 0.419 | 4 |
| 109 | N | S | 1 | 3.483 | 3 | 28 | 0.306 | 1.128 |  | X |  |  | X | X | 0.6765 | 1 |
|  |  |  | 2 | 3.287 | 4 | 32 | 0.209 | 1.267 |  | X |  |  | X | X | 0.5195 | 2 |
|  |  |  | 3 | 3.589 | 3 | 30 | 0.28 | 1.294 |  | X |  |  | X | X | 0.6145 | 1 |
|  |  |  | 4 | 3.191 | 3 | 24 | 0.368 | 1.145 |  |  | X |  | X | X | 0.608 | 1 |
|  |  |  | 5 | 3.881 | 4 | 36 | 0.223 | 1.253 |  | X |  |  | X | X | 0.587 | 1 |
| 110 | N | S | 1 | 4.057 | 5 | 34 | 0.287 | 1.335 |  | X |  |  | X | X | 0.583 | 1 |
|  |  |  | 2 | 3.972 | 5 | 30 | 0.273 | 1.17 |  | X |  |  | X | X | 0.628 | 5 |
|  |  |  | 3 | 4.167 | 5 | 38 | 0.269 | 1.358 |  | X |  |  | X | X | 0.7325 | 5 |
|  |  |  | 4 | 4.118 | 5 | 34 | 0.307 | 1.359 |  | X |  |  | X | X | 0.6175 | 3 |
|  |  |  | 5 | 3.83 | 5 | 28 | 0.352 | 1.296 |  | X |  |  | X | X | 0.6285 | 1 |
| 111 | S | S | 1 | 3.155 | 4 | 28 | 0.272 | 1.013 |  | X |  |  | X | X | 0.465 | 2 |
|  |  |  | 2 | 3.365 | 6 | 36 | 0.224 | 1.215 |  | X |  |  | X | X | 0.514 | 3 |
|  |  |  | 3 | 3.004 | 6 | 30 | 0.221 | 1.061 |  | X |  |  | X | X | 0.4575 | 2 |
|  |  |  | 4 | 2.735 | 4 | 24 | 0.237 | 0.611 |  |  | X |  | X | X | 0.415 | 2 |
|  |  |  | 5 | 2.469 | 2 | 24 | 0.238 | 0.977 |  |  | X |  | X | X | 0.447 | 2 |
| 112 | S | S | 1 | 3.695 | 5 | 40 | 0.201 | 1.199 |  | X |  |  | X | X | 0.713 | 6 |
|  |  |  | 2 | 2.969 | 5 | 28 | 0.248 | 1.412 |  | X |  |  | X | X | 0.6545 | 3 |
|  |  |  | 3 | 4.102 | 4 | 38 | 0.271 | 1.427 |  | X |  | X |  | X | 0.6295 | 4 |
|  |  |  | 4 | 3.365 | 6 | 28 | 0.234 | 0.76 |  | X |  |  | X | X | 0.6225 | 3 |
|  |  |  | 5 | 3.604 | 5 | 30 | 0.3 | 1.139 |  | X |  |  | X | X | 0.611 | 5 |
| 113 | N | S | 1 | 3.247 | 3 | 24 | 0.256 | 1.088 |  |  | X |  | X | X | 0.596 | 3 |
|  |  |  | 2 | 3.495 | 4 | 30 | 0.238 | 1.25 |  | X |  |  | X | X | 0.6545 | 3 |
|  |  |  | 3 | 3.786 | 3 | 30 | 0.303 | 1.002 |  | X |  |  | X | X | 0.6835 | 3 |
|  |  |  | 4 | 3.901 | 4 | 34 | 0.279 | 1.267 |  | X |  |  | X | X | 0.6225 | 4 |
|  |  |  | 5 | 3.589 | 5 | 34 | 0.263 | 1.219 |  | X |  |  | X | X | 0.5765 | 4 |
| 114 | N | S | 1 | 3.309 | 3 | 26 | 0.469 | 1.19 |  | X |  |  | X | X | 0.576 | 2 |
|  |  |  | 2 | 4.017 | 4 | 30 | 0.303 | 1.401 |  | X |  |  | X | X | 0.5645 | 1 |
|  |  |  | 3 | 3.921 | 4 | 36 | 0.281 | 1.462 |  | X |  |  | X | X | 0.626 | 1 |
|  |  |  | 4 | 3.998 | 5 | 36 | 0.227 | 1.444 |  | X |  |  | X | X | 0.6665 | 1 |
|  |  |  | 5 | 3.627 | 5 | 34 | 0.301 | 1.248 |  | X |  |  | X | X | 0.5915 | 1 |
| 115 | N | S | 1 | 3.105 | 3 | 36 | 0.252 | 0.985 |  | X |  |  | X | X | 0.6805 | 4 |
|  |  |  | 2 | 3.386 | 3 | 34 | 0.213 | 1.398 |  | X |  |  | X | X | 0.705 | 3 |
|  |  |  | 3 | 3.303 | 3 | 36 | 0.224 | 1.383 |  | X |  |  | X | X | 0.543 | 5 |
|  |  |  | 4 | 3.241 | 3 | 28 | 0.301 | 0.887 |  | X |  |  | X | X | 0.641 | 2 |
|  |  |  | 5 | 3.63 | 4 | 34 | 0.254 | 0.953 |  | X |  |  | X | X | 0.5775 | 3 |
| 116 | N | S | 1 | 3.498 | 4 | 34 | 0.287 | 1.19 |  | X |  |  | X | X | 0.556 | 4 |
|  |  |  | 2 | 3.417 | 4 | 32 | 0.265 | 1.25 |  | X |  |  | X | X | 0.4915 | 5 |
|  |  |  | 3 | 3.886 | 3 | 36 | 0.213 | 1.336 |  | X |  | X |  | X | 0.472 | 3 |
|  |  |  | 4 | 3.196 | 4 | 30 | 0.237 | 1.085 |  | X |  |  | X | X | 0.597 | 3 |
|  |  |  | 5 | 3.172 | 3 | 26 | 0.222 | 1.131 |  | X |  |  | X | X | 0.536 | 2 |
| 120 | S | S | 1 | 3.479 | 2 | 52 | 0.165 | 1.176 | X |  |  | X |  | X | 0.5935 | 7 |
|  |  |  | 2 | 2.891 | 2 | 40 | 0.17 | 1.146 |  | X |  |  | X | X | 0.414 | 2 |
|  |  |  | 3 | 3.145 | 3 | 44 | 0.172 | 1.174 |  | X |  |  | X | X | 0.588 | 7 |
|  |  |  | 4 | 2.909 | 2 | 34 | 0.178 | 0.818 |  | X |  |  | X | X | 0.77 | 7 |
|  |  |  | 5 | 2.922 | 2 | 36 | 0.166 | 0.914 |  | X |  |  | X | X | 0.6065 | 7 |







| $\mathbf{3 2 6}$ | N | S | $\mathbf{1}$ | 4.011 | 4 | 38 | 0.246 | 1.244 |  | X |  |  |  | X |  |  | X | 0.3885 | 1 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  | $\mathbf{2}$ | 3.18 | 4 | 42 | 0.2 | 1.014 |  | X |  |  |  | X |  |  | X | 0.394 | 2 |
|  |  |  | $\mathbf{3}$ | 3.563 | 5 | 48 | 0.127 | 1.372 | X |  |  |  | X |  |  |  | X | 0.4715 | 2 |
|  |  |  | $\mathbf{4}$ | 2.799 | 4 | 32 | 0.173 | 0.834 |  | X |  |  |  | X |  |  | X | 0.3235 | 1 |
|  |  |  | $\mathbf{5}$ | 2.959 | 5 | 42 | 0.168 | 0.992 |  | X |  |  |  | X |  |  | X | 0.404 | 2 |
| $\mathbf{3 2 9}$ | N | S | $\mathbf{1}$ | 3.6 | 5 | 44 | 0.155 | 0.757 |  | X |  |  |  | X |  |  | X | 0.5495 | 6 |
|  |  |  | $\mathbf{2}$ | 2.57 | 3 | 34 | 0.15 | 0.953 |  | X |  |  |  | X |  |  | X | 0.3235 | 4 |
|  |  |  | $\mathbf{3}$ | 2.777 | 4 | 42 | 0.157 | 0.814 |  | X |  |  |  | X |  |  | X | 0.414 | 4 |
|  |  |  | $\mathbf{4}$ | 3.465 | 5 | 46 | 0.146 | 1.264 |  | X |  |  |  | X |  |  | X | 0.704 | 7 |
|  |  |  | $\mathbf{5}$ | 3.028 | 4 | 42 | 0.169 | 0.987 |  | X |  |  | X |  |  |  | X | 0.515 | 5 |
| $\mathbf{3 3 4}$ | N | S | $\mathbf{1}$ | 3.948 | 5 | 44 | 0.206 | 1.325 |  | X |  |  |  | X |  |  | X | 0.5275 | 3 |
|  |  |  | $\mathbf{2}$ | 3.712 | 6 | 42 | 0.179 | 1.22 |  | X |  |  | X |  |  |  | X | 0.631 | 6 |
|  |  |  | $\mathbf{3}$ | 3.701 | 6 | 38 | 0.15 | 1.526 |  | X |  |  | X |  |  |  | X | 0.5185 | 5 |
|  |  |  | $\mathbf{4}$ | 3.795 | 6 | 42 | 0.158 | 1.371 |  | X |  |  |  | X |  |  | X | 0.4005 | 6 |
|  |  |  | $\mathbf{5}$ | 3.112 | 5 | 38 | 0.16 | 1.34 |  | X |  |  |  | X |  |  | X | 0.464 | 5 |
| $\mathbf{3 3 6}$ | N | S | $\mathbf{1}$ | 3.186 | 4 | 44 | 0.129 | 1.191 |  | X |  |  |  | X |  |  | X | 0.6655 | 6 |
|  |  |  | $\mathbf{2}$ | 3.379 | 3 | 44 | 0.126 | 1.021 |  | X |  |  |  | X |  |  | X | 0.5815 | 6 |
|  |  |  | $\mathbf{3}$ | 2.873 | 3 | 40 | 0.128 | 0.992 |  | X |  |  |  | X |  |  | X | 0.557 | 7 |
|  |  |  | $\mathbf{4}$ | 3.875 | 5 | 48 | 0.19 | 1.738 |  | X |  |  | X |  |  |  | X | 0.467 | 4 |
|  |  |  | $\mathbf{5}$ | 3.502 | 5 | 48 | 0.145 | 0.889 | X |  |  |  |  | X |  |  | X | 0.5405 | 6 |

## APPENDIX D - BOXPLOT GRAPH OF MEASURED VARIABLES



Corallite diameter


First cycle septum thickness


Number of septa


Columella larger diameter


Columella smaller diameter


## APPENDIX E - SIDERASTREA COLONIES COLLECTION

| LEOM | 3 | LEOM | 20 | LEOM | 44 | LEOM | 102 | LEOM | 113 | LEOM | 125 | LEOM | 135 | LEOM | 176 |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| LEOM | 4 | LEOM | 23 | LEOM | 46 | LEOM | 103 | LEOM | 114 | LEOM | 126 | LEOM | 136 | LEOM | 177 |
| LEOM | 5 | LEOM | 25 | LEOM | 54 | LEOM | 104 | LEOM | 115 | LEOM | 127 | LEOM | 137 | LEOM | 179 |
| LEOM | 9 | LEOM | 27 | LEOM | 55 | LEOM | 105 | LEOM | 116 | LEOM | 128 | LEOM | 160 | LEOM | 182 |
| LEOM | 10 | LEOM | 34 | LEOM | 59 | LEOM | 107 | LEOM | 117 | LEOM | 129 | LEOM | 165 | LEOM | 186 |
| LEOM | 11 | LEOM | 38 | LEOM | 60 | LEOM | 108 | LEOM | 120 | LEOM | 130 | LEOM | 168 | LEOM | 192 |
| LEOM | 12 | LEOM | 39 | LEOM | 71 | LEOM | 109 | LEOM | 121 | LEOM | 131 | LEOM | 170 | LEOM | 193 |
| LEOM | 13 | LEOM | 40 | LEOM | 88 | LEOM | 110 | LEOM | 122 | LEOM | 132 | LEOM | 171 | LEOM | 195 |
| LEOM | 17 | LEOM | 42 | LEOM | 89 | LEOM | 111 | LEOM | 123 | LEOM | 133 | LEOM | 172 | LEOM | 196 |
| LEOM | 18 | LEOM | 43 | LEOM | 101 | LEOM | 112 | LEOM | 124 | LEOM | 134 | LEOM | 175 | LEOM | 197 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| LEOM | 199 | LEOM | 214 | LEOM | 247 | LEOM | 279 | LEOM | 303 | LEOM | 326 | LEOM | 360 | LEOM | 381 |
| LEOM | 201 | LEOM | 220 | LEOM | 248 | LEOM | 282 | LEOM | 307 | LEOM | 327 | LEOM | 361 | LEOM | 382 |
| LEOM | 202 | LEOM | 221 | LEOM | 252 | LEOM | 284 | LEOM | 309 | LEOM | 329 | LEOM | 362 | LEOM | 383 |
| LEOM | 204 | LEOM | 222 | LEOM | 253 | LEOM | 286 | LEOM | 310 | LEOM | 330 | LEOM | 363 | LEOM | 384 |
| LEOM | 205 | LEOM | 223 | LEOM | 254 | LEOM | 291 | LEOM | 311 | LEOM | 333 | LEOM | 364 | LEOM | 385 |
| LEOM | 206 | LEOM | 224 | LEOM | 259 | LEOM | 296 | LEOM | 313 | LEOM | 334 | LEOM | 365 |  |  |
| LEOM | 209 | LEOM | 227 | LEOM | 270 | LEOM | 298 | LEOM | 316 | LEOM | 336 | LEOM | 366 |  |  |
| LEOM | 210 | LEOM | 231 | LEOM | 272 | LEOM | 300 | LEOM | 320 | LEOM | 357 | LEOM | 367 |  |  |
| LEOM | 211 | LEOM | 238 | LEOM | 274 | LEOM | 301 | LEOM | 321 | LEOM | 358 | LEOM | 368 |  |  |
| LEOM | 213 | LEOM | 243 | LEOM | 276 | LEOM | 302 | LEOM | 322 | LEOM | 359 | LEOM | 369 |  |  |

LEOM: Laboratory of Evolution of Marine Organisms, $\mathrm{N}^{\circ}$ : colony number

