Image Classification Methods Based on Texture Analysis and Characterization

by

Steve Tsham Mpinda ATAKY

THESIS PRESENTED TO ÉCOLE DE TECHNOLOGIE SUPÉRIEURE
IN PARTIAL FULFILLMENT FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
Ph.D.

MONTREAL, OCTOBER 12, 2022

ÉCOLE DE TECHNOLOGIE SUPÉRIEURE
UNIVERSITÉ DU QUÉBEC

BOARD OF EXAMINERS

THIS THESIS HAS BEEN EVALUATED

BY THE FOLLOWING BOARD OF EXAMINERS

M. Alessandro Lameiras Koerich, thesis supervisor
Software and Information Technology Engineering Department, École de technologie supérieure

M. Ismail Ben Ayed, president of the board of examiners
Software and Information Technology Engineering Department, École de technologie supérieure

M. Jose Dolz, member of the jury
Software and Information Technology Engineering Department, École de technologie supérieure

Ms. Aura Conci, external independent examiner
Department of Computer Science, Fluminense Federal University

THIS THESIS WAS PRESENTED AND DEFENDED

IN THE PRESENCE OF A BOARD OF EXAMINERS AND THE PUBLIC

ON AUGUST 24, 2022

AT ÉCOLE DE TECHNOLOGIE SUPÉRIEURE

ACKNOWLEDGEMENTS

May those whose names have not been mentioned receive the expression of all my thoughts, respect, and appreciation.

Praise ye the LORD: For thy lovingkindness is better than life; my lips shall praise thee.

To my advisor/supervisor Dr. Alessandro Lameiras Koerich, for the trust and credibility placed upon me since the beginning of this journey. My gratitude for his instructions, adequate guidance, and inspiration. Likewise, his patience, the opportunity for growth, understanding under challenging times, and commitment to always be present when requested. I will perpetually have respect, friendship, and admiration for this great man.

"As a matter of self-preservation, a man needs good friends or ardent enemies, for the former instruct him, and the latter take him to task." Diogenes. As of the former, Besides Ataky's family, I am humbled by having Patrick A. Bungama, Daniel-Charles E. Onan, Genick Masongele, Mbamu's family, Helvécio W. Pereira, Lívio Linhares, and Diego Saqui in my life.

I have been morally, emotionally or scientifically supported by , but not limited to, Jered Ataky, Dr. Aristófanes C. Silva, Dr. Laíse Nayra, Reagan Mampuya, Ali M. Kumakamba, Didier M. Kazadi, Jonathan Nyembwa, Maglore Tshama Noella, Plamedi Lusembo, Rosin Ngueveu, Dr. Ulrich Matchi Aivodji, Dr. Arthur Sawadogo, Ndiaye Diop, Mouhammad Dieye, Jonathan de Matos, Pr. Norton Lages, Pedro Filho, Mélanie Maltais, Fernand B. Bi-Tonye, Dr. Thiago Paixão and Thiago Lemos Fonseca, Samir Souza, Raíssa Everton, Paulo José, Olavo and Débora Castro, and Tio Domingos (InMemoriam), to whom I extend my profound gratefulness.

I could not forget the contributions I received from LIVIA's members, to whom I am truly obliged.

This thesis was funded by the Regroupement Strategique REPARTI - Fonds de Recherche du Québec - Nature et Technologie (FRQNT) and by the Natural Sciences and Engineering Research Council of Canada (NSERC).

Méthodes de classification d'images basées sur l'analyse et la caractérisation de la texture

Steve Tsham Mpinda ATAKY

RÉSUMÉ

La texture peut être définie comme des changements dans l'intensité de l'image qui forment certains motifs répétitifs. Ces motifs peuvent être causés par les propriétés physiques à la surface d'un objet ou par des différences de réflexion, telles que la couleur à la surface. Bien que la reconnaissance de texture soit relativement simple pour la perception humaine, il n'en est pas de même dans les procédures automatiques, où cette tâche nécessite fréquemment des techniques de calcul complexes. L'analyse de texture joue un rôle important dans la vision par ordinateur, et son fondement est l'extraction des propriétés intrinsèques d'une image, avec lesquelles sa texture sera caractérisée. Une telle analyse est remarquable dans plusieurs applications telles que la télédétection, la médecine, l'agriculture, l'analyse d'images, la microscopie, etc. Comprendre comment les humains font la distinction entre différentes textures permet le développement de techniques capables d'accomplir cette tâche. Considérant que la perception des textures par l'homme ne change pas avec les rotations, les translations ou les changements d'échelle, toute caractérisation numérique de ces dernières devrait avoir les propriétés fondamentales suivantes : invariance aux changements de contraste et invariance aux transformations monotones. Néanmoins, dans l'ensemble, les approches actuelles de l'état de l'art présentent encore des problèmes de performances, d'une part pour le manque de caractère invariant aux transformations géométriques, telles que les transformations de similarité, et un composant pertinent qu'un descripteur de texture devrait également remplir, à savoir l'invariance aux changements d'intensité tels que les transformations d'intensité monotones ; d'autre part, en outre, de telles approches subissent les contrecoups de fluctuations aléatoires ou de bruit, car les propriétés intrinsèques de l'image en question ne sont pas conservées. Compte tenu des problèmes susmentionnés et parce que la texture forme un système de modèles non déterministe, la mesure théorique de l'information de la diversité écologique, une branche de la biologie, peut aider à sa caractérisation au maximum possible. Pour accomplir cette tâche, les concepts de diversité, de richesse, d'uniformité et de distinction taxonomique des espèces ont été adaptés et appliqués, et utilisés comme descripteurs de texture dans cette thèse. Les résultats obtenus sur des jeux de données naturelles et histopathologiques ont montré les avantages des méthodes proposées, qui sont compétitifs avec les descripteurs de l'état de l'art.

Mots-clés: Reconnaissance de Formes, Descripteur de Caractéristiques, Mesures de Diversité Écologique, Indices de Biodiversité et Taxonomiques, Analyse de Texture, Théorie de l'information, Classification d'images.

Image Classification Methods Based on Texture Analysis and Characterization

Steve Tsham Mpinda ATAKY

ABSTRACT

Texture can be defined as changes in image intensity that form specific repetitive patterns. Such patterns can be caused by the physical properties on the surface of an object or by differences in a reflection, such as color on the surface. Although texture recognition is relatively simple for human perception, it is not likewise in automatic procedures, where this task frequently necessitates complex computational techniques. Texture analysis plays an essential role in computer vision, and its foundation is the extraction of intrinsic properties from an image, with which its texture will be characterized afterward. Such an analysis is noteworthy in several remote sensing, medicine, agriculture, image analysis, microscopy applications, etc. Understanding how humans discriminate between different textures allows developing techniques to perform this task. Considering that the perception of textures by humans does not change with rotations, translations, or changes in scale, any numerical characterization of the latter should, likewise, have the following fundamental properties: invariance to changes in contrast and monotonic transformations. Nonetheless, by and large, current state-of-the-art approaches still show performance issues, namely the lack of invariance character to geometric transformations - such as similarity transformations - and a relevant component that a texture descriptor should also fulfill, to wit, invariance to intensity changes such as monotonic intensity transformations. As a further matter, such approaches suffer the aftereffects of random fluctuations or noise, for intrinsic properties of the image are not preserved. Given the issues mentioned above and because texture forms a non-deterministic system of patterns, the information-theoretical measure of ecological diversity, a branch of biology, can aid in its characterization to the maximum extent. Accordingly, concepts of species diversity, richness, evenness, and taxonomic distinctiveness were adapted and employed to build robust texture descriptors in this thesis, which are generic, independent of macro-level variations in terms of contrast, invariant to in-plane rotations of the image, explainable and interpretable based on biology concepts, and lend themselves to fast computation. The results achieved on natural and histopathologic datasets have shown the advantages of the proposed methods, which are competitive with state-of-the-art descriptors.

Keywords: Pattern Recognition, Feature Descriptor, Ecological Diversity Measurements, Biodiversity and Taxonomic Indices, Texture Analysis, Information Theory, Image Classification.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

Page

LIST OF ABREVIATIONS

BiT	Bio-Inspired Texture Descriptor
EBiT	Extended Bio-Inspired Texture Descriptor
HIs	Histopathological Images
GLCM	Gray-Level Co-occurrence Matrix
LBP	Local Binary Patterns
ROI	Region of Interest
CNN	Convolutional Neural Network
T-CNN	Texture Convolutional Neural Network
BIF	Basic Image Feature
GLP	Gaussian-Laplacian Pyramid
GP	Gaussian Pyramid
LP	Laplacian Pyramid
SIFT	Short Time Fourier Transform
CWT	Continuous Wavelet Transform
DW	Discrete Wavelet
DWT	Discrete Wavelet Transform
IDWT	Inverse Discrete Wavelet Transform
PRFB	Perfect Reconstruction Filter Bank
LL	Approximation

LH	Horizontal Detail
HL	Vertical Detail
HH	Diagonal Detail
D_{Mg}	Margalef's Diversity Index
D_{Mn}	Menhinick's Diversity Index
d_{BP}	Berger-Parker Dominance
d_F	Fisher's Alpha Diversity Metric
d_{KT}	Kempton-Taylor Index of Alpha Diversity
w_M	McIntosh's Evenness Measure
d_{SW}	Shannon-Wiener Diversity Index
d_{HB}	Brillouin Index
d_{Dw}	Strong's Dominance Index
d_C	Simpson's Index
d_{ENS}	Enspie Index
d_{McInt}	McIntosh Dominance Diversity Index
e_{CR}	Chao1 Richness Estimator
e_{GC}	Gini Coefficient
e_{HE}	Heip's Evenness
$e_{J'}$	Pielous Evenness
e_{SE}	Simpsons Evenness

S_{PD}	Sum of Phylogenetic Distances
d_{NN}	Average Distance from Nearest Neighbor
e_{IQ}	Intensive Quadratic Entropy
e_{EQ}	Extensive Quadratic Entropy
d_{TT}	Total Taxonomic Distinctness
PCA	Principal Component Analysis
RGB	Red Green Blue
SVM	Support Vector Machine
XGBCB	eXtreme Gradient Boosting
HistoB	Histogram-Based Gradient Boosting Ensembles
LightB	Light Gradient Boosting Machine
SuperL	Super Learner
k-NN	k-Nearest Neighbors
CatBoost	Gradient Boosting on Decision Trees
LDA	Linear Discriminant Analysis
RF	Random Forest
GB	Gradient Boosting
ExtraT	Extra trees
HOG	Histogram of Oriented Gradients
CAD	Computer-aided Detection

CADx	Computer-assisted Diagnostic
ML	Machine Learning
BiTW	Biodiversity Measures, Information Theory and Taxonomic Indexes Extracted From Wavelet Subband
CV	Cross-Validation
ACC	Accuracy
AUC	Area Under ROC Curve
Tio	Three-in-One (TiO)

LIST OF SYMBOLS AND UNITS OF MEASUREMENTS

Δ	Taxonomic Diversity
Δ^*	Taxonomic Distinctiveness
Δ^+	Average Taxonomic Distinctiveness

INTRODUCTION

Classifying an image's pattern is one of the most challenging processes in digital image processing. Because the information extracted from existing patterns can aid decision-making in various areas, the proper adjustment and the improved techniques for their classification are fundamental so that an image can become a valuable auxiliary tool for decision-making. The texture is helpful information that can be associated with interpreting patterns related to various areas of object recognition. The texture of an object is a descriptor of its shape, size, shade, and tonality, allowing the visual impression of the roughness or smoothness of a given surface.

Likewise, the texture is a way to interpret visual information using a natural characteristic of human eyes, for it carries information about the spatial distribution of an object's tonal variations or the tonal repetition of some groups of objects that are not individually identifiable - luminosity -, and structural arrangement of the surface concerning neighboring regions (Petrou & Sevilla, 2006). The term texture is frequently used to describe the sensation of touching an object. This experience gave rise to terms like rough, smooth, and soft for classification purposes. There is a texture present in the images known as visual texture, for which there is still no agreement on its definition (Gonzalez & Woods, 2009). It is referred to as the repetition of simple or complex patterns on a surface. These patterns can be identical duplicates that appear repeatedly or have minor or non-deterministic changes. There is also a link between the visual texture and the texture of the physical objects captured by these images. As a result, analyzing this type of texture is highly beneficial to object recognition. Accordingly, there is a need for computational techniques that allow the classification of such patterns in an image, which not only improve texture characterization - to the maximum extent - but also have a low computational cost.

The texture analysis seeks to determine the neighborhood relationship of the texture elements and their position to the others (connectivity), the number of features per spatial unit (density), and the regularity (homogeneity) therefrom (Traina, 2001). Texture approaches developed in

the literature to characterize an image can be divided into statistical and geometric methods, according to Tuceryan & Jain (1993). The former seeks to determine how and to what extent some image properties related to texture may be distributed and then derives numerical texture measures. The latter investigates the various types of periodicity in an image and characterizes a texture using the relative spectral energy at different periodicities.

Texture characterization is not trivial because this process presents considerable challenges, one of which is related to the conditions in which an image was captured. Accordingly, lighting geometry or intensity changes can significantly impact a textural image's appearance. Moreover, in both the microtexture and macrotexture, different types of texture may restrain the characterization process and the development of methods robust enough to represent textural information from/in an image, mainly when both are present. Indeed, each texture mentioned above contains intrinsic properties that may require a specific statistical method to be represented to the utmost extent. In computer vision, this process involves image processing techniques used to characterize texture properties from an image. These techniques allow the extraction of descriptors from an image or a region related to characteristics that refer to intrinsic properties such as roughness, regularity, smoothness, and so forth. Thus, choosing which texture analysis method to employ for extracting features becomes challenging for the success of the classification phase.

A variety of classical and novel approaches apropos of texture information extraction from images have been developed, namely gray-level co-occurrence matrix (GLCM) (Haralick, Shanmugam & Dinstein, 1973), Haralick descriptors (Haralick, 1979), local binary patterns (LBP) (Pietikäinen, Hadid, Zhao & Ahonen, 2011), wavelet transform (Arivazhagan & Ganesan, 2003), Markov random fields (Cross & Jain, 1983), Gabor texture discriminator (I & Sagi, 1989), local phase quantization (LPQ) (Ojansivu & Heikkilä, 2008), local tera pattern (LTP) (Saxena, Teckchandani, Pandey, Dutta, Travieso, Alonso-Hernández et al., 2015), binarized statistical

image features (BSIF) (Kannala & Rahtu, 2012), and fractal models (Kaplan, 1999), amongst many others that can be found in the literature.

Texture descriptors, like in natural images, are becoming increasingly popular in medical image analysis, particularly in histopathologic images (HIs), due to the variability of texture that such images exhibit. As a result, researchers have investigated a wide range of textural descriptors for classifying HIs, which are expected to be insensitive to translation, scale, rotation, and intensity changes. One of the most significant challenges in extracting features from such images is characterizing morphological features from structures observed in HIs and exploring higher-level representations capable of capturing relevant information for medical diagnosis purposes. The characteristics listed above are related to recognizing tissue changes (such as cell density or abnormal cell quantity) or cellular changes (such as malformed nuclei) caused by mitotic phases. Furthermore, morphological characteristics are linked to how pathologists investigate HIs, looking for specific reasons to classify them. In contrast, high-level features are generalizations of all structures in HIs, not just cell structures. As a result, most researchers exploit frequency-domain representations or texture descriptors (de Matos, Ataky, de Souza Britto, Soares de Oliveira & Lameiras Koerich, 2021a). Several approaches were developed to detect breast cancer through HIs—nevertheless, accurate classification of HIs remains a challenge.

Most often, such approaches concentrate on specific information in the image. Accordingly, most select a finite set of texture features based on contextual information in the form of a region of interest extracted from the local image. Nonetheless, relying solely on local texture information may harm their subsequent classification because the texture is defined not only by local information but also by its global appearance, representing the repetition and relationship between local patterns (Liu & Fieguth, 2012). Furthermore, the presence of noise, which distorts the observed data, is another factor that can penalize the classification of textural images. Typically, noise is inherent in the acquisition of real-world images.

Into the bargain, when it comes to classifying images with noisy textures, traditional classification methods and some current ones still have performance issues in most cases.

Despite extensive research into texture analysis, there is still a need to investigate increasingly precise and efficient techniques that meet the requirements of a diverse range of applications, that is, a context-free or generic technique, allowing for utmost texture characterization. Therefore, the main objective of this thesis is to develop approaches that, through the exploration of texture analysis, can produce discriminant characteristics to overcome the limitations found in existing methods of texture characterization for both natural images and images that contain other structures in addition to texture.

0.1 Problem Statement and Motivation

In recent years, the concept of information has become increasingly present and relevant at all levels of modern society. Because the volume of generated data has never been greater, being able to decode the symbols present in this vast ocean of information is an essential step in learning, understanding, and analyzing the rules that govern the most diverse complex phenomena that are part of nature. Some of the burdensome challenges in dealing with this scenario are the mining, identification, processing, and classification of patterns and symbols in the data produced by a diverse range of observable phenomena.

Along the same lines, object recognition is one of the most critical tasks in applications involving a computer vision system. The objective is to obtain a description that contains enough information to distinguish between different objects of interest. Typically, the recognition process relies on the objects' gray levels or colors, shape, and texture characteristics. Some examples of applications involving computer systems are listed below:

Remote Sensing: Aerial photographs must be evaluated and cataloged with fidelity and efficiency to obtain cartographic maps and geographic studies, such as soil analysis and mapping of forest and urban areas, as quickly as possible.

Quality control: Image processing has a wide range of applications in the field of quality control. Almost any industrial process that requires optical or visual monitoring can be automated. For example, the inspection of circuit boards for defects and the separation of parts on an assembly line are two typical applications.

Medicine: Images captured by magnetic resonance imaging (MRI), computed tomography (CT), and ultrasound can help with medical diagnoses. They help locate the damaged area but do not work at the cellular level; for that, a microscope level is required. In microscopic images, for instance, the analysis and interpretation help search for cells that have been damaged by cancer and other congenital abnormalities, as well as counting blood cells.

Microscopy: The analysis of images captured by optical or electron microscopes in fields ranging from biology to metallurgy to identify and classify cells or particles.

In the bargain, when it comes to texture characteristics of objects, it represents valuable information that can be associated with interpreting patterns related to the areas described above and numerous others. Notwithstanding, texture analysis, in turn, is critical for such applications necessitating an examination of an object's surface properties, an understanding of how humans distinguish between different textures, and modeling techniques capable of performing this task. Unfortunately, even though the texture is an essential aspect of surface appearance, it is difficult for most approaches to manipulate its appearance while considering the effects of illumination, viewpoint, and image surface shape.

Along the same lines, texture and color are the main aspects of any natural image. Both components have proven to be significant in various computer vision applications. One of the

assumptions apropos of color-texture integration is that both aspects are mutually dependent attributes of an image. Therefore, their extraction should be accomplished either from correlated pairs of color channels or individual color channels. In the matter of color, the challenge is to convey aspects of surface color appearance, such as the amount and the spatial relationships for each color. Even if most of the existing texture descriptors have proven to be discriminative for texture classification, some do not exploit the color information that may exist in images, which could bring - to some extent- relevant information. Human color perception appears so complex and nonlinear that it cannot be fully described by any single computational model, let alone a single linear model.

On the other hand, the demand for visual inspection automation in various tasks has grown significantly. Human inspection performance is frequently variable, making it unsuitable for processes requiring accuracy and speed of identification. In many cases, the same type of texture must be analyzed repeatedly to detect anomalies, necessitating the observer's undivided attention. Several studies show that humans demonstrate the inability to perform monotonous and repetitive tasks. Furthermore, there are inappropriate environments where human presence can be hazardous. These factors illustrate the need for, as well as the advantages of, using automated systems. Among them is an increase in productivity, an improvement in the quality of the products or services provided, the absence of personal risks, a high rate of inspection, and a reduction in costs (Tobias, Seara, Soares & Bermudez, 1995).

The texture attribute is an essential source of information for identifying objects or regions of interest in an image, contributing to increased classification accuracy. On the other hand, texture-based segmentation and classification techniques necessitate the development of a set of measures that accurately represent it. Thus, this thesis presents developed methods that allow texture characterization and classification tasks to meet this need. To that end, it was necessary to address:

1. The development of techniques for texture analysis and characterization in natural images and those containing other structures in addition to texture;
2. The adaptation and development of strategies that enable analysis with as much relevant information as possible;
3. The image representation scheme that allows the biological concepts from the branch of the study of species diversity, namely ecological diversity indices between species, richness, abundance, and evenness of species, to be adapted for the extraction of features that measure and characterize texture;
4. The exploitation of multi-scale and multi-resolution analysis to supply characteristics that describe texture from an image to a great extent under distinct differentiations for bettering the classification performance.

0.2 Objectives

Considering the points raised in the preceding section, this thesis seeks new tools and novel approaches for the characterization of textures, intending to extract relevant information and then recognize patterns thereof. The overarching objective of this thesis was to present the development of approaches for extracting features from texture images. Such approaches are generic, provide a high accuracy rate, are independent of macro-level variations in terms of contrast, are invariant to in-plane rotations of the image, lend themselves to fast computation, and can be explained and interpreted. Results similar to or better than those provided by traditional or state-of-the-art methods available in the literature are considered to have high accuracy levels.

The approaches proposed in this thesis to achieve the aforementioned general objective are based on developing new texture analysis descriptors founded on biological diversity measurements (Clarke & Warwick, 1998a; Da Silva & Batalha, 2006; Magurran, 2004b; Vandamme, Pot, Gillis, De Vos, Kersters & Swings, 1996).

More specifically, the research objectives of this thesis are as follows:

- To investigate, adapt, and apply biological diversity measurements for texture characterization of textural images and images containing not solely texture;
- To propose new texture descriptors based on biological diversity measurements;
- To establish a parallelism between restoration ecology and multi-scale, multiresolution, and integrative methods for leveraging temporal/spatial trends of the texture patterns;
- To compare the performance of the proposed descriptor against classical methods and approaches based on deep learning.

0.3 Research Hypothesis

Our point of view is based on the principle that data distribution of textures forms a non-deterministic complex system; that is, its behavior cannot (or not always can) be predicted. We stated that textural patterns behave similarly to ecological patterns, in which large populations of units can self-organize into aggregations that generate patterns from processes that are nonlinear and non-deterministic.

Therefrom, this thesis investigated the following hypotheses:

- The use of information-theoretical measures of ecological diversity indices in conjunction with measures of biodiversity can provide a robust tool for quantifying such a complex system of diverse patterns;

- The combination of species richness, abundance, evenness, and taxonomic indices can integrate properties of statistical and structural approaches to texture analysis and take advantage of ecological patterns' invariance characteristics to permutation, rotation, and scale;

- The exploitation of integrative methods, multi-scale, multiresolution analysis can supply as many intrinsic properties as possible that reflect the local and global descriptors and summarize values that describe texture from color and gray-scale images to a great extent under distinct differentiations.

0.4 Contributions

The major contributions of this thesis are related to the characterization of texture patterns based on ecological diversity indices. In general, the following three contributions are defined:

1. Adaptation of Biological concepts, specially:
   a. Ecological concepts from the branch of the study of species diversity: (i) species richness, abundance, and evenness; (ii) evolutionary relationships between species, topology, and shortest path; and
   b. Restoration ecology and referent.
2. Development of descriptors: generic, independent of macro-level variations in terms of contrast, invariant to in-plane rotations of the image, lend themselves to fast computation, explainable and interpretable based on biology concepts.
3. Modeling of texture image as an ecosystem of organisms for applying Biological concepts; and
4. Establishment of a parallelism between restoration ecology and multi-scale and multiresolution analysis.

0.5 Thesis Organization

This thesis is organized in six chapters as follows:

Chapter 1 presents the theoretical foundation used in the development of this thesis, which is necessary to understand the techniques employed in the proposed methods for the characterization

and classification of texture images. Likewise, the literature review regarding texture analysis is presented. In a nutshell, we explore the concept of texture and its analysis, features extraction, multiresolution analysis of images, phylogenetic diversity indexes, information theory, and feature selection.

Chapter 2 presents and discusses novel approaches to quantifying patterns for texture characterization using ecological diversity measures, namely species diversity, richness, evenness, and taxonomic indices. The proposed methods consider an image as a species ecosystem, where it becomes possible to extract and compute ecological diversity measures to describe the texture. Section 2.1 introduces a novel bio-inspired texture (BiT) descriptor based on biodiversity measurements (species richness and evenness) and taxonomic distinctiveness. Section 2.2 put forth an extension of the BiT descriptor that integrates a few sets of diversity and evenness measures widely used in ecology to resemble the completeness of alpha diversity to build a robust representation for texture characterization and classification.

Related publications:

1. A novel bio-inspired texture descriptor based on biodiversity and taxonomic measures. Pattern Recognition (Journal). Volume 123, March 2022, 108382.
   Status: Published.
2. E-BiT: Extended bio-inspired texture descriptor for texture analysis and characterization. Image and Vision Computing (Journal).
   Status: Submitted.

Chapter 3 employs the multi-scale and multiresolution analysis for texture analysis of natural and histopathologic images. We establish a parallelism between restoration ecology and multi-scale and multiresolution. In the former, a site (soil, hydrology, or climate) may change over time - yet such changes in primary resources have to be taken into account to strongly define/characterize a community, for such differentiations reveal the temporal/spatial trends of

the biodiversity conservation capacity. In the latter, analyses such as wavelet decomposition and pyramids are employed to supply a characteristic value that can reflect the background condition and summarize values that describe texture from an image to a great extent under distinct differentiations.

Related publications:

1. Multiresolution Texture Analysis of Histopathologic Images Using Ecological Diversity Measures.

   Expert Systems With Applications (Journal).

   Status: Minor Review.

2. Multiscale Analysis for Improving Texture Classification.

   Applied Science (Journal).

   Status: Submitted.

Conclusion presents final considerations, the limitations discovered in this thesis, and discusses future works.

CHAPTER 1

THEORETICAL FOUNDATIONS AND LITERATURE REVIEW

This chapter presents the theoretical foundation used in developing this thesis, which is necessary to understand the techniques used in the proposed methods for the classification of texture images. Likewise, the literature review regarding texture analysis is also presented. The following sections explore the concept of texture and its analysis, features extraction, multi-resolution analysis of images, phylogenetic diversity indexes, information theory, and feature selection.

1.1 Texture

The definition of texture has different flavors in the literature on computer vision. A common one considers texture as changes in the image intensity that form specific repetitive patterns (Tuceryan & Jain, 1993). These patterns may result from the physical properties of the object's surface (roughness) that provide different types of light reflection. A smooth surface reflects the light at a defined angle (specular reflection), while a rough surface reflects it in all directions (diffuse reflection). Although texture recognition is easy for human perception, it is not the case with automatic procedures, where this task sometimes requires complex computational techniques. In computer vision, texture analysis is of a notable role, whose basis is extracting relevant information from an image to characterize its texture. This process involves a set of algorithms and techniques. Since humans' perception of texture is not affected by rotation, translation, and scale changes, any numerical image texture characterization should be invariant to those aspects and any monotonic transformation in pixel intensity.

Microtexture: Primitives or micro-patterns that form the panorama of a given surface define microtexture. These irregular micropatterns make it challenging to develop a structural model that adequately describes them because information about the shape is difficult to obtain. As a result, it is common to describe it using a stochastic approach. That is, employing a descriptor that can statistically summarise the relationship between the patterns in the material that a surface is made of.

Macrotexture: The contrast between shadows and well-lit areas and the shape of patterns in macrotexture textures produce primitives with larger dimensions when compared to microtexture dimensions. Furthermore, such patterns and their shape are quite repetitive in the material that makes up the surface under consideration.

1.2 Texture Analysis

The analysis of texture in an image is critical. Its purpose is to investigate an object's surface properties, understand how humans discriminate between different textures, and model algorithms capable of performing this task. The following question may arise before analyzing an image's texture: Is there any texture in the image?

Karu, Jain & Bolle (1996) addressed this question in their work by determining whether or not a given image contains texture for analysis. However, most texture definitions consider it to be a measurement of the roughness of an image. Images with no texture can be classified as noisy or containing distinct objects too large to characterize the texture. Thus, for an image to have texture, it must fall somewhere between these two extremes.

The texture is frequently confounded and used as a synonym for the pattern. The term pattern is used herein to describe a certain spatial regularity in an area of interest. On the other hand, texture is a measurement of spatial variability within an area of interest, whether random or not. Because a technique is not spatial in the measurement scale, many texture measurements may fail to distinguish patterns from other textures. Moreover, although numerous pattern recognition techniques can be applied in various domains, an optimal and specific individual approach for a given application has yet to be discovered.

Broadly speaking, one of the important aims of texture analysis is to compare textures and determine whether they are the same or different. Thereby, Materka, Strzelecki et al. (1998) presented four elements as part of this set: feature extraction, texture classification, texture segmentation, and texture reconstruction of shapes. The first two elements of this set are the focus of this thesis.

1.2.1 Feature Extraction

Extracting features provides information for further steps, such as classification and analysis. Various information or characteristics can be obtained in this process, such as histogram, size, area, perimeter, texture, contours, etc. Several image classification systems use color, texture, and shape to represent an image. Although color is a reliable attribute, situations in which color information does not discriminate images well require using texture or shape attributes. Besides, classifiers based on a single image attribute may not achieve adequate accuracy, so it is necessary to use multiple image attributes.

Texture analysis is the process of determining the texture content of a region or regions of an image. A feature vector is a result of characterizing the texture of an image's region of interest (ROI). It contains all numerical measurements that can perform various tasks like classification, segmentation, and image retrieval by content. Furthermore, the length of the feature vector will determine its dimension (n-dimensional vector). Such a new representation, nevertheless, should meet the following three considerations to wit reduce the dimensionality of the data, emphasize aspects of the image to facilitate human perception, and be invariant to the transformations of the image (Loew, 2000).

Several methods for image classification using texture features have been proposed in the literature. Regardless, no standard method or formal approach can be applied to a wide range of images. In addition, different types of descriptors are classified according to each approach adopted for their elaboration. Haralick *et al.* (1973) defined texture as the image's uniformity, density, roughness, regularity, and intensity, among other things. They define texture as a two-dimensional concept, with one dimension containing the fundamental properties of tonality and the other corresponding to their spatial relationships. They indicated that the concepts of tonality and texture are not independent, as tonality is dominant in some images while texture is dominant in others.

Materka *et al.* (1998) and Bharati, Liu & MacGregor (2004) distinguish four main approaches used in image classification for texture description: (i) statistical methods, (ii) structural methods,

(iii) model-based methods, and (iv) transform-based methods. Whereas, Gonzalez & Woods (2009) and Dougherty (2009) categorize them as follows: (i) statistical approach, (ii) structural approach, and (iii) spectral approach. On the other hand, Rao (2012) and Tuceryan & Jain (1993) portray classification for texture description in only two groups: (i) statistical or stochastic approach and (ii) structural approach. In addition to the approaches to texture descriptors mentioned above, some approaches have come to describe the texture over time and through extensive research in this area. These methods rely on auto-regressive models, Gaussian-Markovian random fields, wavelets, fractals, and other techniques. These models provide more powerful tools for analyzing invariant textures.

The three main approaches to describing textures in image classification are statistical, structural, and spectral. A set of local measurements extracted from the pattern defines the texture in statistical approaches. Common statistical measures include entropy, correlation, contrast, and variance. On the other hand, structural approaches are based on the idea that textures are made up of primitives arranged in an approximately regular and repetitive manner according to well-defined rules. For example, consider the texture description based on regularly spaced parallel lines. The spectral approaches, in turn, are based on Fourier spectrum properties and are primarily used to detect global periodicity in an image by identifying high-energy peaks in the spectrum.

1.2.1.1 Statistical Approaches

A texture can be described statistically in a suitable way for statistical pattern recognition. As a result, each texture is represented by a feature vector representing a point in a multi-dimensional feature space. The goal is to find a probabilistic decision rule that associates the texture with a specific class (Sonka, Hlavac & Boyle, 2014). Statistical approaches do not attempt to understand the texture's hierarchical structure explicitly. Instead, they represent texture indirectly by defining the distributions and relationships between an image's grey levels using non-deterministic properties. In other words, they attempt to determine how a texture-related image property may be distributed and then derive numerical texture measures from the computed distributions.

Several attempts to characterize textures have been made. However, most of them have relied on extracting the first- and second-order properties of grayscale and color levels. The first order is concerned with properties derived directly from individual pixels, i.e., without cross-comparisons. In contrast, second-order properties involve comparing two pixels at the same time. In other words, it entails determining how one pixel at a reference location statistically relates to another pixel at a location distant from the reference location.

The human visual system's recognition of textures was investigated using statistical properties (first and second-order statistics) (Julesz, 1975). Later research by Bergen & Julesz (1983) and Julesz (1981) gave rise to this statistical model. The co-occurrence matrix (GLCM) is the most widely used texture descriptor based on second-order statistics (Haralick, 1979). Once computed, this matrix contains all information about the spatial dependence of pixels in an image. The relationships between shades of grey, according to Haralick *et al.* (1973), are characterized by a function of distances and angles and represent all texture information when contained in these structures.

1.2.1.2 Structural Approaches

The structural approach, explored by Haralick (1979) and Levine (1985) refers to texture analysis by decomposing the image into primitives known as texels (texture elements). In fact, when the texture primitive is large enough to be individually segmented and described, structural approaches are considered appropriate. The texture primitive is the fundamental geometric structure that gives rise to the texture. For example, it could be the pixel itself in a very fine texture of digital images. The texture in this method is defined by sub-patterns called primitives that repeatedly appear within a pattern according to well-defined rules (Nadler & Smith, 1993).

If a reasonable set of texture primitives can be extracted from an image, a texture can be described using statistics on the properties of these primitives. The mean or standard deviation of grey levels, area, perimeter, orientation, and so on are some examples. A more straightforward approach is to extract blocks with the highest degree of homogeneity and describe the image

in terms of block size statistics. This description can be hierarchical, which means that the primitives can be made up of sub-primitives. In this case, the texture is described using stochastic grammars, in which a probability density function determines the rules to be applied.

When a texture is described structurally, the basic idea is that a simple texture primitive can be used to form complex texture patterns by following rules that limit the number of possible arrangements (Gonzalez & Woods, 2009). Furthermore, the structural approach has the main advantage of providing a good symbolic description of an image, making it more suitable for analysis tasks than texture synthesis (Materka *et al.*, 1998). However, because structural techniques are limited to well-defined macro-textures, they do not perform well in analyzing natural textures.

1.2.1.3 Model-based Approaches

A texture is modeled as a probabilistic mathematical model or as a linear combination of a set of basis functions in the context of model-based methods. As a result, they are referred to as stochastic and generative models, respectively. In this context, texture analysis begins with estimating the coefficients of these models, which are then used to characterize texture images. The central issue is always how to estimate the model coefficients and select the best model for a given texture (Zhang & Tan, 2002).

Gaussian Markovian and fractal random fields are classic examples of these paradigms. The Gaussian Markov random field model can encapsulate spatial dependencies between a pixel and its neighbors due to its local conditional probability distribution. That is, the probability distribution establishes that the value of each pixel is directly dependent on its neighbors (Zhao, Zhang, Li & Huang, 2007). Fractals are based on systematic construction rules at different scales. They have demonstrated good performance in models and represent natural surfaces because patterns in nature have similar qualities to each other (with some statistical variations) at different levels of scale (Dougherty, 2009; Jain & Farrokhnia, 1990; Mandelbrot, 1967; Tuceryan & Jain, 1993).

1.2.1.4 Transform-based Approaches

The transform-based descriptors reflect an image in space whose coordinate system has an interpretation that is closely related to the features (for example, frequencies) of the texture being analyzed (Materka *et al.*, 1998).

The texture analysis in this approach is done based on the frequencies that make up the image. The Fourier transform, Gabor filters, and wavelet transform stand out among these methods. Fourier descriptors use the signal's power or energy spectrum to obtain texture information. According to Gonzalez & Woods (2009), this spectrum plays a critical role in texture characterization. He can describe the directionality of an image's periodic patterns. It is worth noting that methods based on Fourier transforms do not perform well in practice due to a lack of spatial location information, that is, there is a lack of information indicating exactly when a given frequency appears.

Because the texture is highly scale-dependent, its sensitivity can be reduced by describing it in multiple resolutions. To achieve the best texture discrimination, choose an appropriate scale. Gabor and wavelet transform are suitable for multi-scalar texture characterization (Sonka *et al.*, 2014). Both approaches represent an image in a space whose coordinate system has an interpretation closely related to texture characteristics such as frequency or size.

Gabor filters improve spatial localization, but their practical utility is limited because there are no simple resolution filters that can locate a spatial structure in natural textures. Compared to the Gabor transform, the wavelet transform has some advantages: (i) by varying the spatial resolution, textures can be represented on a more appropriate scale; (ii) the wavelet function has a wide range of options, making this approach more suitable for texture analysis in a specific application. As a result, the wavelet transform is appropriate for texture segmentation (Materka *et al.*, 1998).

There are also hybrid methods that combine more than one approach, resulting in a texture descriptor partially related to one or another technique. A multi-level texture description is an

example that is based on the definition of primitives and the spatial description of relationships between primitives. The method, which consists of several steps, considers grayscale and structural intensities. Texture primitives must first be extracted, described, and classified. A classifier learns to classify texture primitives due to this processing stage. In a second learning stage, known textures are presented to the recognition system. Texture primitives are extracted from the image, and the first-level classifier recognizes their classes. For each texture in the training set, spatial relationships between primitive classes are evaluated using the recognized texture primitives. A feature vector is used to fit a second-level classifier to describe spatial relationships between texture primitives. The second-level learning process begins when the second-level classifier is chosen, and unknown textures can be presented to the recognition system. The first-level classifier classifies the primitives, the properties of spatial primitives are calculated, and the texture is assigned to one of the texture classes by the second-level classifier (Sonka *et al.*, 2014).

1.2.2 Texture Characterization and Classification

Any numerical characterization of image textures must have certain properties to be effective in practical applications. It must be invariant to variations in visual contrast caused by shifting or uneven illumination of a scene to the greatest extent possible, providing that the texture remains nearly the same according to human perception. At the very least, the numerical characterization must be invariant to the grayscale's monotonic transformation. This is significant because the lighting circumstances under which the training data for a machine-learning algorithm was gathered are unlikely to be the same as those under which the test data was acquired. Furthermore, it must be invariant to in-plane rotations of the image to the greatest extent possible. Because the texture orientation used to train a machine-learning algorithm is unlikely to be identical to the orientations of the same texture in the test images. Finally, it must be capable of quick computation.

A texture classification task's main goal is to determine which class (two or more classes specified in advance) a texture sample belongs to based on a similarity criterion. Two procedures

are involved in this task: feature extraction and pattern categorization. The usage of supervised machine learning methods is directly associated with pattern classification. That is, it is first and foremost important to prepare a set of samples for training. The values of these samples are then mapped to the labels of their respective classes, allowing the inference of new unknown samples, known as test samples. These test samples can be obtained in two ways: through cross-validation or by taking samples from outside the training process. Cross-validation, according to Barrow & Crone (2016), is a statistical approach that determines how well the results of an estimate can be generalized to an independent database.

1.3 Literature Review on Texture Analysis and Characterization

This section provides an overview of the existing descriptors and approaches employed in texture analysis, focusing on feature extraction from textural images. Some existing texture descriptors considered state-of-the-art in texture analysis of natural and histopathological images are presented.

Several comprehensive literature reviews apropos of texture representation as texture analysis are presented in Simon & Uma (2018b) and Liu, Chen, Fieguth, Zhao, Chellappa & Pietikäinen (2019). Whence a considerable number of methods have been developed over time, exploring each, therefore, a different approach to extracting the texture information of an image. In fact, there is a variety of classical and novel approaches developed for such a purpose, and some of them are Gray-Level Co-occurrence Matrix (GLCM) (Haralick *et al.*, 1973), Haralick descriptors (Haralick, 1979), Markov Random Fields (Cross & Jain, 1983), Gabor Texture Discriminator (I & Sagi, 1989), Wavelet Transform (Arivazhagan & Ganesan, 2003), Fractal Models (Kaplan, 1999), Local Phase Quantization (LPQ) (Ojansivu & Heikkilä, 2008), Local Binary Patterns (LBP) (Pietikäinen *et al.*, 2011), Binarized Statistical Image Features (BSIF) (Kannala & Rahtu, 2012), and Local Tera Pattern (LTP) (Saxena *et al.*, 2015). Some recent works can be found in Simon & Uma (2018a), Wang, Li, Li, Gupta & Choi (2020), Zhang, Wang, Huang & Zhang (2018), and Nsimba & Levada (2019).

Furthermore, researchers have recently been focused on convolutional neural networks (CNNs) due to their effectiveness in object detection and recognition tasks. However, the shape information extracted by CNNs is of minor importance in texture analysis (Andrearczyk & Whelan, 2016). Andrearczyk & Whelan (2016) developed a simple texture CNN (T-CNN) architecture for analyzing texture images that pools an energy measure at the last convolution layer and discards the overall shape information analyzed by classic CNNs. Despite the promising results, the trade-off between accuracy and complexity is not so favorable. Other T-CNN architectures have also achieved moderate performance on texture classification (de Matos, de Souza Britto Jr., de Oliveira & Koerich, 2019; Fujieda, Takayama & Hachisuka, 2017; Vriesman, Britto Junior, Zimmer & Koerich, 2019). Another disadvantage of CNNs is the lack of explainability and interpretability. On that matter, Gilpin, Bau, Yuan, Bajwa, Specter & Kagal (2018) argued that interpretability and explainability are two distinct concepts as explainable models are interpretable a priori, but the reverse is not always true. Thus, interpretability alone is insufficient. Therefore, there is a need for explainable models capable of summarizing the reasons for deep learning behavior, gaining the trust about the causes of their decisions.

Besides natural images, texture descriptors are becoming quite popular in biological imaging analysis, particularly in Histopathologic Images (HI) analysis due to the different types of texture found in HIs. By analyzing an HI, it is clear that different areas of interest in the image, such as high/low concentration of nuclei and stroma present quite different patterns of textures. For this reason, several researchers have been investigating a large spectrum of textural descriptors for HI classification. The descriptors based on the Grey-Level Co-occurrence Matrix (GLCM) had been used by several authors to represent texture in HI. Kuse, Sharma & Gupta (2010) used GLCM as features with a pre-segmentation process based on unsupervised mean-shift clustering. Such a method reduces color variety to segment the image using thresholds. Afterward, nuclei are identified and have the overlapping removed by contour and area restrictions.

Additionally, Caicedo, González & Romero (2011) combined seven feature extraction methods, including GLCM, and created a kernel-based representation of the data on each feature type. Kernels are used inside an SVM to find similarities between data and to implement a content

retrieval mechanism. Fernández-Carrobles, Bueno, Déniz, Salido, García-Rojo & Gonzández-López (2015) presented a feature extraction method based on the frequency and spacial textons. The use of textons implies that images are represented by a reduced vocabulary of textures. Features used for the classification are histograms of textons and GLCM features extracted from texton maps. They also evaluated the impact of different colormaps on these procedures. Despite the fact that GLCM requires a gray-level image, the conversion of the hematoxylin and eosin (H&E) color image to gray-level is affected by the variability of the staining color, so in the end, GLCM is also affected. Leo, Lee, Shih, Elliott, Feldman & Madabhushi (2016) used a Random Forest to analyze whether GLCM features are susceptible or not to image variations. The work also highlighted the importance of color normalization.

The work presented by Reis, Gazinska, Hipwell, Mertzanidou, Naidoo, Williams, Pinder & Hawkes (2017) focused on stroma maturity to evaluate breast cancer. The features for the stroma are Basic Image Features (BIF), obtained by convolving images with a bank of derivatives-of-Gaussian filters, and LBP with multiple scales for the neighborhood. Das, Mitra, Chakraborty, Chatterjee, Maiti & Bose (2017) proposed the so-called geometric- and texture-aware features, which are based on Hu moments and fractal dimensional, respectively. Such a set of features was applied to detect geometrical and textural changes in nuclei to discriminate between mitotic and non-mitotic cells. Cruz-Roa, Caicedo & González (2011) proposed a patching method on HI slides to create small regions and extract SIFT, luminance level, and discrete cosine transform features to create a bag-of-words.

Texture analysis aims at grouping the texture patterns belonging to the same class. In addition to what is presented above, various texture descriptor approaches for feature extraction are found in the literature, and they are proven to be promising. Table 1.1 presents a summary of some of the texture descriptors and datasets used in classification tasks.

In fact, the list of works presented in the Table 1.1 is not exhaustive, more works can be found in Ghalati, Nunes, Ferreira, Serranho & Bernardes (2021), Singh, Sunkaria & Kaur (2022), Simon & Uma (2018b), and Tuceryan & Jain (1993).

Table 1.1 Summary of some texture descriptors and datasets used in classification tasks.

Author	Techniques	Datasets
Nsimba & Levada (2020b)	*Infomation theory and GMRF*	KTH-TIPS2b, Salzburg
Wang, Zhao, Cai, Li & Yan (2016)	*LBP and Zernike moments features*	Outex, CURet
Bashier, Hoe, Hui, Azli & Han (2016)	*Grand structure and histogram-based*	UIUC, XU
Hao, Wang, Li & Zhang (2016)	*Gaussian components and image features*	KTH-TIPS2, FMD, UIUC
Junior & Backes (2016)	*Extreme Learning Machine*	Brodatz, Outex, Vistex
Qiu, Thompson & Calderbank (2015)	*Based on binary Heisenberg-Weyl group*	Fabric texture
Ayed, Larousi & Masmoudi (2014)	*Local and global information*	Outex
Mehta & Egiazarian (2014)	*Granularity at multiscale.*	KTH-TIPS, UMD, Curet
Quan, Xu & Sun (2014)	*Fractal based*	KTH-TIPS, UMD, UIUC
Zhang, Zhao & Liang (2013)	*Gaussian derivatives filters*	CUReT, KTH-TIPS2
Backes, Casanova & Bruno (2013)	*Texture as pixel Network*	Outex, VisTex, Brodatz
Wang, Bichot, Zhu & Li (2013)	*LBP and neighboring gray-scale props*	Outex
Zhang, Zhou & Li (2012)	*Gabor filter and LBP*	CURet
Guo, Zhang & Zhang (2010)	*Completed LBP*	CUReT, Outex
Xu, Yang, Ling & Ji (2010)	*Wavelet and fractal*	UIUC, UMD

The descriptors previously mentioned have proven to be discriminative in terms of texture pattern classification. Considering that most are employed but to gray-scale images, their application on natural images as well as microscopic images may be constrained as color information is not exploited. In order to prevail over the presented limitation, an extension of LBP as well as other approaches have been proposed to incorporate the local patterns with color features. Likewise, in the work of Qi, Qiao, Li & Guo (2013), it is proposed an approach that encodes cross-channel texture correlation regarding color texture classification. The research conducted in Nsimba & Levada (2020a) presented a novel approach to compute information theory measures for color texture classification task, an approach based on the capacity of information theory measures to capture meaningful textural information of an input color image. In our methods, we will deal with these limitations in order to improve the metrics that evaluate performance, such as accuracy, recall, and Kappa, by considering and exploiting color information.

1.4 Multi-resolution

When looking at an image, we usually see regions of similar textures, colors, or levels of gray that combine to form objects. If the objects are small or have low contrast, it may be necessary to examine them in high resolution; if they are large or have high contrast, a coarser view is

sufficient. If both types of objects appear in an image, it can be helpful to analyze them in multiple resolutions (Gonzalez & Woods, 2009). Changing the resolution can also lead to the creation, deletion, or merging of image features. This serves as a motivation for an important paradigm in computer vision and image processing: multiresolution processing. In addition, there is evidence that the human visual system processes visual information in a multiresolution manner (Blakemore & Campbell, 1969), sensors can provide data in various resolutions, and multiresolution algorithms for image processing offer advantages from a computational point of view and are generally robust.

1.4.1 Gaussian-Laplacian Pyramid (GLP)

When analyzing an image, it can sometimes be useful to break it down into separate parts so that there is no loss of information. The pyramid theory provides ways to decompose images at multiple levels of resolution (Goutsias & Heijmans, 1998).

Consider a collection of representations of an image in different spatial resolutions, stacked on top of each other, with the highest resolution image at the bottom of the stack and subsequent images appearing over it in descending order of resolution. This generates a pyramid-like structure, as can be seen in Figure 1.1(a). The traditional procedure for obtaining a lower resolution image is to perform low-pass filtering followed by sampling (Jolion & Rosenfeld, 2012).

In signal processing and computer vision, pyramid representation is the main type of multi-scale representation for computing image features in different scales. The pyramid is obtained by repeated smoothing and subsampling of an image or a signal. This concept is frequently used because it expresses computational efficiency approximation compared to other representations such as scale-space representation and multi-resolution analysis (Burt, 1981; Crowley & Riff, 2003; Lowe, 2004). For generating the pyramid representation, different smoothing kernels have been brought forward and the binomial one strikingly shows up as useful and theoretically well-founded (Crowley, 1981; Lindeberg, 2013).

Accordingly, for a bi-dimensional image, the normalized binomial filter may be applied (1/4, 1/2, 1/4) in most cases twice or even more along all spatial dimensions, afterward, the subsampling of the image by a factor of two, which lead to efficient and compact multi-level representation. There are two main types of pyramids, namely, low-pass and band-pass (Adelson, Anderson, Bergen, Burt & Ogden, 1984; Burt & Adelson, 1983). To develop filter-based representations by decomposing images into information on multiple scales as well as to extract features/structures of interest from an image, Gaussian pyramid (GP), Laplacian Pyramid (LP), and Wavelet pyramid are examples of the most frequently used pyramids.

The GP illustrated in Figure 1.1 consists of low-pass filtered, reduced density, where subsequent images of the preceding level of the pyramid are weighted down using Gaussian average or Gaussian blur and scaled-down. The base level is defined as the original image. Formally speaking, assuming that $I(x, y)$ is a two-dimensional image, the GP is recursively defined in (1.1).

$$G_0(x, y) = \begin{cases} I(x, y), & \text{for level, } l = 0 \\ \sum_{m=-2}^{2} \sum_{n=-2}^{2} w(m, n) G_{l-1}(2x + m, 2y + n), & \text{otherwise.} \end{cases} \tag{1.1}$$

where $w(m, n)$ is a weighting function (identical at all levels) termed the generating kernel which adheres to the following properties: separable, symmetric and each node at level n contributes the same total weight to nodes at level $l + 1$. The pyramid name arose from the fact that the weighting function nearly approximates a Gaussian function. This pyramid holds local averages on different scales, which has been leveraged for target localization and texture analysis (Anderson, Burt & Van Der Wal, 1985; Burt, 1983; Larkin & Burt, 1983). Moreover, assuming the GP $[I_0, I_1, \ldots, I_k]$, the LP is obtained by computing $b_k = I_k - EI_{k+1}$, where EI_{k+1} represents an up-sampled, smoothed version of I_{k+1} of the same dimension. In the literature, LP is used for image compression, enhancements, analysis and graphics (Adelson *et al.*, 1984).

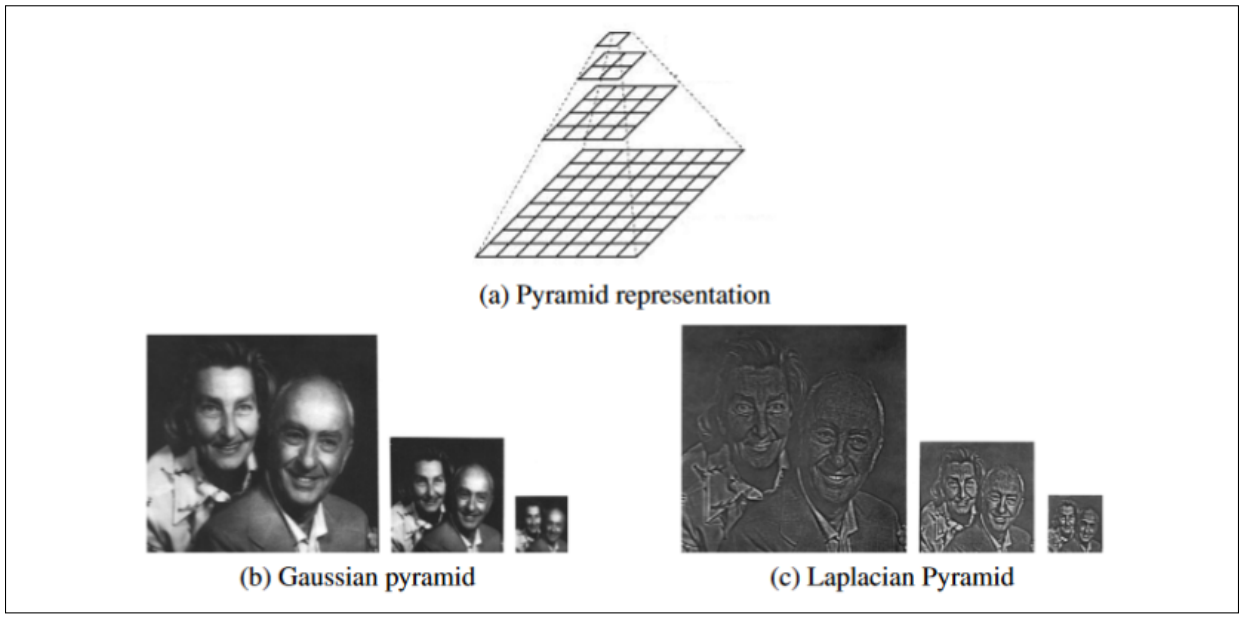

Figure 1.1 An example of Gaussian and Laplacian Pyramids from the same input image. (b) First three levels of Gaussian pyramid; (c) First three levels of Laplacian pyramid Adapted from Ataky *et al.* (2020)

1.4.2 Wavelet Analysis

The multi-resolution analysis is a signal processing strategy where a set of filters is used, specialized in extracting the signal information, such as the frequencies present in it and their location depending on the duration of the signal, in different resolutions (Castleman, 1996).

The brief description of the multi-resolution analysis allows presenting the two functions responsible for the generation of the entire wavelet system: the scale function and the primary wavelet (or mother wavelet). The term mother comes from the fact that functions with different sizes are used in the transformation process and all of them originate from the main wavelet, the mother wavelet.

The scale functions $\Phi_{j,k}$ and wavelets $\Psi_{i,j}$, are said to be orthogonal because they respect the following condition:

$$\int_{-\infty}^{+\infty} \Phi_{j,k}(x)\Psi_{j,k}(x)dx = 0 \tag{1.2}$$

where, $j \in Z$ corresponds to the parameter that represents the scale to which the function is represented, and $k \in Z$ corresponds to the translation of $k/2^j$ concerning the scale function and the primary wavelet, given by $j = 0$ and $k = 0$. Both the scale function and the wavelet are defined in the set of reals ($\mathbb{R}$), through scaling and translations of the presented functions.

The translation parameter corresponds to the time information in the transform domain and the scaling parameter corresponds to the signal compression and expansion process (Mallat, 1999).

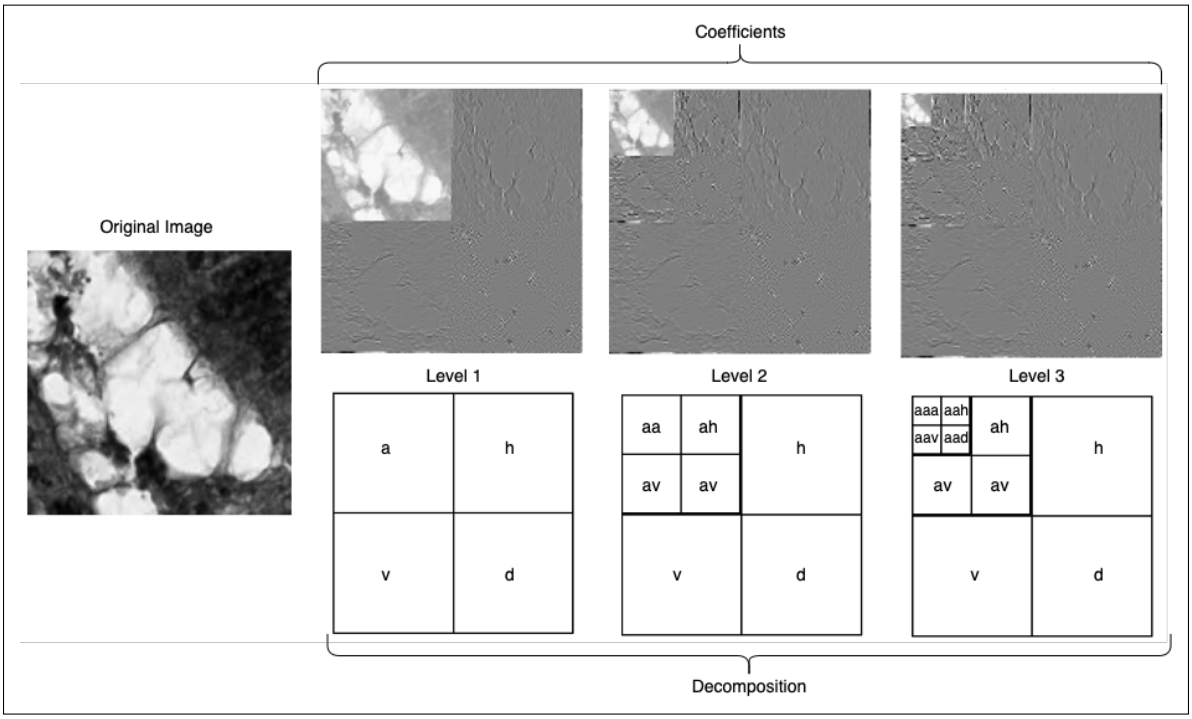

Figure 1.2 Multi-resolution representation of an image

Figure 1.2 shows an initial image and the degree of refinement applied to it through the wavelet transform. On a smaller scale, the useful information obtained from the original image is found in the six squares adjacent to the image. The information contained in these squares is called detail or resolution, which is the information needed to move from one degree of refinement,

or degree of "sharpness", to another. Adding the information related to the squares makes it possible to recompose the image. This way of decomposing and recomposing images can be implemented quickly and effectively, employing wavelet transforms.

1.4.2.1 Wavelets

Traditional methods of signal analysis, based on the Fourier transform, can determine all frequencies present in the signal, but their relationship with the temporal domain does not exist. To overcome this problem, the Gabor transform (or STFT - Short Time Fourier Transform) emerged; the main idea of this transform is to introduce a new parameter of local frequency as if the “local transform” observed the signal through a short window within which the signal remains approximately stationary (Oliveira, 2007).

The Wavelet transform was developed as an alternative to the Gabor transform to solve the problem of resolution. Wavelets are mathematical functions that separate signals into different components and extract each component with a resolution appropriate to its scale. They have an advantage over the Fourier transform because they analyze the signal at different scales and move around analyzing each point of the signal. The Continuous Wavelet Transform (CWT) can be expressed in the form:

$$CWT(\tau, a) = \int_{-\infty}^{+\infty} f(t) \frac{1}{\sqrt{a}} \Psi^* \left(\frac{t-1}{s} \right) dt \tag{1.3}$$

where τ and a are the translation and scale parameters, respectively.

1.4.2.2 Discrete Wavelet

The CWT is calculated by making continuous translations and scaling of a function over a signal. In practice, this transformation is not feasible, as it requires infinite translations and scaling, demanding a lot of time, computational effort, and redundancy. Discrete wavelets (DWs) were introduced to overcome this obstacle. DWs are not translated or scaled continuously, but at discrete intervals, which is achieved by modifying the continuous wavelet:

$$\Psi_{s,\tau}(t) = \frac{1}{\sqrt{|s|}}\Psi\left(\frac{t-\tau}{a}\right) \tag{1.4}$$

$$\Psi_{j,k}(t) = \frac{1}{\sqrt{\left|s_0^j\right|}}\Psi\left(\frac{t-k\tau_0 s_0^j}{s_0^j}\right) \tag{1.5}$$

where j and k are integers; $s_0 > 1$ is a fixed expansion parameter; τ_0 is the translation factor, which depends on the expansion factor.

Generally, $s_0 = 2$ is chosen to have a frequency sampling called dyadic sampling and $\tau_0 = 1$ is chosen for temporal sampling, also dyadic. This results in (Oliveira, 2007):

$$\Psi_{j,k}(t) = \sqrt{2^j}\Psi(2^j - k) \tag{1.6}$$

When discrete wavelets are used to analyze a signal, the result is a series of wavelet coefficients, also called a wavelet decomposition series (Oliveira, 2007). Since a wavelet can be seen as a low-pass filter, the series of scaled wavelets can be seen as a bank of band-pass filters with a Q factor (filter bank fidelity factor). To computationally implement wavelet transform, we need a notion of Discrete Wavelet Transform (DWT). DWT is used to decompose a signal into two other signals through low-pass filters l (scaling signals) and high-pass filters h (wavelet signals). Therefore, DWT can be implemented as a perfect reconstruction filter bank (PRFB) which is completely characterized by a pairwise Quadrature Mirror Filters (QMF) l and h (Nsimba & Levada, 2019).

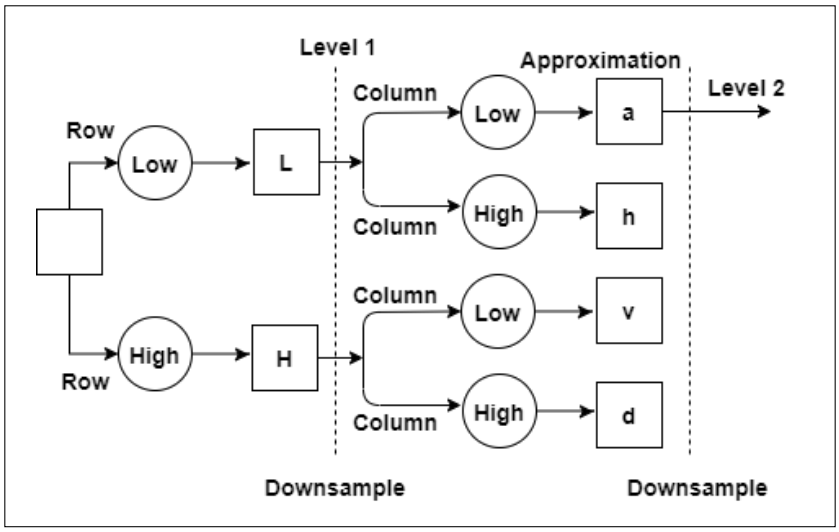

Figure 1.3 Wavelet decomposition for two-dimensional images

For a complete DWT specification, two other filters are required l and h, known as synthesis filters, and are used to reconstruct the original signal from wavelet coefficients. This process is known as the Inverse Discrete Wavelet Transform (IDWT). Figure 1.3 shows the decomposition of a two-dimensional image.

Furthermore, any wavelet decomposition of a bi-dimensional image involves four sub-bands, to wit, LL (Approximation), LH (Horizontal Detail), HL (Vertical Detail), and HH (Diagonal Detail). The sub-band image LL is but used for DWT computation in the next scale. Additionally, there are many types of the wavelet transform, such as Haar, Daubechies, Symlets, etc. The choice of which to be used depends on how it fits best in a given situation.

1.5 Ecological Diversity Indices

As stated in Magurran (2004a), and Magurran (2013), diversity is a term often used in ecology, and the purpose of its indices is to describe the variety of species present in a community or region. The concept of community is presented as a set of species that occur in a certain place and time.

Measurements such as variance and standard deviation, which are calculated in statistical studies, show values that measure quantitative variability, while diversity indices describe qualitative variability.

Diversity is measured through two variants, namely, species richness (which represents the number of species contained in a given region) and relative abundance (which refers to the number of individuals of a given species in a region) (Pianka, 2011; Rousseau, Van Hecke, NIjssen & Bogaert, 1999a).

Phylogeny is a branch of biology responsible for studying the evolutionary relationships between species, for verifying the relationships between them, in order to determine possible common ancestors. A phylogenetic tree, or simply phylogeny, is a tree where leaves represent organisms and internal nodes represent supposed ancestors. The edges of the tree denote evolutionary relationships.

In general, diversity cannot be measured only with the use of data such as abundance and species richness; thus, the phylogenetic parameter is increasingly being inserted in this calculation (Clarke & Warwick, 1998a).

Phylogenetic diversity is a measure of community diversity that incorporates the phylogenetic relationships of species (Magurran, 2013). The combination of species abundance with phylogenetic proximity to generate a diversity index is denoted taxonomic diversity. Taxonomy is the science that deals with classification (creating new taxa), identification (allocation of lineage within species) and nomenclature (Vandamme *et al.*, 1996).

One way to represent the phylogenetic tree is through the dendrogram, which is a diagram that represents the ancestral relationships between organisms. This type of tree describes the evolutionary sequence of some primates.

The phylogenetic tree combined with phylogenetic diversity indices is used in biology to compare behavior patterns between species in different areas. The phylogenetic diversity indices were chosen due to their potential in characterizing a given region/image.

The richness of details obtained with each group of indices is essential for the composition of the descriptors being proposed in this research.

It is important to mention that these indices are complementary, that is, a group of indices is able to measure some property that another group cannot achieve. For example, the phylogenetic diversity indices based on species richness (biodiversity measurements), and the group of indices based on the distance between pairs of species (taxonomic indices). The first is able to measure properties directly related to the species, such as its relative abundance, that is, the number of individuals of a species. The second, on the other hand, is capable of measuring the kinship relationships that certain species have, such as the number of common ancestors that exist between certain species.

1.5.1 Biodiversity and its Measurements

Biodiversity is defined as a variety within and among life forms on an ecosystem or a site; and is measured as a combination of two components, to wit, richness, and evenness across species (Rousseau *et al.*, 1999a). The former component is also referred to as species richness, standing for the number of groups of functionally related individuals, and the latter denotes the proportions of species or functional groups present in an ecosystem or community. Besides these components, another type of indices is taxonomic indices, which consider the taxonomic relationships between different organisms in an ecosystem. Moreover, taxonomic diversity reflects the average taxonomic distance between any two organisms, randomly chosen from a sample. Such a distance can be understood as the length of the path connecting these two organisms along the branches of a phylogenetic tree (Sohier, 2019).

The diversity can be employed to represent variation in several forms, to wit, genetic, life form, and functional group. It is worth mentioning that diverse communities are often a sign of fragmented sites where much of species richness is contributed by disturbance species (Rousseau *et al.*, 1999a; Solow & Polasky, 1994). Different objective measures have been brought into existence as a means to empirically measure biodiversity. The fundamental idea of a diversity

index is to quantify biological variability, which, in turn, can be used to compare biological entities, composed of direct components, in whether space or time (Sohier, 2019). Biodiversity can be expressed or monitored at different scales/spaces: (i) alpha diversity, which is the diversity within a particular ecosystem, that is, the richness and evenness of individuals within a community; (ii) beta diversity, which is the diversity between ecosystems, involving comparing the number of taxa that are unique to each of the ecosystems. In other words, it expresses the diversity between habitats; and (iii) gamma diversity, which measures the overall diversity for different ecosystems within a region, that is, the diversity of habitats within a region. More details concerning these three types of indices can be found in Jost (2007). Figure 1.4 illustrates an example of Alpha, Beta, and Gamma diversities in different sites.

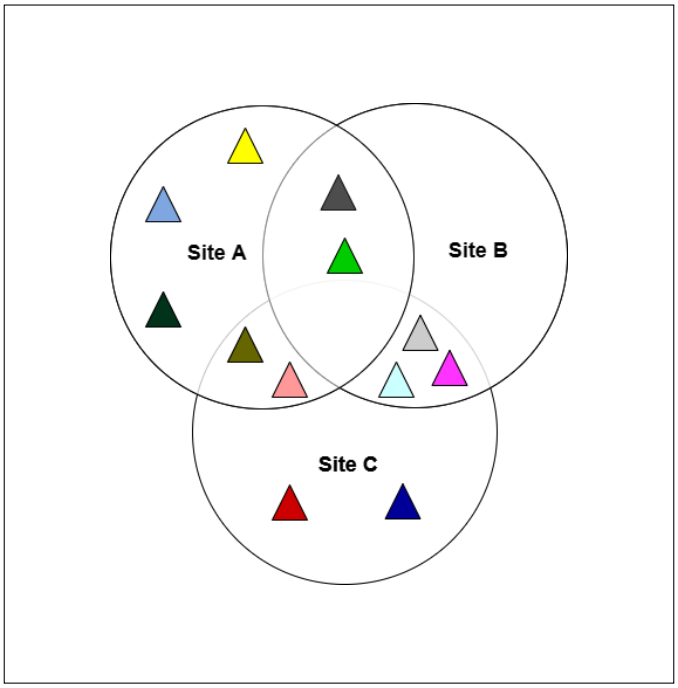

Figure 1.4 Schematic visualization of the species richness. Alpha Diversity of Site A = 7 species, Site B = 5 species, Site C = 7 species. Beta Diversity is observed between Site A and C with 10 species that differ between them and only 2 species in common. Gamma diversity is 3 habitat with 12 species total diversity

Some alpha diversity measures, including measures of richness, dominance, and evenness (SDR-IV, 2020) are described as follows:

Margalef's diversity index ($\mathbf{D_{Mg}}$) (Clifford, Stephenson et al., 1975; Magurran, 2004a) and **Menhinick's diversity index ($\mathbf{D_{Mn}}$)** (Whittaker, 1972) are both the ratio between the number of species recorded (S) and the total number of individuals in the sample (N):

$$\mathrm{D_{Mg}} = \frac{S - 1}{\ln N} \tag{1.7}$$

$$\mathrm{D_{Mn}} = \frac{S}{N} \tag{1.8}$$

Berger-Parker dominance (May, Cody & Diamond, 1975) is the ratio between the number of individuals in the most abundant species N_{max} and the total number of individuals in the sample (N):

$$\mathrm{d_{BP}} = \frac{N_{max}}{N} \tag{1.9}$$

Fisher's alpha diversity metric (Fisher, Corbet & Williams, 1943a; Magurran, 1988; SDR-IV, 2020) is:

$$\mathrm{d_F} = \alpha \ln\left(1 + \frac{N}{\alpha}\right) \tag{1.10}$$

where F denotes the number of operational taxonomic unit (groups of closely related species) and N is the total number of individuals in the sample, and the index is the alpha parameter, and α is approximately equal to the number of species represented by a single individual.

Kempton-Taylor index of alpha diversity ($\mathbf{d_{KT}}$) (Kempton & Taylor, 1976) measures the interquartile slope of the cumulative abundance curve, where n_r is the total number of species with abundance R; S is the total number of species in the sample; R_1 and R_2 are the 25% and 75% quartiles of the cumulative species curve; n_{R_1} is the number of individuals in the class where R_1 falls; n_{R_2} is the number of individuals in the class where R_2 falls:

$$\mathrm{d_{KT}} = \frac{\frac{1}{2}n_{R_1} + \sum_{R_1+1}^{R_2-1} n_r + \frac{1}{2}n_{R_2}}{\log(\frac{R_2}{R_1})} \tag{1.11}$$

McIntosh's evenness measure ($\mathbf{e_M}$) (Heip & Engels, 1974) is the ration between the number of individuals in the i^{th} species and the total number of individuals (N) plus the number of species in the sample (S):

$$\mathrm{e_M} = \sqrt{\frac{\sum n_i^2}{(N - S + 1)^2 + S - 1}} \tag{1.12}$$

where n_i denotes the number of individuals in the i^{th} species, N is the total number of individuals, and S is the number of species in the sample.

Shannon-Wiener diversity index ($\mathrm{d_{SW}}$) (SDR-IV, 2020) is defined as the proportion of individuals of species i in terms of species abundance (S):

$$\mathrm{d_{SW}} = -\sum_{i=1}^{S} (p_i \ \ln \ p_i) \tag{1.13}$$

where p_i denotes the proportion of individuals in the i-th species.

The **Brillouin index** ($d_{\mathbf{HB}}$) is defined as:

$$d_{\mathrm{HB}} = \frac{\ln N! - \sum_{i=1}^{S} \ln n_i!}{N} \tag{1.14}$$

where N is defined as the number of individuals, S is the number of species, and n_i is defined as the number of individuals in the i-th species.

The **Strong's dominance index** (d_{Dw}) is defined as:

$$d_{\mathrm{Dw}} = \max_i [(\frac{b_i}{N}) - \frac{i}{S}] \tag{1.15}$$

where b_i is the sequential cumulative totaling of the i-th distinct species values ranked from largest to smallest.

The expression in brackets is computed for all species, and max_i denotes the maximum value in brackets for any species.

The **Simpson's index** ($d_{\mathbf{C}}$) and the **Enspie index** ($d_{\mathbf{ENS}}$) are defined as:

$$d_{\mathrm{C}} = 1 - \sum_{i=1}^{S} p_i^2 \tag{1.16}$$

$$d_{\mathrm{ENS}} = \frac{1}{\sum_{i=1}^{S} p_i^2} \tag{1.17}$$

where p_i is the proportion of the community represented by species i.

The **McIntosh dominance diversity index** ($d_{\mathbf{McInt}}$) (McIntosh, 1967) is defined as:

$$d_{\mathrm{McInt}} = \frac{N - U}{N - \sqrt{N}} \tag{1.18}$$

where N represents the total number of individuals in the sample, $U = \sqrt{\sum {n_i}^2}$, and n_i represents the number of individuals in the i-th species.

The **Chao1 richness estimator** ($e_{\mathbf{CR}}$) (Chao, 1984; Eren, Chao, Hwang & Colwell, 2012) uses only the number of singletons (F_1) and doubletons (F_2) and the observed richness (S_{obs}) to write the following estimator for the class richness:

$$e_{\mathrm{CR}} = S_{obs} + \frac{F_1^2}{2F_2} \tag{1.19}$$

where F_1 and F_2 are the count of singletons and doubletons species in the sample, respectively.

According to Chao, Shen & Hwang (2006), in the presence of many class abundance distributions, this estimator, originally derived as an estimate of minimum possible richness (number of individuals pertaining to that specific species), is much sharp if the reference sample size is large enough. This corroborates the reason for its use as a valid estimator for large number of species.

The **Gini coefficient** ($\mathbf{e_{GC}}$) (Gini, 1912) is defined as:

$$\mathrm{e_{GC}} = \frac{2}{mS^2}\left(\sum_{i=1}^{n}(S+1-i)x_i\right) - \frac{1}{S} \tag{1.20}$$

where x_i is the number of individuals of the i-th species ranked from least to most abundant, $i \in [1, S]$ and m is the mean abundance of a species – the mean of the x_i values. The Gini coefficient measures income inequality, but can also be used to measure any form of uneven distribution. It ranges between 0 and 1, where 0 denotes a perfect inequality and 1 denotes a perfect equality, where each species has the same number of individuals.

The **Heip's evenness** ($\mathbf{e_{HE}}$) is defined as:

$$\mathrm{e_{HE}} = \frac{e^H - 1}{S - 1} \tag{1.21}$$

where H is the Shannon-Wiener entropy of counts using logarithm base e.

The **Pielous evenness** ($\mathbf{e_{J'}}$) is defined as:

$$\mathrm{e_{J'}} = \frac{H}{\ln S} \tag{1.22}$$

Simpsons evenness e_{SE} is defined as:

$$e_{SE} = \frac{\frac{1}{D}}{S_{obs}} \tag{1.23}$$

where D is dominance and S_{obs} is the number of observed species.

1.5.2 Taxonomic Indices

The quantification of ecosystem diversity is generally required for practices of environmental planning. The traditional ecological diversity indices presented in previous section are based on the abundances of species present in the community. Nevertheless, such indices may be insensitive to taxonomic or similar differences. With equal species abundances they measure but the species richness (species number) (Izsák & Papp, 2000). Assemblages with the same species richness may either comprise species that are closely related to one another taxonomically or they may be more distantly related (Rogers, Clarke & Reynolds, 1999).

Taxonomic indices consider the taxonomic relation between different individuals in an ecosystem. The diversity thereof reflects the average taxonomic distance between any two individuals, randomly chosen from a sample. The distance can represent the length of the path connecting these two individuals along the branches of a phylogenetic tree (Ricotta, 2002; Rogers *et al.*, 1999).

Gibson, Barnes & Atkinson (2001) proposed the distinctiveness index describing the average taxonomic distance between two randomly chosen individuals through the phylogeny of all the species in a dataset. This distinctiveness possesses different forms (Sohier, 2019), to wit, taxonomic diversity, and taxonomic distinctness, which are defined as follows:

Taxonomic diversity: this index includes aspects of taxonomic relatedness and evenness.

$$\Delta = \frac{\sum \sum_{i<j} w_{ij} x_i x_j}{N\,(N-1)\,/2} \tag{1.24}$$

Taxonomic distinctiveness: the measure of pure taxonomic relatedness.

$$\Delta^* = \frac{\sum \sum_{i<j} w_{ij} x_i x_j}{\sum \sum_{i<j} x_i x_j} \tag{1.25}$$

where $x_i = N_{pi}$, N and w_{ij} represent the abundance of the i-th species in the sample, the number of individuals in the sample, and the 'distinctness weight' given to the path length linking species i and j in the hierarchical classification, respectively.

Average taxonomic distinctiveness: in case only presence or absence data is considered, both Δ and Δ^* converge to the same statistic Δ^+, which can be seen as the average taxonomic path length between any two randomly chosen species (Clarke & Warwick, 1998b),

$$\Delta^+ = \frac{\sum \sum_{i<j} w_{ij}}{S(S-1)/2} \tag{1.26}$$

where S is the number of species.

The studies that verify the relationship of distances between pairs are based on a distance matrix between all species of a community. The distances can be based on morphological or functional differences (Izsáki & Papp, 1995), on the length of the branches of the phylogenetic relationships based on molecular data (Pavoine, Ollier & Dufour, 2005; Solow, Polasky & Broadus, 1993) or, if the lengths of the branches are not known, on the number of nodes that separate each pair of species (Faith, 1992).

The values within the distance matrix can be interpreted as the distinction between each pair of species or of each species in particular for all others (Izsáki & Papp, 1995; Rao, 1982).

These indices make one up in the distances between pairs of species. Sum of Phylogenetic Distances; intensive quadratic entropy and extensive quadratic entropy (Izsáki & Papp, 1995); average distinctness (Clarke & Warwick, 1998a); total taxonomic distinctness (Clarke, Gorley, Somerfield & Warwick, 2014); and average distance (Faith, 1994).

Sum of Phylogenetic Distances: it represents the sum of phylogenetic distances between pairs of species.

$$s_{PD} = \left(\frac{s(s-1)}{2}\right) \frac{\sum \sum_{m<n^2} mn^a m^a n}{\sum \sum_{m<n^a} m^a n} \tag{1.27}$$

Average Distance from Nearest Neighbor (Vellend, Cornwell, Magnuson-Ford & Mooers, 2011): it represents the average distance to the nearest taxon.

$$d_{NN} = \sum_{s}^{m} min\,(d_{mn} a_m) \tag{1.28}$$

where d_{mn} $(m, n = 1, \ldots, s)$ is the distance from species m to species n; a is the species abundance and s is the number of species.

Intensive Quadratic Entropy: it represents the number of species and their taxonomic relationships.

$$e_{IQ} = \frac{\sum d_{ij}}{s^2} \tag{1.29}$$

Extensive Quadratic Entropy: it represents the sum of the differences between species.

$$e_{EQ} = \sum d_{ij} \tag{1.30}$$

Total Taxonomic Distinctness: it represents the average phylogenetic distinctiveness added across all species.

$$\mathrm{d_{TT}} = \sum i \frac{\sum_{i \neq j} d_{ij}}{s - 1} \tag{1.31}$$

where d_{ij} $(i, j = 1, \ldots, s)$ is the distance from species i to species j and s is the number of species.

As a further factor of consideration, to apply the taxonomic indexes of a set of species (the joint dissimilarity of species or set of pairwise distances between the species in the set), the species distance is computed by means of the topology taxonomic tree. The topological distance, which is the number of the edge between two species in the Linnaean taxonomic tree, is the cumulative branch length of the full phylogenetic tree. An example of a taxonomic tree along with its species distance matrix is shown in Figure 1.5.

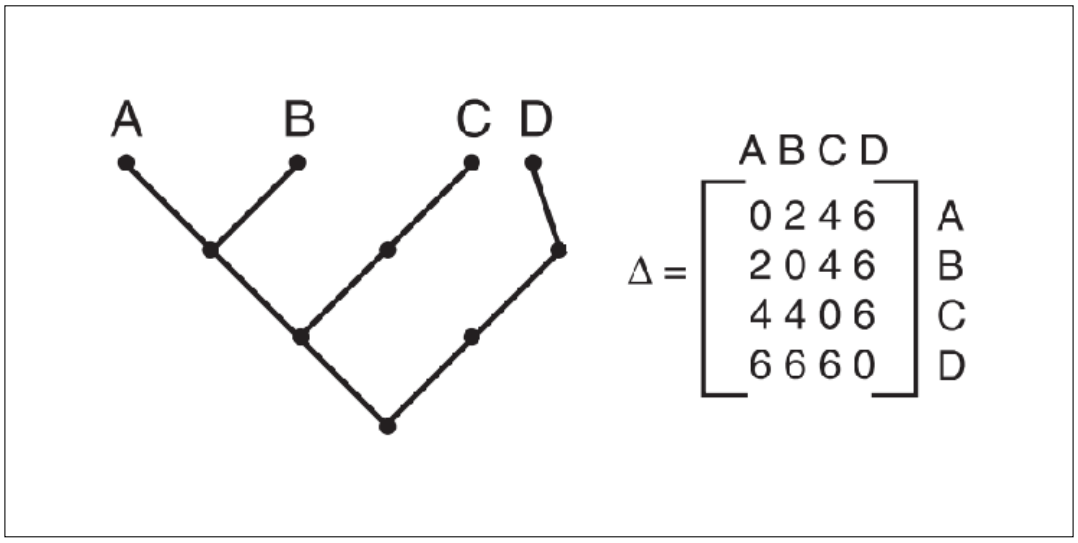

Figure 1.5 Generic example of taxonomic tree and its respective distance matrix. This matrix shows how cumulative branch length, which corresponds to taxonomic distances, is calculated
Adapted from Ricotta (2004)

1.6 Measurements based on Shannon Entropy and Multi-information

1.6.1 Shannon Entropy

The concept of entropy is used in thermodynamics, statistical mechanics, and information theory. Entropy is considered a measure of uncertainty and of the information necessary for, in any process, to be able to limit, reduce, or eliminate the uncertainty.

It turns out that the concept of information and that of entropy are basically related to each other. In the field of information theory, Shannon entropy is related to a random variable and is used to measure the uncertainty of an information source (Shannon, 1948).

Suppose that an event (random variable) has an initial degree of indeterminacy equal to k (i.e. there are k possible states) and suppose all states are equiprobable. Then the probability that one of these combinations occurs will be $p = \frac{1}{k}$. Then we can represent the expression c_i as:

$$c_i = \log_2(k) = \log_2\left[\frac{1}{\frac{1}{k}}\right] = \log_2(\frac{1}{p}) = log_2(1) - \log_2(p) = -\log_2(p) \tag{1.32}$$

If now each of the k states has a probability p_i, then the entropy will be given by the weighted sum of the amount of information:

$$H = -p_1 \log_2(p_1) - -p_2 \log_2(p_2) - \cdots - p_k \log_2(p_k) = -\sum_{i=1}^{k} p_i \log_2(p_i) \tag{1.33}$$

Therefore, the entropy of a message X, denoted by $H(X)$, is the weighted average value of the amount of information of the various states of the message; in other words, it represents a measure of the mean uncertainty about a random variable and therefore the amount of information.

$$H(X) = -\sum_{i=1}^{k} p(x_i) \log_2 p(x_i) = \sum_{i=1}^{k} p(x_i) \log_2(\frac{1}{p(x_i)}) \tag{1.34}$$

1.6.2 Multi-information

In information theory, multi-information (Studenỳ & Vejnarová, 1998) (a.k.a total correlation (Watanabe, 1960)), is among the various generalizations of the mutual information, and it quantifies the dependency among a group of random variables.

Given discrete random variables $X_1, \cdots, X_n$, the multi-information $T(X_1, \cdots, X_n)$ is defined as:

$$T(X_1, \cdots, X_n) = \left(\sum_{i=1}^{n} H(X_i) \right) - H(X_1, \cdots, X_n) \tag{1.35}$$

where $H(x)$ denotes the entropy and $H(x_i, \cdots, x_n)$ denotes the joints entropy.

Both Shannon entropy and total multi-information are found in the literature to be used for clustering and feature selection algorithms (Watanabe, 1960). In this thesis, both measurements are used for feature extraction, where an image is considered an array containing discrete random variable realizations.

1.7 Feature Selection

Feature selection is the process of selecting a subset of relevant features for use in model building in machine learning. It is desired that the subset of features promotes better performance than the original set. This is necessary for simplifying models to make them easier for users to interpret, avoiding the curse of dimensionality, and reducing data correlation and data volume. The fundamental assumption while using a feature selection technique is that the data comprises various redundant or irrelevant information and can be removed without incurring much loss of information. A feature selection algorithm can be seen as a coupling of search techniques to propose new feature subsets and an evaluation that measures scores provided by different subsets of features.

Automatic feature selection methods are helpful when a large set of features is available and a suitable subset must be chosen. In addition to being a type of dimensionality reduction, data fusion from multiple data models is an important application. Given a set of features, automatic feature selection attempts to select a subset of size m, where ($m < N$), which maximizes a criterion function $f(x)$.

To illustrate, let $f(x) = 1 - E$, where E is the rate or probability of error of a classifier. The greater the criterion function, the less redundancy between the characteristics and the easier it is to discriminate patterns of different classes. The feature selection algorithm will be able to reduce the dimensionality in this manner, resulting in the smallest possible drop in the power of class discrimination by a classifier in the feature space. Furthermore, applying a good feature selection algorithm reduces the number of training samples required to obtain good results with a classifier, thereby reducing the dimensionality problem.

In addition to selecting the criterion function, an appropriate dimensionality of the reduced feature space must also be determined. A simple solution to this problem is to perform feature selection for different values of m. According to Jain, Duin & Mao (2000), the authors argue that in practical problems where $|T|$ is the size of the set training, the dimensionality problem can be avoided by using fewer than $|T|/10$ features.

Despite the importance of feature selection, there are no clear rules or methods for doing so in each application, especially when the number of features provided is vast. As a result, many feature selection methods have been developed over time (Kumar & Minz, 2014; Remeseiro & Bolon-Canedo, 2019; Venkatesh & Anuradha, 2019).

To get the best performance out of a classifier, one must first determine the ideal dimensionality for a given pattern recognition problem. To accomplish this, a simple trial-and-error dimensionality strategy can be used, employing a dimensionality reduction method (including feature extraction and selection) until the maximum performance point of a classifier is reached. In this strategy, dimensionality reduction tests are run to generate feature subspaces of various sizes until the dimensionality that minimizes the classification error is found.

The following sections present a few categories of feature selection and feature transformation techniques.

1.7.1 Linear and Non-linear Techniques

One of the commonly used techniques for feature transformation and dimensionality reduction is principal component analysis (PCA) (Dunteman, 1989), a family of techniques for handling high-dimensional data that exploits the dependencies between variables to represent them more compactly without losing relevant information. It is a second-order statistical method and optimal by maximizing the variance of the new compact representation Y and minimizing the mean square error between the original data X and the new compact representation Y. Although the dimensionality reduction by the PCA is optimal from the point of view of data compression, owing to the fact that it minimizes the mean squared error between the original representation and the new representation, it is not optimal for discriminating data in classes.

Besides being a linear method, PCA assumes that data are found in a Euclidean subspace of R^n. In many cases, however, linear methods cannot learn the geometric structure of the data. Therefore, the hypothesis that data are found in Euclidean space is not valid. It is, therefore, necessary to find a more suitable metric than the Euclidean distance (metric learning). Thus, manifold learning can be applied to overcome this limitation. A manifold is a subspace with curvature, and Euclidean distance fails to capture the geometric properties of such sets. Some of the manifold learning techniques are Laplacian eigenmaps, ISOMAP, locally linear embedding (LLE), and multidimensional scaling (MDS) (Cox & Cox, 2008). Manifold learning is a popular approach for nonlinear dimensionality reduction, and it uses geodesic distance to understand the data, amongst other advantages.

1.7.2 Filter, Wrapper, and Embedded Strategies

There is another subcategory of methods for feature selection known as Filter, Wrapper, and Embedded. The methods based on the Filter strategy establish evaluation metrics and then

eliminates a subset of features before the construction of the classification model (Tuv, Borisov, Runger & Torkkola, 2009). Thus, the feature selection is carried out first. Afterward, the necessary steps for training the classifiers are followed. This is in line with the principle that methods based on the Filter strategy do not incorporate a classifier in the search process and, therefore, do not require labeled data (classes) (Habermann, 2018). Notwithstanding, methods based on the Filter strategy are not limited to being unsupervised. It is possible to use labeled samples to establish metrics for the attributes without using a classifier. Otherwise, they would be characterized as based on the strategy Wrapper.

Wrapper methods have been widely used for hyperspectral images for band selection. Wrapper-based methods use a classifier and directly evaluate the weight of each feature (Tuv *et al.*, 2009). In this strategy, the selection occurs during the training of the classifier. Then, for each band or set of bands added or removed from the possible combinations, the classifier must be trained again, generating a model that is evaluated afterward (Habermann, 2018). A disadvantage of this strategy relative to Filter is its slowness due to repetitive training processes. In addition, sets with a large number of samples may be needed to be performed, especially when using parametric classifiers.

Typically, researchers using supervised band selection have to devise methods capable of handling little training data. This can be challenging for Filter methods based on classification and class separation measures. However, the scarcity of training data can worsen the problem for methods based on the Wrapper strategy since the methods rely exclusively on classifiers to generate results. In compensation for this disadvantage, due to the wrapper methods using the classifier incorporated in the band selection process, the selected bands can help to promote a classifier model with much higher accuracy performance compared to methods based on Filter strategy (Molina, Belanche & Nebot, 2002; Shahana & Preeja, 2016). However, bands selected by such methods can operate well only with the classifier incorporated into the band selection process. Therefore, it is recommended to use the selection with the same classifier to be used in the application for which the bands were selected (Kohavi, John et al., 1997).

Another category found in the literature for feature selection is Embedded strategy. This strategy tries to combine the qualities of the two previous strategies, using classification algorithms that have their own feature selection methods. Methods based on the Embedded strategy use all variables to generate a classification model and then analyze this model by inferring the importance of its variables (Tuv *et al.*, 2009). The best-known representatives of this strategy are based on decision trees or artificial neural networks (Mitchell, 1982).

1.7.3 Multi-objective Optimization Algorithms

Besides the previous strategies, there are feature selection methods based on multiobjective optimization algorithms. Xu, Shi & Pan (2017) proposed an incorporated rank-based multiobjective band selection (IRMOBS), a method based on the Filter strategy, unsupervised and with the objectives based on entropy, variance, and the number of bands. Saqui, Saito, De Lima, Cura & Ataky (2019) proposed an incorporated Decision-maker-based multiobjective band selection (IDMMoBS), a method based on the Wrapper strategy, and supervised for multiobjective selection of bands of hyperspectral images.

Some recent methods with a single objective, which present the most current components explored in the literature, are the trivariate mutual information-clonal selection algorithm (TMI-CSA) (Feng, Jiao, Zhang & Sun, 2013) and are based on the adequacy of the mutual information (MI) that considers the correlation between three variables, the class label, and two bands. It uses CSA as a search engine for optimization based on the Filter strategy and supervised learning. The Semi-supervised band selection approach based on TMI and graph regulation (STMIGR) proposed in (Feng, Jiao, Liu, Sun & Zhang, 2014) is similar to TMI-CSA but uses a semi-supervised learning strategy that allows the propagation of labels. Finally, the maximum information and minimum redundancy-clonal selection algorithm (MIMRCSA) proposed by Feng, Jiao, Liu, Sun & Zhang (2016) uses CSA to optimize a function derived from MI and entropy called Maximum information and minimum redundancy (MIMR). This method operates unsupervised, based on the Filter strategy, and uses the search for optimization. Xie, Li, Lei & Ke (2018) proposed the information gain - gray wolf optimizer (IG-GWO),

which operates as a supervised method with a fitness function based on the information gain of the bands, but based on the Filter strategy, as it does not incorporate a classifier in the process of band selection. The Single-layer neural network (SLN) proposed by Habermann, Fremont & Shiguemori (2019) is the method considered Filter because it uses SLN to select bands and other algorithms for classification; and genetic algorithm with Support Vector Machines (GA-SVM) (Nagasubramanian, Jones, Sarkar, Singh, Singh & Ganapathysubramanian, 2018; Zhuo, Zheng, Li, Wang, Ai & Qian, 2008) is a method based on the Wrapper strategy, supervised and uses a fitness function with a weighted average between classification accuracy and the number of bands. This combination is traditional in the literature, being explored in different ways as presented in (Nagasubramanian *et al.*, 2018; Zhuo *et al.*, 2008).

An important and common point of the presented strategies is that the majority has only one objective, feature selection, which does not imply the guarantee of effective classification performance. Methods based on multiobjective optimization demonstrated an interesting ability to deal with conflicting objectives. As in most of the related works presented above, it allows searching for solutions that achieve a good classification ability with fewer bands. In the proposed method, we explore the limitations and advantages of both single and multiobjective optimization algorithms, intending to find a solution that makes a trade-off between the number of features and classification accuracy.

1.8 Final Considerations

This chapter presented the theoretical basis used in the development of this thesis. The concepts presented are of paramount importance for the portentous understanding of the proposed methods. Concepts about texture images, digital image processing, texture analysis, multi-resolution analysis, phylogenetic diversity indices, and information theory were presented.

In the following chapters, the materials and methods applied in constructing the approaches for analysis and characterization of texture images are discussed, aiming to increase/improve the accuracy performance in classification.

CHAPTER 2

TEXTURE DESCRIPTORS BASED ON ECOLOGICAL DIVERSITY MEASURES

Given that texture constitutes a non-deterministic system of patterns, we hypothesize that textural patterns behave similarly to ecological patterns. Large populations of units can self-organize into aggregations that generate patterns from non-deterministic nonlinear processes in an ecosystem. This chapter discusses novel approaches to quantifying such a complex system of diverse patterns using ecological diversity measures, namely species diversity, richness, evenness, and taxonomic indices. The proposed methods consider an image as a species ecosystem, in which case it becomes possible to extract and compute ecological diversity measures to describe the texture. Furthermore, such approaches take advantage of the invariance characteristics of ecological patterns to construct permutation, rotation, and translation invariant descriptors. Thus, the following sections investigate how leveraging information-theoretical measures of ecological diversity indices in conjunction with measures of biodiversity can provide a robust approach for texture characterization and classification.

2.1 A Novel Bio-Inspired Texture Descriptor based on Biodiversity and Taxonomic Measures

This section introduces a novel bio-inspired texture (BiT) descriptor based on biodiversity measurements (species richness and evenness) and taxonomic distinctiveness. Ecology primarily exploits these concepts by considering patterns in ecosystems. In this thesis, textural image is considered an ecosystem, where both the biodiversity measurements and taxonomic indices are computed and quantified. The proposed approach exploits both sides of ecological diversity indices, to wit, distance-based phylogenetic diversity indices and species richness, abundance, and evenness - as a generalization.

The BiT descriptor is a generic descriptor that can characterize texture information on various images. Into the bargain, the BiT relies on the values of the indices, which can be explained and interpreted based on the related ecological concepts. Furthermore, the proposed approach also

exploits color information. We represent and describe biodiversity as the interaction of pixels with their neighborhood within each image channel (R, G, or B) and on a single RGB image. Besides, taxonomic indices and species richness measures on which the novel BiT descriptor relies are of an underlying use as they capture the all-inclusive behavior of texture image patterns. In ecology, they capture the intrinsic properties of the whole ecosystem, although the latter forms a non-deterministic complex system. The complexity, in this case, surges when causality breaks down. The aim is that this approach performs well regardless of this texture nature since biodiversity indices measurements cope with such complexity from the ecosystem perspective.

The main contribution is a novel bio-inspired descriptor that leverages species diversity, richness, and taxonomic distinctiveness to build a texture representation for classification purposes. This contribution of the thesis highlights the following points:

- Modeling each channel of a color image as an ecosystem;
- A novel bio-inspired texture (BiT) descriptor combining measurements of species diversity and richness and taxonomic distinctiveness;
- The BiT descriptor is invariant to scale, translation and permutation;
- The BiT descriptor is easy to compute and has low computational complexity;
- The BiT descriptor is a generic texture descriptor that performs well on different image categories, such as natural textures and medical images.

In the following sections, we describe how these ideas were applied to extract efficient features for the texture classification task.

2.1.1 Images as Ecosystems

In order for the ecological concepts to be employed in this approach, we assume that an image is an abstract model of an ecosystem where:

- Gray levels of pixels in an image correspond to the species in an ecosystem;
- Pixels in an image correspond to the individuals in an ecosystem;

- The number of different gray levels in an image corresponds to species richness in an ecosystem;
- The number of pixels per gray level corresponds to species abundance;
- The number of distinct gray levels in a specific region of an image corresponds to relative abundance in that region.

Another factor is that both the patterns in an ecosystem and texture images form a non-deterministic system.

Figure 2.1 illustrates an example of an ecosystem with three species, where there are six individuals of white species, five individuals of gray species, and five individuals of black species.

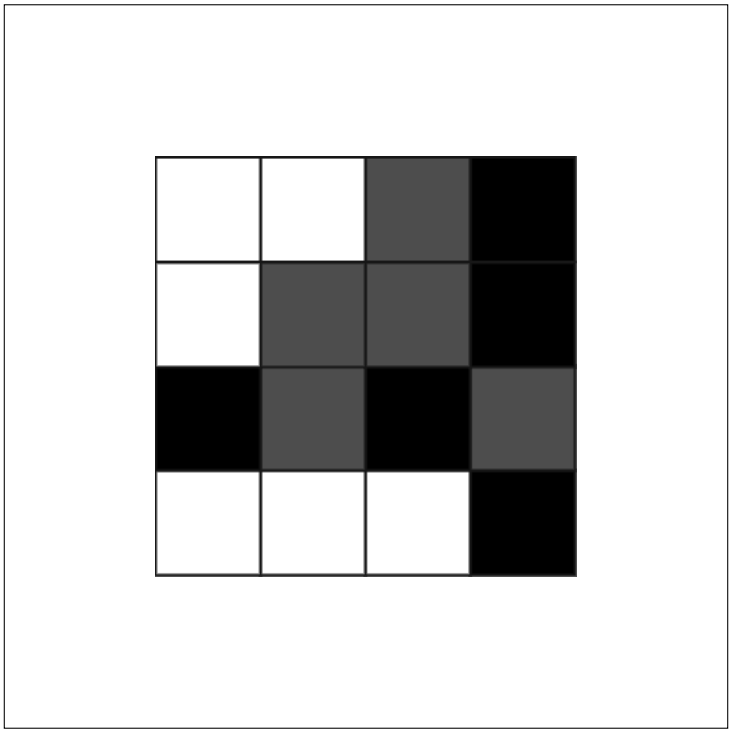

Figure 2.1 A gray-level image as an abstract model of an ecosystem of three species (three gray levels): white (6 individuals), gray (5 individuals) and black (5 individuals)

2.1.2 Biodiversity and its Measurements

Biodiversity is defined as the variety within and among life forms on an ecosystem or a site, and it is measured as a combination of richness and evenness across species (Rousseau *et al.*, 1999a). Diversity can represent variation in several forms, such as genetic, life form, and functional groups. It is worthy of mention that diverse communities are often a sign of fragmented sites where much species richness is contributed by disturbance species (Rousseau *et al.*, 1999a). Different objective measures have been proposed as a means to measure biodiversity empirically. The fundamental idea of a diversity index is to quantify biological variability, which, in turn, can be used to compare biological entities composed of direct components in space or time (Sohier, 2019). Biodiversity can be expressed or monitored at different scales and spaces: alpha diversity, beta diversity, and gamma diversity. Jost (2007) presents more details on these three types of indices.

In texture analysis, the statistical approach defines texture as a set of local measurements extracted from the pattern, favoring image description through statistical rules that govern the description and the relationship between the different gray levels. To characterize texture, we use phylogenetic diversity indices, most of which are statistical measures.

In ecology, diversity cannot be generally measured solely through data such as abundance and species richness; the phylogenetic parameter is increasingly being incorporated into this analysis (Clarke & Warwick, 1998b). Phylogenetic diversity measures a community's diversity that considers species relationships (Magurran, 2004b). The combination of species abundance and phylogenetic proximity yields a diversity index known as taxonomic diversity (Da Silva & Batalha, 2006).

The set of distances between pairs of species accumulated from taxonomic trees is used to calculate taxonomic diversity. In biology, phylogenetic trees and phylogenetic diversity indices are used to compare behavior patterns between species from different areas.

It is important to note that phylogenetic diversity indices based on species richness and the distance between species pairs can measure species-specific properties such as relative abundance and parenthood relationships.

The following sections present both sides of ecological diversity. It is noteworthy that diversity indices based on species richness capture global information, whereas those based on parenthood relationships, capture local (spatial) information. In other words, diversity measures are calculated globally. Nevertheless, taxonomic indices capture the spatial arrangement of pixels.

2.1.2.1 Diversity Measures

Diversity measurements rely on three assumptions (Magurran, 2004a): (i) all species are equal – richness measurement makes no distinctions among species and considers species exceptionally abundant in the same way as those extremely rare. In other words, no species is excluded from computing the ecosystem richness due to its abundance; (ii) all individuals are equal – there is no distinction between the largest and the most minor individual; however, in practice, the least animals can often escape, for instance, by sampling with nets. This does not necessarily apply to taxonomic and functional diversity measures; (iii) species abundance is recorded using appropriate and comparable units.

We can translate such assumptions to our abstract model as (i) all gray levels are equally taken into account regardless of the number of pixels – richness measurement makes no distinctions among gray levels and treats the gray levels that are exceptionally abundant in the same way as those significantly less represented; In other words, all gray levels within an image are taken into account for further calculation, regardless of how non-representative some of them are; (ii) all pixel values are equal – there is no distinction between the largest and the smallest pixel value; (iii) gray-level abundance has to be recorded using appropriate and comparable units such as the intensity.

Some alpha diversity measures used to build the BiT descriptor, including measures of richness, dominance, and evenness (Magurran, 2004b; SDR-IV, 2020), are described as follows. They

represent the diversity within a particular ecosystem: the richness and evenness of individuals within a community. All these indices are computed on a gray-level image of dimensions m and n denoted as $\mathbf{I}_{m\times n}$. It is worth noting that we will henceforth use image analysis terms instead of ecosystem ones.

Margalef's ($\mathrm{d_{Mg}}$) (Clifford *et al.*, 1975; Magurran, 2004a) and **Menhinick's** ($\mathrm{d_{Mn}}$) (Whittaker, 1972) **diversity index** are both the ratio between the number of species (S) and the total number of individuals in the sample (N):

$$\mathrm{d_{Mg}} = \frac{S-1}{\ln N} \tag{2.1}$$

$$\mathrm{d_{Mn}} = \frac{S}{N} \tag{2.2}$$

where, S and N denote the number of different gray levels and the total number of pixels in an image, respectively.

Berger-Parker dominance ($\mathrm{d_{BP}}$) (May *et al.*, 1975) is the ratio between the number of individuals in the most abundant species (N_{max}) and the total number of individuals in the sample:

$$\mathrm{d_{BP}} = \frac{N_{max}}{N} \tag{2.3}$$

where N_{max} denotes the number of pixels belonging to the most frequent gray level in an image.

Fisher's alpha diversity metric ($\mathrm{d_F}$) (Fisher, Corbet & Williams, 1943b; SDR-IV, 2020) denotes the number of operational taxonomic units, that is, groups of closely related individuals, and it is defined as:

$$\mathrm{d_F} = \alpha \ln\left(1 + \frac{N}{\alpha}\right) \tag{2.4}$$

where α is approximately equal to the number of gray levels represented by a single pixel.

Kempton-Taylor index of alpha diversity ($\mathrm{d_{KT}}$) (Kempton & Taylor, 1976) measures the interquartile slope of the cumulative abundance curve. R_1 and R_2 are the 25% and 75% quartiles

of the cumulative species curve, respectively, n_r is the number of species with abundance R, n_{R_1} is the number of individuals in the class where R_1 falls, and n_{R_2} is the number of individuals in the class where R_2 falls:

$$\mathrm{d_{KT}} = \frac{\frac{1}{2}n_{R_1} + \sum_{R_1+1}^{R_2-1} n_r + \frac{1}{2}n_{R_2}}{\log \frac{R_2}{R_1}} \tag{2.5}$$

where n_r denotes the number of gray levels with abundance R; R_1 and R_2 are the 25% and 75% quartiles of the cumulative gray level curve; n_{R_1} is the number of pixels in the class where R_1 falls; n_{R_2} is the number of pixels in the class where R_2 falls.

McIntosh's evenness measure ($\mathrm{e_M}$) (Heip & Engels, 1974) is the ratio between the number of individuals in the i-th species and the total number of individuals, and the number of species in the sample:

$$\mathrm{e_M} = \sqrt{\frac{\sum_{i=1}^{S} n_i^2}{(N - S + 1)^2 + S - 1}} \tag{2.6}$$

where n_i denotes the number of pixels of the i-th gray level (the summation is over all gray levels).

Shannon-Wiener diversity index ($\mathrm{d_{SW}}$) (SDR-IV, 2020) is defined as the proportion of individuals of species i in terms of species abundance (S):

$$\mathrm{d_{SW}} = -\sum_{i=1}^{S} (p_i \ \ln \ p_i) \tag{2.7}$$

where p_i denotes the proportion of pixels in the i-th gray level.

2.1.2.2 Taxonomic Indices

The ecological diversity indices presented in the previous section rely on the richness and abundance of species present in a community. Nevertheless, such indices may be insensitive

to taxonomic differences or similarities. With equal species abundances, they measure merely the species richness. Assemblages with the same species richness may either comprise species closely related taxonomically or more distantly related (Rogers *et al.*, 1999)

Taxonomic indices consider the taxonomic relation between different individuals in an ecosystem. The diversity thereof reflects the average taxonomic distance between any two individuals randomly chosen from a sample. The distance can represent the length of the path connecting these two individuals along the phylogenetic tree branches (Rogers *et al.*, 1999). Taxonomic diversity and taxonomic distinctiveness define the relationship between two organisms in an existing phylogeny in a community (Clarke & Warwick, 1998a; Gibson *et al.*, 2001), and three key factors characterize them: (i) the number of individuals; (ii) the number of species; (iii) the structure of species connection, that is, the number of edges. Furthermore, Gibson *et al.* (2001) also proposed the distinctiveness index describing the average taxonomic distance between two randomly chosen individuals through the phylogeny of all species in a sample. This distinctiveness may be represented as taxonomic diversity and taxonomic distinctness (Sohier, 2019), which is described as follows.

Taxonomic diversity (Δ) (Clarke & Warwick, 1998a) includes aspects of taxonomic relatedness and evenness. In other words, it considers the abundance of species (number of different gray levels) and the taxonomic relationship between them, and whose value represents the average taxonomic distance between any two individuals (pixels) chosen at random from a sample.

$$\Delta = \frac{\sum_{i=0}^{S} \sum_{i<j}^{S} w_{ij} x_i x_j}{\frac{N\,(N-1)}{2}} \tag{2.8}$$

where x_i, x_j, and w_{ij} represent the number of pixels that have the i-th gray level in the image, the number of pixels that have the j-th gray level in the image, and the 'distinctness weight' (distance) given to the path length linking pixels i and j in the hierarchical classification, respectively, and $i, j = 0, \ldots, S$.

Taxonomic distinctiveness (Δ^*) is a measure of pure taxonomic relatedness. It represents the average taxonomic distance between two individuals (pixels), constrained to different species (gray levels).

$$\Delta^* = \frac{\sum_{i=0}^{S}\sum_{i<j}^{S} w_{ij}x_i x_j}{\sum\sum_{i<j} x_i x_j} \tag{2.9}$$

Different ecological studies, particularly large-scale ones, employ species richness to measure biodiversity. Nevertheless, species richness as the sole reflection of biodiversity can present limitations as all species are treated equally without considering phylogenetic relationships. The literature shows that phylogenetic relationships are essential factors determining species' extinction. Thus, phylogenetic information may be a better indicator of the preservation value than merely the species richness. Studies verifying the distance relationship between the species pairs depend on a distance matrix computed for all community species. In ecology, this distance matrix relies on either functional or morphological differences (Izsáki & Papp, 1995), on the length of the branches of the phylogenetic relationships based on molecular data (Pavoine *et al.*, 2005). Accordingly, if the branches' length is unknown, such distances rely on the number of nodes that separate each pair of species (Faith, 1992). Therefore, the distance matrix values can be interpreted as the distinctness between each pair of species or between each particular species vis-à-vis all others (Izsáki & Papp, 1995).

Sum of Phylogenetic Distances (s_{PD}) represents the sum of phylogenetic distances between pairs of species.

$$s_{PD} = \left(\frac{S(S-1)}{2}\right)\frac{\sum\sum_{i<j^2} ij^a i^a j}{\sum\sum_{i<j^a} i^a j} \tag{2.10}$$

where i and j denote two distinct gray levels, and a is the number of pixels that have such gray levels.

Average Distance from the Nearest Neighbor (d_{NN}) (Vellend *et al.*, 2011) represents the average distance to the nearest taxon[1].

$$d_{NN} = \sum_{i}^{S} \min\left(d_{ij}, a_i\right) \tag{2.11}$$

where $d_{ij}(i, j = 1, \ldots, S)$ is the distance between the species (gray levels) i and j; a is the abundance of the species, and S is the total number of species (gray levels).

Extensive Quadratic Entropy (e_{EQ}) represents the sum of the differences between gray levels.

$$e_{EQ} = \sum_{i \neq j}^{S} d_{ij} \tag{2.12}$$

Intensive Quadratic Entropy (e_{IQ}) represents the number of species and their taxonomic relationships. It aims at establishing a possible link between the diversity indices and the biodiversity measurement indices. Thus, expressing the average taxonomic distance between two species chosen at random, the relationships between them influence the entropy, unlike other diversity indices.

$$e_{IQ} = \frac{\sum_{i \neq j}^{S} d_{ij}}{S^2} \tag{2.13}$$

Total Taxonomic Distinctness (d_{TT}): represents the average phylogenetic distinctiveness added across all species (gray levels).

$$d_{TT} = \sum i \frac{\sum_{i \neq j}^{S} d_{ij}}{S - 1} \tag{2.14}$$

[1] Taxon is a group of one or more populations of an organism or organisms seen by taxonomists to form a unit.

Accordingly, in image processing, such indices allow for each individual's (pixel) specific location (physical point in space) and the spatial distribution of gray levels. The following indices are based on the distances between pairs of species.

It is worth noting that Equations 2.1 to 2.9 are based on species richness, abundance, and evenness, whereas Equations 2.10 to 2.14 are based on the pairwise distance between pairs of species. All measurements described in Equations 2.1 to 2.14 can be computed from an image – in this section, from each channel of a color image – and they result in scalar values. Normalization is required because dynamic ranges of such scalars are related to species richness, abundance, and their relationship within an image or a region of an image, either directly or inversely. Therefore, these scalars are concatenated and normalized within the interval $[0, 1]$ using min-max mapping to form a d-dimensional feature vector named the BiT descriptor.

The taxonomic indices require a taxonomic tree to compute species' joint dissimilarity (different gray levels) or pairwise distances between species (different gray levels). The topological distance, defined as the number of edges between two species in the Linnaean taxonomic tree, is the entire phylogenetic tree's cumulative branch length. An example of a taxonomic tree and its species distance matrix is shown in Figure 2.2.

Based on the example mentioned above (Figure 2.2), we can derive an instance of the taxonomic tree and its corresponding distance matrix of gray levels (Figure 2.3). We have represented the taxonomic tree as a matrix, where the distance between two gray levels represents the distance between two species. The division of species in the rooted tree shows the phylogenetic relationship between ancestor species. Such a division allows computing indices connecting diversity, richness, and parenthood. Furthermore, a dendrogram can describe the evolutionary relationships between species: the parenthood relationship between gray levels, where the leaves represent the species and the internal nodes represent the common ancestors of the species. This relationship allows establishing an evolutionary connection between the gray levels (species) (Vane-Wright, Humphries & Williams, 1991), which, in this work, relies on the intrinsic properties of the texture present in an image. Thus, the division of an image or a patch

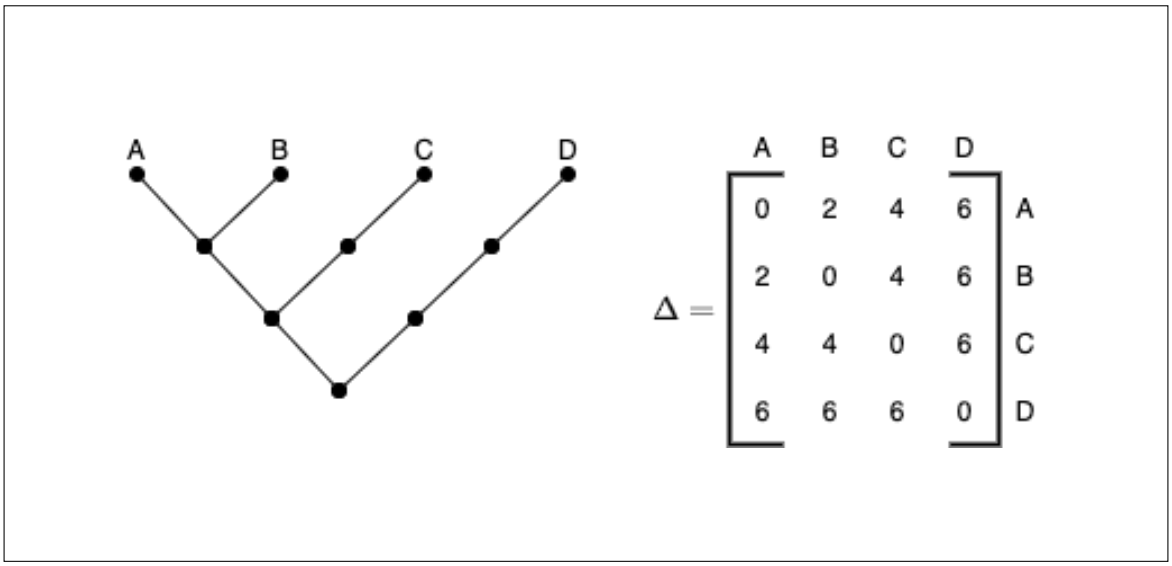

Figure 2.2 Generic example of a four-species taxonomic tree for four species (A, B, C, and D) and its respective distance matrix. This matrix shows how cumulative branch length corresponding to taxonomic distances is calculated
Adapted from Ricotta (2004)

for generating a dendrogram should rely on the parenthood, the similarity between pixels.

Figure 2.3 illustrates the process of division performed in an image or part of it to assemble a phylogeny tree (dendrogram) based on the similarity between gray levels for computing the taxonomic indexes. In this case, some iterations are needed to divide the original region/image until a single gray level remains on each leaf. The division is carried out based on a threshold – considering all the pixels in the entire image – which splits recursively an image into two parts, each containing pixels of gray levels above (right) and below (left) the threshold until the number of species present in each region is 1. Thus, the threshold can be the average of gray levels, where a gray level is a set of pixels with the same intensity value. From the original image, having five gray levels (6, 75, 117, 141, 230), with average $\frac{(6+75+117+141+230)}{5} = 113.8$, in the first iteration (step 1), gray levels 6 and 75 (left) are below the threshold 113.8, whereas gray levels 117, 141, and 230 (right) are above the threshold.

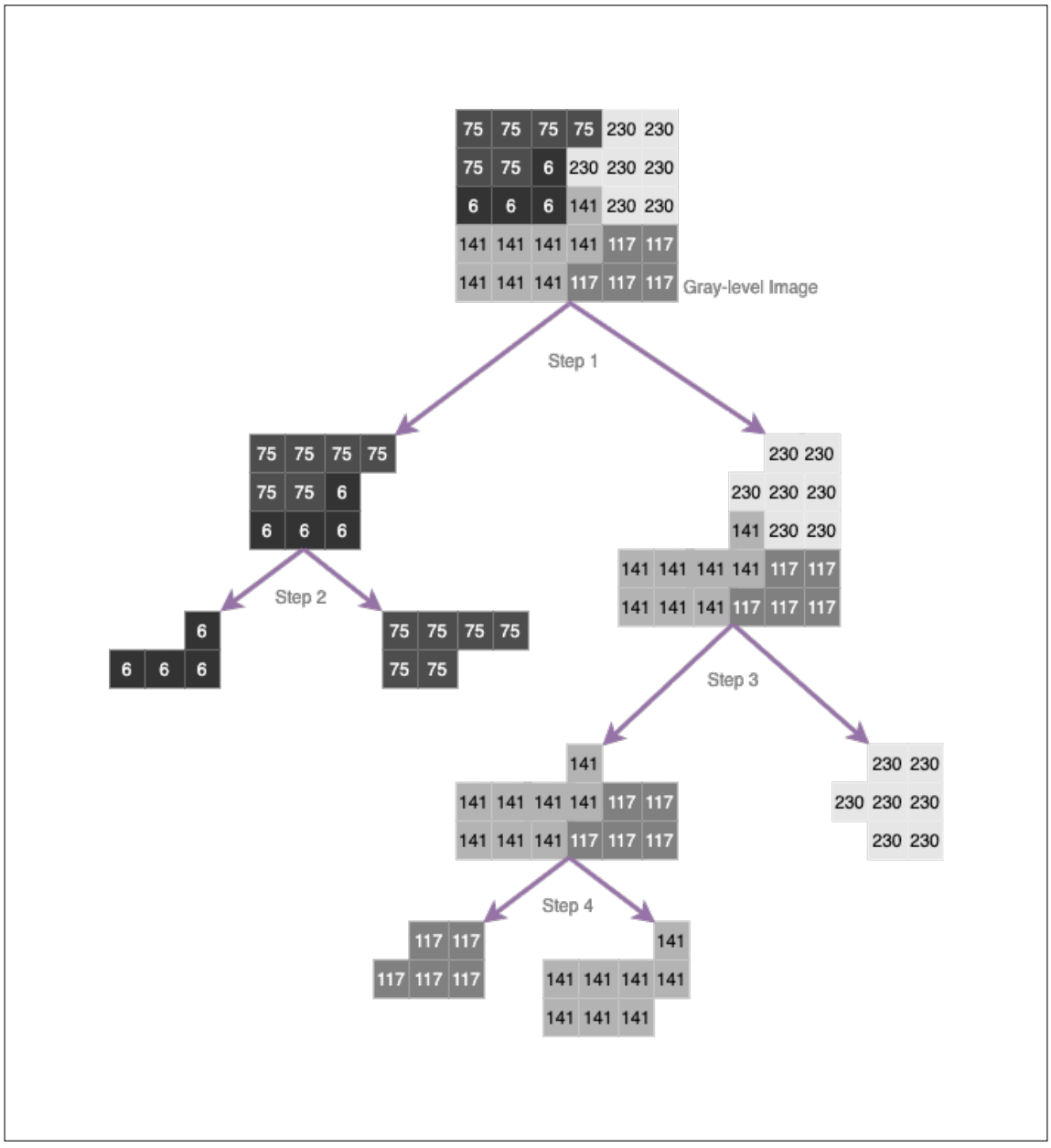

Figure 2.3 Construction of a phylogenetic tree for computing the taxonomic indexes. In each iteration (step), the image is divided based on species (gray levels). The average species value is used as a threshold at each step

The second iteration (step 2) splits the left part resulting from step 1, that is, gray levels 6 (left) and 75 (right), into two parts. Since there are single gray levels in each region resulting from step 2, these regions become leaves.

The third iteration (step 3) separates the right part resulting from step 1 into two parts: pixels of gray levels 141 and 117, which are below the threshold (162.6) go to the left, while pixels of gray-level 230 to the right.

Finally, the fourth iteration (step 4) separates the left part from step 3 into two parts: pixels of gray levels 141 and 117. Figure 2.4 illustrates the rooted tree, the dendrogram, and the respective species (gray levels), as well as their characteristics.

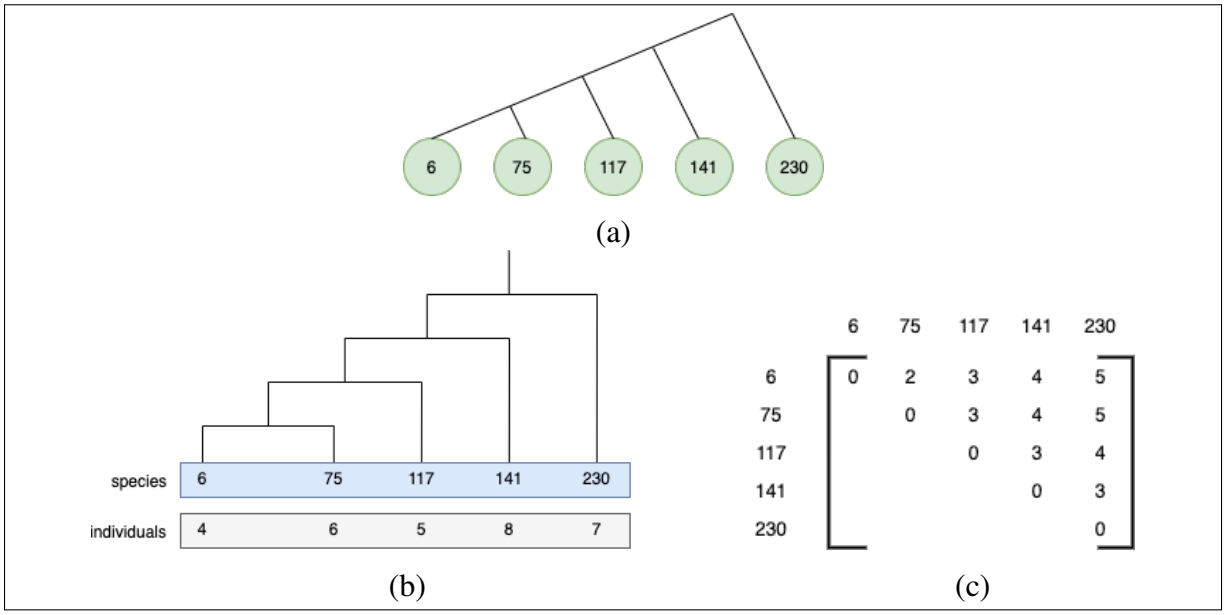

Figure 2.4 Example of (a) rooted tree; (b) a dendrogram; (c) and the respective distance matrix of gray levels computed from the image in Figure 2.3. Note that (a) and (b) are equivalent. The dendrogram allows computing the phylogenetic indexes to infer the phylogenetic relationship between existing gray levels in the original image. Therefrom, the taxonomic indexes are likewise computed

2.1.3 Properties of BiT Descriptors

For many applications, a texture descriptor should have essential properties such as invariance to rotation, translation, and scale. Furthermore, the descriptor should be easy to calculate. The diversity indices based on species richness measure properties directly related to species, such as their relative abundance and evenness. These measurements are invariant to in-plane

rotations and scale (because the true essence of the pattern is invariance). The fundamental idea of diversity indices is to quantify biological variability, which, in turn, can be used to compare biological entities composed of direct components, whether space or time (Sohier, 2019). Biodiversity can be expressed or monitored at different scales and spaces. It is assumed that all species are equal, meaning that richness measurement makes no distinctions among species and treats the exceptionally abundant species in the same way as scarce species. All individuals are equal, meaning there is no distinction between the largest and the minor individual (Magurran, 2004a).

In our abstract model, these assumptions may be expressed as pixels of any gray level are equal. Therefore, the richness measurement makes no distinctions among gray levels and treats pixels that are exceptionally abundant in the same way as significantly less represented pixels. In other words, pixels of all gray levels present in an image are considered for further calculation, regardless of how non-representative some are; and all pixel values are equal. There is no distinction between the largest and the smallest pixel value.

In ecology, a pattern is subject to how form remains invariant to changes in measurement. Some patterns retain the same shape after uniformly stretching or shrinking the measurement scale. The rotational invariance in the ecological pattern has been stated by Frank & Bascompte (2019), being the most general way to understand commonly observed patterns. From there, species abundance distributions provide a transcendent example in which the maximum entropy and neutral models can succeed in some cases because they derive from invariance principles. Likewise, as presented by Daly, Baetens & De Baets (2018), diversity is invariant to the species abundance vector's permutation. Rousseau, Van Hecke, NIjssen & Bogaert (1999b) emphasizes that there is a one-to-one correspondence between abundance vectors and Lorenz curves. Consequently, abundance vectors can be partially ordered according to the Lorenz order, which is permutation-invariant (rotation) and scale-invariant.

Therefore, the BiT descriptor combines statistical and structural approaches and takes advantage of ecological patterns' invariance characteristics to permutation, rotation, and scale by combining species richness, abundance, evenness, and taxonomic indices.

2.1.4 BiT and other Texture Descriptors

The BiT descriptor shares some characteristics of both GLCM (Haralick *et al.*, 1973) and LBP (Pietikäinen *et al.*, 2011) descriptors. The BiT descriptor also characterizes textures based on second-order statistical properties, which involves comparing pixels and determining how a pixel at a specific location relates statistically to pixels at different locations.

In ecology, taxonomic indices approximate second-order statistics at the species level. These indices are based on group analysis, thus enabling a behavioral exploration of the neighborhood of regions displaced from a reference location. For example, given a distance measurement between pairs of species (pairs of pixels of different gray levels), a classical approach to solving the phylogeny issue can be finding a tree that predicts the observed set of adjoining distances. Such distances are represented in a matrix that indicates the existing phylogenetic distance, reducing it to a simple table of pairwise distances (Rogers *et al.*, 1999; Vane-Wright *et al.*, 1991).

Furthermore, the BiT descriptor also shares some characteristics of Gabor filters (Fogel & Sagi, 1989). Gabor filters explore different periodicities in an image and attempt to characterize a texture at these different periodicities. This analysis is confined to the adjacent neighborhoods of the individual pixels. These within-neighborhood periodicity properties can be used to recognize texture differences between the different regions. Accordingly, phylogenetic trees combined with diversity indices are used in biology to compare behavioral patterns between species in different areas and within-neighborhood. Besides, diversity indices based on species richness are of an underlying use when defining an all-inclusive behavior of an ecosystem, forming a non-deterministic complex system.

2.1.5 Case Study

This section presents how the proposed bio-inspired texture descriptor can be integrated with image processing and machine learning algorithms for classification tasks. The proposed classification scheme is structured into five stages: image channel splitting, preprocessing, feature extraction, normalization, and training/classification. Figure 2.5 shows an overview of the proposed scheme. Algorithm 2.1 integrates the first three steps, and it receives an RGB image as input and provides a d-dimensional feature vector of BiT descriptors. An implementation of this algorithm is available as a Python module[2].

The five stages are described as follows.

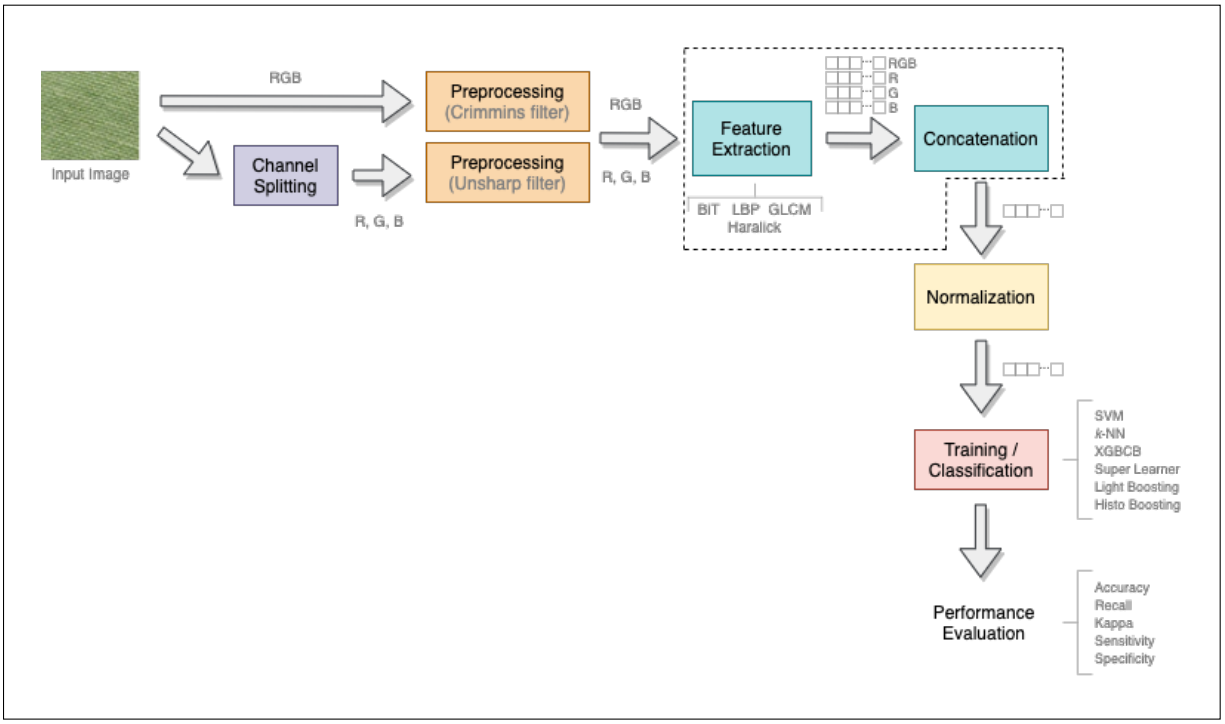

Figure 2.5 An overview of the proposed scheme to evaluate the BiT descriptor and compare it with other texture extractors

[2] https://github.com/stevetmat/BioInspiredFDesc. The Python class **BiT**(*image, b_feat = True, t_feat = True, unsharp_filter = True, crimmins_filter = True, normalization = False*) generates a 56-dimensional feature vector. The library may be found in https://pypi.org/project/Bitdesc/.

2.1.5.1 Channel Splitting

Besides the original RBG image, each image channel (R, G, B) is considered a separate input. Notwithstanding that the texture descriptors presented in Chapiter 1 have shown a discriminative ability to classify texture patterns, their performance on natural and microscopic images may be bounded on the condition that they are applied only to gray-level images. Thus, they do not exploit color information. Mäenpää & Pietikäinen (2004) and Cernadas, Fernández-Delgado, González-Rufino & Carrión (2017) evaluated the impact of color on texture analysis, and they stated that they are distinct phenomena that should be treated individually. Additionally, comparing texture features extracted from gray-level and color images, it has been put forth that color information improves the accuracy under static illumination conditions. However, employing texture and color in parallel is not as powerful as either color or gray-level texture alone. The channel splitting aims to capture the textural information of color images based on the principle that most ecosystems work in a cause-effect relationship. Such a relationship implies that when one resource is added or lost, it affects the entire ecosystem. Some of the most marked temporal/spatial fluctuations in species abundances are linked to this cause-effect relationship (Shimadzu, Dornelas, Henderson & Magurran, 2013). Therefore, we characterize the biodiversity in an image by local descriptors generated from the interaction between a pixel and its neighborhood in each channel (R, G, B) and the RGB image.

The conversion from RGB to grayscale was carried out using:

$$\text{RGB[A] to Gray:} \quad Y \leftarrow 0.299 \cdot R + 0.587 \cdot G + 0.114 \cdot B \tag{2.15}$$

2.1.5.2 Preprocessing

It consists of an unsharp masking filter to highlight image characteristics and a Crimmins filter[3] to remove speckles (Crimmins, 1985). The unsharp[4] filter (Polesel, Ramponi & Mathews, 2000) is applied to each image channel, and the Crimmins filter is applied to the RGB image to improve their quality for the feature extraction step.

Unsharp masking, a method well-known to photographers, is used to modify the relative high-pass content of an image by deleting a version of the image that has been softened (lowpass filtered). The latter can be accomplished optically by developing a blurry image onto a negative film and then using this film as a mask during a second development stage. When processing digital picture data, it is preferable to maintain the image's local mean. Contrary to its name, an "unsharp mask" is used to sharpen an image. In the post-processing of the vast majority of digital photos, sharpening is essential for highlighting texture and detail. Unsharp masks are the most popular type of sharpening and may be applied by practically all image editing programs. An unsharp mask cannot provide extra details, but it can considerably improve the appearance of existing details by enhancing small-scale acuity. The sharpening technique employs a slightly blurred version of the original image, which is removed from the original to create the unsharp mask (effectively a high-pass filter) to detect the presence of edges. Using this mask, contrast is deliberately raised along these edges, resulting in a sharper final image. It turns out that a blurry mask employs a trick performed by the human visual system. At the borders of sharp transitions, the human eye perceives what is known as "Mach bands," named after their discovery in the 1860s by physicist Ernst Mach. These boost our ability to distinguish edge details.

On the other hand, the Crimmins speckle removal algorithm works by putting an image through a speckle and removing the filter that uses complimentary hulling to reduce an image's speckle index. The algorithm uses a non-linear noise reduction technique that compares each pixel intensity to its eight nearest neighbors. Based on the relative values, it adds or decrements the

[3] More details about the implementation: https://homepages.inf.ed.ac.uk/rbf/HIPR2/crimmins.htm

[4] Parameters: radius=5, amount=2

pixel value to make it more representative of its surroundings. Crimmins' approach for noisy pixel change (and detection) is more sophisticated than the non-linear median filter's ranking procedure. In a search for intensity spikes, the "middle" pixel value in each neighboring window is compared with each group of neighbors (N-S, E-W, NW-SE, NE-SW).

2.1.5.3 Feature Extraction and Concatenation

After the preprocessing step, the images undergo feature extraction, which looks for informative and discriminative characteristics. Images are represented by several measurements organized in feature vectors. From each image, we extract biodiversity measurements (Equations 2.1 to 2.7) and taxonomic indices (Equations 2.8 to 2.14), which are concatenated into a single vector. This process is repeated for all the images of a dataset.

2.1.5.4 Normalization

Before the training step, we first split the feature vectors into training and test sets and perform feature normalization over the training data, where values are normalized to the range $[0, 1]$ using the min-max normalization. Afterward, we perform normalization on testing instances using the minimum and maximum values of the training variables. In the case of k-fold cross-validation, we used the same procedure by splitting the feature vectors into k folds and computing the min-max pairs in the merged training folds. The training data min-max pairs are then used to normalize both the training and the test fold. This procedure is repeated for each new training/test fold during the cross-validation procedure.

2.1.5.5 Training/Classification

The final step of the proposed scheme consists of using a shallow approach where feature vectors are split into training and test partitions to train different classification algorithms, as detailed in Section 2.1.6. Finally, the learned models are evaluated and the results obtained are presented and discussed in Section 2.1.7.

Algorithm 2.1 Feature_Extraction_Procedure

Description : Compute BiT descriptor
Input : A RGB image $\mathbf{I}_{m \times n \times 3}$
Output : A d-dimensional feature vector $\mathbf{x}$

1 Separate the RGB image $\mathbf{I}$ in channels $\mathbf{I}^{\mathrm{R}} = \mathbf{I}[1 \ldots n, 1 \ldots m, 1]$, $\mathbf{I}^{\mathrm{G}} = \mathbf{I}[1 \ldots n, 1 \ldots m, 2]$, $\mathbf{I}^{\mathrm{B}} = \mathbf{I}[1 \ldots n, 1 \ldots m, 3]$;
2 Convert $\mathbf{I}$, $\mathbf{I}^{\mathrm{R}}$, $\mathbf{I}^{\mathrm{G}}$, and $\mathbf{I}^{\mathrm{B}}$ to gray-level images $\mathbf{I}^{\mathrm{g}}$, $\mathbf{I}^{\mathrm{Rg}}$, $\mathbf{I}^{\mathrm{Gg}}$, and $\mathbf{I}^{\mathrm{Bg}}$;
3 Apply unsharp filter to $\mathbf{I}^{\mathrm{Rg}}$, $\mathbf{I}^{\mathrm{Gg}}$, and $\mathbf{I}^{\mathrm{Bg}}$;
4 Apply Crimmins filter to $\mathbf{I}^{\mathrm{g}}$;
5 Compute biodiversity measurements (Equations 2.1-2.7) and taxonomic indices (Equations 2.8-2.14) for $\mathbf{I}^{\mathrm{Rg}}$, $\mathbf{I}^{\mathrm{Gg}}$, $\mathbf{I}^{\mathrm{Bg}}$, and $\mathbf{I}^{\mathrm{g}}$;
6 Concatenate the computed measures and indices into a single vector $\mathbf{x}$;
7 Return $\mathbf{x}$;

2.1.6 Experimental Protocol

This section presents the datasets used to assess the performance of the proposed BiT descriptor, which includes natural texture images and histopathological images (HIs), and the experimental protocol to evaluate the properties of the BiT descriptor and its performance on classification tasks. We compare the BiT descriptor's performance with classical texture descriptors such as LBP, GLCM, and Haralick. It is worthy of mention that our contribution relies on the combination of biodiversity measurements and taxonomic indices to build a discriminative descriptor capable of efficiently classifying textures.

2.1.6.1 Texture Datasets

We use three texture datasets that have already been employed for evaluating texture descriptors, such as LBP, GLCM, and Haralick (Simon & Uma, 2018a).

- The **Salzburg** dataset[5] contains 476 color texture images of resolution 128×128, captured around Salzburg in Austria. These images belong to 10 different classes, and 70% of the images are used for training and validating the classification algorithms, while the remaining

[5] http://www.wavelab.at/sources/STex/

30% are used for testing. Figure 2.6(a) shows some samples from the Salzburg texture dataset.

- The **Outex_TC_00010_c** dataset[6] has a training set consisting of 20 non–rotated color images of each of the 24 classes (480 in total) of illuminant "inca", color counterpart of the original Outex_TC_00010 dataset. The test set consists of 3,840 color images of eight orientations (5, 10, 15, 30, 45, 60, 75, and 90 degrees). Each image has the resolution of 128×128 pixels. Figure 2.6(b) shows some samples from the training set of the Outex dataset.
- The **KTH-TIPS** dataset[7] contains a collection of 810 color texture images of 200×200 pixels of resolution. The images were captured at nine scales, under three different illumination directions and three different poses, with 81 images per class. Seventy percent of images are used for training, while the remaining 30% are used for testing. Figure 2.6(c) shows some samples from the KTH-TIPS dataset.

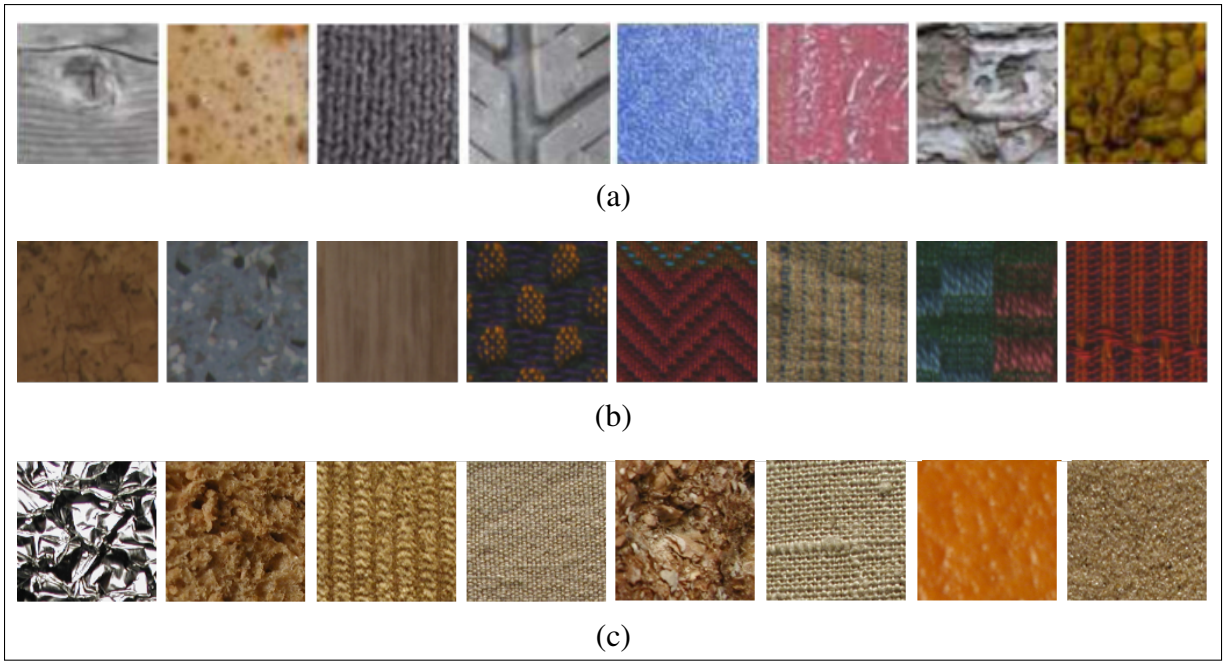

Figure 2.6 Samples from the texture datasets: (a) Salzburg, (b) Outex_TC_00010_c, and (c) KTH-TIPS

[6] http://lagis-vi.univ-lille1.fr/datasets/outex.html

[7] https://www.csc.kth.se/cvap/databases/kth-tips/

2.1.6.2 Histopathological Image (HI) Datasets

HIs are more challenging than pure texture images since HIs usually have other structures, such as nuclei (shape) and tissue variations (colors) within the same class.

- The CRC dataset (Kather, Weis, Bianconi, Melchers, Schad, Gaiser, Marx & Zöllner, 2016a) encompasses colorectal cancer histopathology images of dimension 5,000×5,000 pixels cropped into 150×150 patches and labeled according to the structure they contain. Eight types of structures are labeled: tumor (T), stroma (ST), complex stroma (C), immune or lymphoid cells (L), debris (D), mucosa (M), adipose (AD), and background or empty (E). Each structure detailed in the CRC dataset has a specific textural characteristic. For example, few shape characteristics are found in cell nuclei formation, which has a rounded shape but different coloring due to hematoxylin. The total number of images is 625 per structure type, resulting in 5,000 images. Figure 2.7 shows samples of each class from the CRC dataset. The experiments were performed with stratified 5-fold and 10-fold cross-validation.

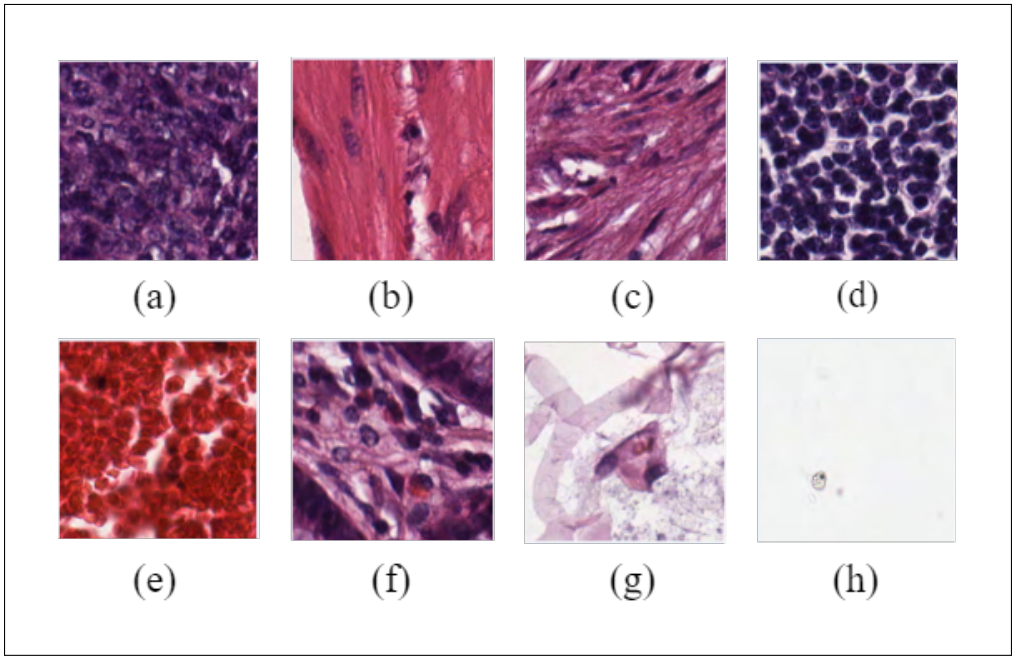

Figure 2.7 Samples of the CRC dataset: (a) tumor, (b) stroma, (c) complex, (d) lympho, (e) debris, (f) mucosa, (g) adipose, (h) empty

- The **BreakHis** dataset (Spanhol, Oliveira, Petitjean & Heutte, 2016a) comprises 7,909 microscopic images of breast tumor tissue collected from 82 patients using different magnification factors (40×, 100×, 200×, and 400×). The breast tissues extracted from biopsy usually have some basic structures, such as glands, ducts, and supporting tissue. For example, there will be a difference in textures between an area with a malignant tumor ductal carcinoma and a healthy area. There will be a significant presence of nuclei in the region with carcinoma, identified by the purple color of hematoxylin's reaction with its proteins. The nuclei and many cells in a reduced region make the apparent texture noisier. In an area without carcinoma, the epithelial tissue is thin and delimits two regions, lumen and stroma, with different textural characteristics due to the excess of epithelial cells. The lumen generally presents itself as a homogeneous and whitish region. Due to its reaction to eosin, the stroma shows a pink and homogeneous color, with little noise. At this point, a texture descriptor can assist in detecting carcinomas by characterizing a given texture. Nevertheless, the evaluation of types of malignant tumors, that is, differentiation between types of carcinoma on a dataset such as BreaKHis, would need to detect shape to differentiate the papillae from a disorderly cluster of cells. The BreakHis dataset contains 2,480 benign and 5,429 malignant samples (700×460 pixels, 3-channel RGB, 8-bit depth in each channel, PNG format). We used hold-outs with repetition, where 70% of samples are used for training, and 30% of samples are used for testing. Figure 2.8 shows samples from each class of the BreakHis dataset.

2.1.6.3 Description of Experiments

We have carried out three types of experiments to evaluate the proposed BiT descriptor:

- Experiments on texture images in which the accuracy of classification algorithms trained using the BiT descriptor are computed for a comparative analysis with traditional texture descriptors.
- Experiments on texture images to evaluate invariance of the BiT descriptor to rotation, scale, and intensity.

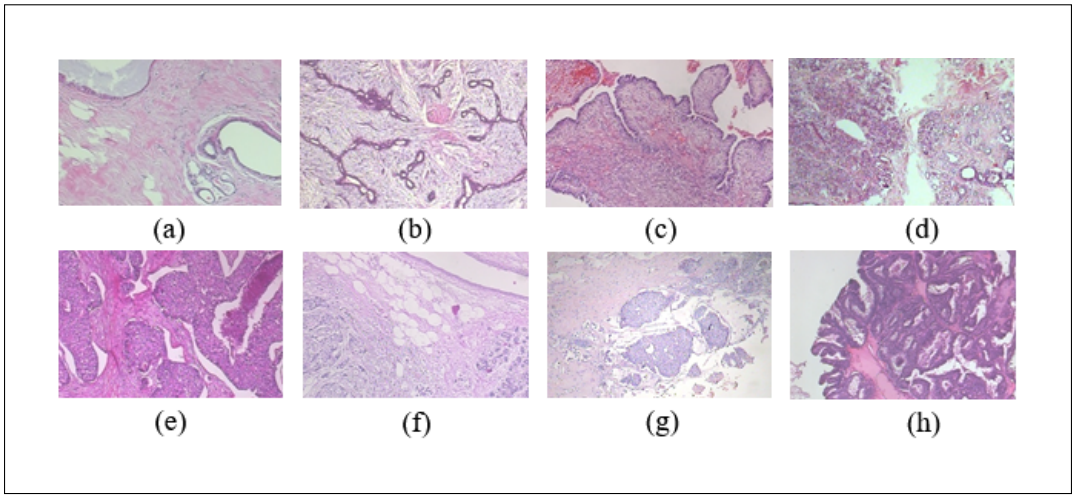

Figure 2.8 Example of HIs: (a) Adenosis, (b) Fibroadenoma, (c) Phyllodes, (d) Tabular adenomaa, (e) Ductal carcinoma, (f) Lobular carcinoma, (g) Mucinous carcinoma, (h) Papillary carcinoma, where (a) to (d) are benign and (e) to (f) are malignant tumors

- Experiments on HIs in which measures frequently used in medical imaging such as sensitivity, specificity, and Kappa scores are computed.

The BiT descriptor is evaluated by the accuracy achieved by different classification algorithms on three texture datasets. The same classification algorithms are trained with other texture descriptors, and their performance is compared with the performance achieved with BiT. For a fair comparison with other texture descriptors, we use the same approach described in Section 2.1.5 for all texture descriptors, including the feature extraction procedure described in Algorithm 2.1 as well as the preprocessing and normalization steps. In addition, we have used SVM and k-NN and four ensemble learning algorithms: decision tree-based ensemble algorithm that uses a gradient boosting framework (XGBCB), a histogram-based algorithm for building gradient boosting ensembles of decision trees (HistoB), light gradient boosting decision trees (LightB), and super learner (SuperL) (van der Laan, Polley & Hubbard, 2007), which involves the selection of different base classifiers and the evaluation of their performances using a resampling technique. SuperL applies a stacked generalization through out-of-fold predictions during k-fold cross-validation. The base classifiers used in SuperL are k-NN, decision trees,

and ensembles of decision trees such as AdaBoost, bagging, extra trees, and random forest. Furthermore, the k-NN was tuned with k values between 1 and 21. The SVM was tuned using a grid search with the hyperparameter c between 0.1 and 2 and linear, polynomial, Gaussian, and sigmoid kernels.

The invariance properties of the proposed BiT descriptor are evaluated on different transformations applied to texture images. We compute the BiT descriptor for each image and compare them to those computed from the transformed images. In this case, feature values should not change with the transformations. Additionally, we have also evaluated the invariance of the BiT descriptor to monotonic intensity transformations.

The BiT descriptor is also evaluated on two HI datasets. Only the classification algorithm that achieved the best performance with the BiT descriptor in the previous experiment is retained. Its performance is compared with the state-of-the-art of these datasets, which includes approaches based on CNNs. The experiments are performed using stratified k-fold cross-validation because the related works employed such a strategy.

2.1.7 Experimental Results and Discussion

This section presents the experimental sets to evaluate the proposed method. We have conducted experiments on texture datasets (of natural images) and HI datasets. Along the same line, we also presented the invariance properties of the BiT descriptor. Moreover, we provided the details of the environment where all the experiments were carried out.

2.1.7.1 Experiments with Texture Datasets

Table 2.1 shows the accuracy achieved by monolithic classifiers and ensemble methods trained with four texture descriptors on Salzburg, Outex, and KTH-TIPS datasets. The proposed BiT descriptor achieved the best accuracy for most of the classification algorithms on the Salzburg dataset, and the best result was achieved with BiT+SuperL (94.23%), which outperformed all other texture descriptors. The difference in accuracy achieved by BiT and the second and

the third-best texture descriptors (Haralick+SVM and GLCM+k-NN) is nearly 6% and 19%, respectively. The BiT descriptor also provided the best accuracy on the Outex dataset, and BiT+SVM (99.88%) achieved the best result. The difference in accuracy achieved by BiT+SVM and the second and the third-best texture descriptors (Haralick+k-NN and GLCM+LightB) is nearly 3% and 5%, respectively. On the KTH-TIPS dataset, the BiT descriptor provided the best accuracy for all the classification algorithms. The best result was achieved with BiT+SVM (97.87%). The difference in accuracy achieved by BiT and the second and the third-best texture descriptor (Haralick+SVM and GLCM+SuperL) is nearly 3% and 11%. Additionally, the McNemar test[8] has shown a different proportion of errors on the test set for the three datasets. Therefore, there is a statistically significant difference between the best results of the BiT descriptor and the three other feature descriptors.

We have also evaluated the importance of the preprocessing step in the final accuracy. Overall, there is no clear advantage of preprocessing texture images because, depending on the dataset and classification algorithm, the preprocessing may improve or harm the accuracy of the feature extraction methods. However, it is worth mentioning that the preprocessing played an important role for GLCM and Haralick descriptors for specific classifiers (mainly k-NN), where significant differences (greater than 10%) were observed between the accuracy achieved on the original dataset (lower) and the preprocessed one (higher). In general, the proposed method does not experience a significant variation in terms of accuracy. For instance, the accuracy of BiT+SVM without preprocessing decreases slightly on the Salzburg (93.78%) and KTH-TIPS (97.57%) datasets, and it increases slightly on the Outex dataset (99.99%).

A direct comparison of the results presented in Table 2.1 with other works that employed the same datasets may not be reasonable due to differences in the experimental protocols. For example, the results reported on the Salzburg dataset omit the subclasses used in the experiments and which samples made up the test set. Several works have also used the Outex dataset for texture classification. For instance, Mehta & Egiazarian (2016) presented an approach based on a rotation-invariant LBP, which achieved an accuracy of 96.26% with a k-NN. Nonetheless, it

[8] Significance level of 95%.

has the downside of not considering color information and global features. Du, Yan & Ma (2016) presented a rotation-invariant, impulse noise resistant, and illumination-invariant approach based on a local spiking pattern, which achieved an accuracy of 86.12% with a neural network. Notwithstanding, it is not extended for color textures, and it requires many input parameters.

The KTH-TIPS dataset has also been used to evaluate approaches for texture classification. Hazgui, Ghazouani & Barhoumi (2021) presented an approach based on genetic programming and fusion of HOG and LBP features. Such an approach achieved an accuracy of 91.20% with a k-NN. Nevertheless, it does not consider color information and global features. Nguyen, Vu & Manzanera (2016) presented statistical binary patterns, which are rotational and noise invariant. Such an approach reached an accuracy of 97.73%, which is slightly lower than the accuracy achieved by BiT+SVM. In addition to being resolution sensitive, this method presents a high computational complexity. Qi *et al.* (2013) studied the relative variance of texture patterns between different color channels using LBP and Shannon entropy to encode the cross-channel texture correlation. They proposed a multi-scale cross-channel LBP (CCLBP), which is rotation-invariant. The CCLBP computes the LBP descriptors in each channel and three scales and computes co-occurrence statistics before concatenating the extracted features. Such an approach achieved an accuracy of 99.01% for three scales with an SVM, which is 1.14% higher than the accuracy achieved by BiT+SVM. Notwithstanding, this method is not invariant to scale.

In addition to the shallow methods, we have also carried out experiments on the three texture datasets with a tiny T-CNN of 11,900 trainable parameters with and without data augmentation (1×, 2×, 4×, and 6×) (de Matos *et al.*, 2019)(Ataky *et al.*, 2020). The T-CNN achieved the best accuracy of 61.06%, 70.6%, and 70.22% for Salzburg, Outex, and KTH-TIPS datasets, respectively. However, these results are far below the accuracy achieved by BiT+SVM and BiT+SuperL reported in Table 2.1.

Finally, we have carried out an empirical evaluation to assess the computational time required by the BiT descriptor for feature extraction and compare it with the other three texture descriptors

on the three texture datasets. The average time[9] per image, including preprocessing, is as follows: (i) KTH-TIPS dataset: 27.2 msec, 79.8 msec, 106.2 msec, and 1.39 sec for GLCM, Haralick, LBP, and BiT, respectively; (ii) Outex dataset: 24.4 msec, 42.3 msec, 51.1 msec, and 461.8 msec for GLCM, Haralick, LBP and, BiT, respectively; (iii) Salzburg dataset: 12.6 msec, 18.7 msec, 15.8 msec, and 881 msec for GLCM, Haralick, LBP, and BiT, respectively. The entire methodology was developed using Microsoft Windows 10 operating system, with the Python programming language running on an Intel Core i7-8850H CPU @ 2.60 GHz, 64-bit operating system, x64-based processor, and 32 GB of RAM memory.

Table 2.1 Average accuracy (%) on the test set of Salzburg, Outex, and KTH-TIPS datasets. The overall best result for each dataset is in boldface. The best result for each texture descriptor is marked with *

Dataset	Texture Descriptors	Classification Algorithms					
		XGBCB	HistoB	LightB	SuperL	k-NN	SVM
Salzburg	LBP	56.13±0.024	56.89±0.029	57.21±0.021	64.12*	32.23	62.12
	GLCM	70.74±0.022	68.33±0.018	71.77±0.016	75.04	75.38*	63.68
	Haralick	81.75±0.015	82.27±0.017	84.27±0.017	86.88	82.54	87.99*
	BiT	88.65±0.015	88.20±0.014	90.17±0.013	**94.23**	88.65	92.33
Outex	LBP	55.30±0.011	57.80±0.013	57.35±0.014	83.21*	48.20	82.41
	GLCM	94.37±0.012	94.13±0.006	95.52±0.008*	94.29	94.37	94.06
	Haralick	96.15±0.003	96.69±0.003	96.53±0.004	95.5	96.92*	96.71
	BiT	99.34±0.006	99.68±0.005	98.53±0.006	99.53	99.83	**99.88**
KTH-TIPS	LBP	57.17±0.021	59.51±0.031	57.18±0.019	64.83*	58.26	61.78
	GLCM	83.53±0.028	83.12±0.017	86.00±0.022	86.83*	74.89	79.83
	Haralick	90.12±0.019	88.83±0.019	90.94±0.020	93.00	89.71	94.89*
	BiT	92.18±0.024	92.59±0.022	94.65±0.024	95.41	95.49	**97.87**

k = 3, 5, and 1 for k-NN on Salzburg, Outex, KTH-TIPS datasets, respectively.
Linear kernel and c = 2.0, 1.3, and 1.7 for SVM on Salzburg, Outex, KTH-TIPS datasets, respectively

2.1.7.2 Invariance of the BiT Descriptor

Figure 2.9 illustrates different transformations applied to texture images (first row) and HIs (second row). We have computed some BiT descriptors for each transformed image, and the non-normalized feature values are presented in Tables 2.2 and 2.4.

[9] Average time in seconds per image for running a Python implementation of Algorithm 1.

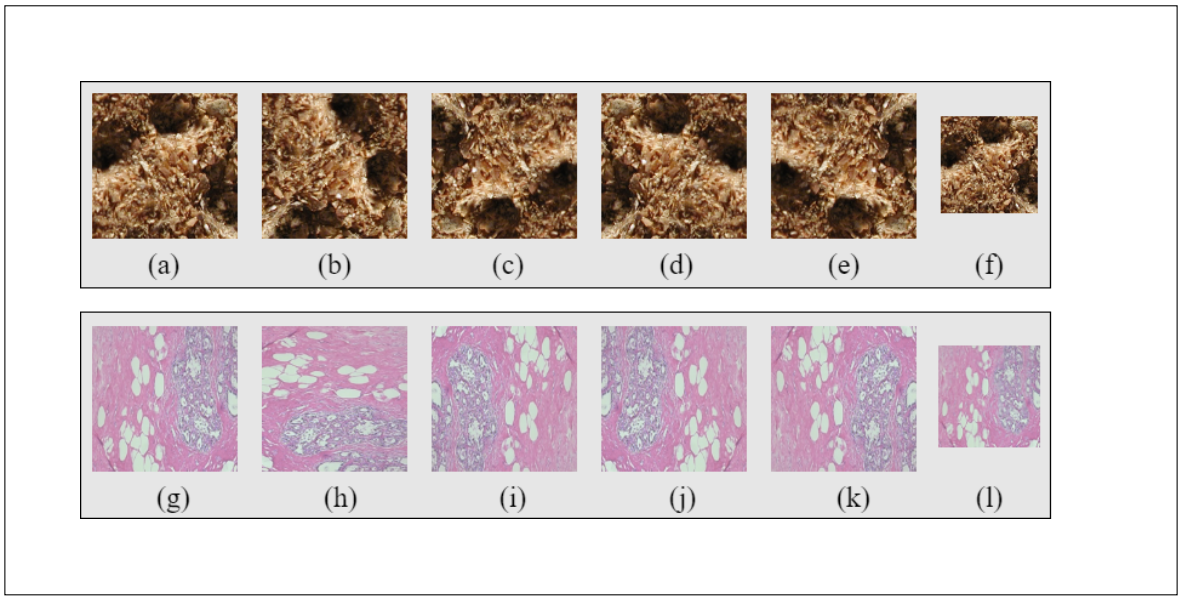

Figure 2.9 Example of texture images: (a) original image, (b) rotation 90°, (c) rotation 180°, (d) horizontal reflection, (e) vertical reflection, (f) rescaled 50%. Example of histopathologic images: (g) original image, (h) rotation 90°, (i) rotation 180°, (j) horizontal reflection, (k) vertical reflection, (l) rescaled 50%

Table 2.2 Non-normalized feature values computed from different image transformations applied to a texture image (Figure 2.9(a))

BiT Features	Original	Transformations				
		Rotation		Reflection		Rescaling
		90^o	180^o	Horizontal	Vertical	50%
d_{Mg}	2636.49	2636.49	2636.49	2636.49	2636.49	725.45
e_M	0.00055	0.00055	0.00055	0.00055	0.00055	0.00109
d_{Mn}	20.3073	20.3073	20.3073	20.3073	20.3073	10.1634
d_{SW}	15.0453	15.0453	15.0453	15.0453	15.0453	14.9963
Δ	101297.7	101297.7	101297.7	101297.7	101297.7	6253.41
Δ^*	2.002325	2.002325	2.002325	2.002325	2.002325	2.003482
e_{IQ}	2.4900637	2.4900637	2.4900637	2.4900637	2.4900637	2.4901419
d_{NN}	4.9999	4.9999	4.9999	4.9999	4.9999	4.9999

Because the computed values for features presented in Table 2.2 (the same for Table 2.4) have considerable variances in their absolute values (significant characters and position after the dot), relying merely on the number does not allow for fair comparisons between them. Table 2.3 shows the relative error among them, along with the values of each variable after they have been rescaled.

Table 2.3 Rescaled feature values computed from Table 2.2

Features	Absolute Values		Error	Relative Error
	Original	Rescaled		
d_{Mg}	2636.4900000	723.4500000	1913.0400000	0.7256011
e_M	0.0005500	0.0010900	0.0005400	0.9818182
d_{Mn}	20.3073000	10.1634000	10.1439000	0.4995199
d_{SW}	15.0453000	14.9963000	0.0490000	0.0032568
Δ	101297.7000000	6253.4100000	95044.2900000	0.9382670
Δ^*	2.0023250	2.0032820	0.0009570	0.0004779
e_{IQ}	2.4900637	2.4901419	0.0000782	0.0000314
d_{NN}	4.9999000	4.9999000	0.0000000	0.0000000

Table 2.4 Non-normalized feature values computed from different image transformations applied to a histopathologic image (Figure 2.9(g))

BiT Features	Original	Transformations				
		Rotation		Reflection		Rescaling
		90^o	180^o	Horizontal	Vertical	50%
d_{Mg}	1975.95	1975.95	1975.95	1975.95	1975.95	548.347
e_M	0.00036	0.00036	0.00036	0.00036	0.00036	0.00072
d_{Mn}	13.2022	13.2022	13.2022	13.2022	13.2022	6.64831
d_{SW}	14.8910	14.8910	14.8910	14.8910	14.8910	14.6985
Δ	214389.7	214389.7	214389.7	214389.7	214389.7	15287.596
Δ^*	2.00673	2.00673	2.00673	2.00673	2.00673	2.00710
e_{IQ}	2.48115	2.48115	2.48115	2.48115	2.48115	2.48099
d_{NN}	4.9998	4.9998	4.9998	4.9998	4.9998	4.9998

The values of BiT descriptors presented in Tables 2.2 and 2.4 show that: (i) all measurements employed are invariant to rotation and reflection as shown in Figures 2.9(a)-(e) and 2.9(g)-(k), since they presented the same values for all texture images or HIs.

This analysis also corroborates the fact that BiT descriptors capture the all-inclusive behaviors of patterns in an image; (ii) Shannon-Wiener diversity index (d_{SW}), taxonomic distinctness (Δ^*), intensive quadratic entropy (e_{IQ}), and the average distance from the nearest neighbor (d_{NN}) are invariant to scale as they provided values of the order of other transformations for each of the images. By way of explanation, from Table 2.3, the latter can be expressed as: (ii) Shannon-Wiener diversity index (d_{SW}), taxonomic distinctness (Δ^*), intensive quadratic entropy (e_{IQ}), and the average distance from the nearest neighbor (d_{NN}) is invariant to scale as they

provided relative error between the original and the rescaled sample less than 0.5% for each image.

On the other hand, the measures based on richness and abundance show some scale dependence. By changing the image scale, we somehow affect the proportion of both factors, which affects the resulting values either directly or inversely. Conjointly, taxonomic indices rely on the parenthood relationship between species. Therefore, they are not affected by the change in scale, as the phylogenetic relationship depends on the intrinsic properties found in the ecosystem (image).

2.1.7.3 Invariance to Intensity Changes

We have also carried out an empirical evaluation to assess the impact of monotonic intensity transformations on the performance of the BiT descriptor. We have applied gamma transformation with values between 0.5 and 3.0 on the test set of Salzburg, Outex, and KTH-TIPS datasets for such an aim. Table 2.5 shows the results for selected combinations of the BiT descriptor and classifiers[10] from Table 2.1.

The difference in accuracy between the original and gamma-transformed images for the Salzburg dataset is above 5%. Notwithstanding, the results are still slightly higher than other texture descriptors, as shown in Table 2.1. On the other hand, the results achieved on the Outex dataset vary less than 1%, which shows some robustness of the BiT descriptor to intensity changes. There is a difference of nearly 3% in terms of accuracy between the original and gamma-transformed images for the KTH-TIPS dataset.

From Table 2.5, the McNemar test has shown similar proportions of errors between the original test set and all transformed test sets of the KTH-TIPS dataset. On the other hand, for the Salzburg dataset, the McNemar test has shown equal proportions of errors with gamma values of 2.0, 2.5, and 3.0, but different proportions of errors with gamma values of 0.5 and 1.5. Finally, for the Outex dataset, the McNemar test has shown similar proportions of errors with gamma values of 0.5, 2.5, and 3.0 and different proportions of errors with gamma values of 1.5 and 2.0. In

[10] The classification models were trained only on the original images.

summary, the proposed BiT descriptor is not invariant to intensity changes for some datasets, even though it achieved promising results for the KTH-TIPS dataset.

Table 2.5 Average accuracy (%) on the three texture datasets with the BiT descriptor applying gamma transformation on the images of the test set

Dataset	Classification Algorithm	Original	Gamma Correction				
			0.5	1.5	2.0	2.5	3.0
Salzburg	SuperL	94.23	92.27	93.19	91.37	91.33	90.88
Outex	SVM	99.88	99.45	99.61	99.45	99.76	99.69
KTH-TIPS	SVM	97.87	97.65	97.23	97.65	97.35	96.41

2.1.7.4 Experiments with HI datasets

Table 2.6 shows the accuracy of monolithic classifiers and ensemble methods trained with BiT descriptors on the CRC dataset. Among all classification algorithms, SuperL provided the best results. We have also computed other important metrics used in medical images for BiT+SuperL. Table 2.7 shows that specificity, sensitivity, and Kappa achieved on the CRC dataset are 94.19%, 94.22%, and 93.53%, respectively. For both cases of Table 2.6, the McNemar test has shown a similar proportion of errors on the test set for the CRC dataset under Super Learner compared with all the classifiers, except the SVM, which presents a different proportion of errors. Therefore, there is no statistically significant difference in the disagreements between the best results of the BiT descriptor on CRC and all other classifiers. However, there is a statistically significant difference in the disagreements between SVM and Super Learner.

Table 2.8 compares the results achieved by BiT+SuperL with the state-of-the-art for the CRC dataset. The proposed descriptor slightly outperforms the accuracy achieved by nearly all other methods. For instance, the difference in accuracy to the second-best method (CNN) is 0.12%, considering an 8-class classification task for those who used 10-fold cross-validation. In contrast, BiT+SuperL reached the second-best accuracy with a slight difference of 0.29% compared with the first-best method (CNN). It is worthy of mention that the success of CNNs relies on the ability to leverage massive labeled datasets to learn high-quality representations. Notwithstanding, data availability for a few fields may be scanty, and therefore CNNs become

prohibitive in several domains. The results achieved by the BiT descriptor on the CRC dataset for HI classification have shown that the proposed descriptor works well on other types of images, which have structures other than textures, with no need for data augmentation.

Table 2.6 Average accuracy (%) of monolithic classifiers and ensemble methods with the BiT descriptor on the CRC dataset

Texture Folds	Classification Algorithms					
	XGBCB	HistoB	LightB	SuperL	k-NN	SVM
5	90.80±0.009	89.60±0.010	90.20±0.013	**92.31**	83.50	90.40
10	90.50±0.010	90.71±0.008	90.23±0.013	**92.52**	83.70	90.98

Table 2.7 More evaluation performance for BiT+SuperL on the CRC dataset

Specificity	Sensitivity	Kappa
94.43	94.47	93.87

Table 2.8 Average accuracy (%) of shallow and deep approaches on the CRC dataset

Reference	Approach	10-fold	5-fold
Ribeiro, Neves, do Nascimento, Roberto, Martins & Tosta (2019)	Shallow	97.60*	–
Kather, Weis, Bianconi, Melchers, Schad, Gaiser, Marx & Zöllner (2016b)	Shallow	87.40	–
Sarkar & Acton (2017)	Shallow	73.60	–
BiT+SuperL	Shallow	**92.52**	92.31
Wang, Shi, Zhang & Ying (2017)	CNN	–	**92.60**
Pham (2017)	CNN	–	84.00
Raczkowski, Możejko, Zambonelli & Szczurek (2019)	CNN	92.40	92.20

*Used 2-classes classification instead.

Table 2.9 shows the accuracy of monolithic classifiers and ensemble methods trained with the BiT descriptor on the BreakHis dataset. The SVM classifier achieved the best accuracy for all magnifications, followed by the SuperLearner. Table 2.10 shows specificity, sensitivity, and Kappa achieved by BiT+SVM.

Table 2.11 compares the results achieved by BiT+SVM with the state-of-the-art for the BreakHis dataset. The proposed descriptor achieved a considerable accuracy of 97.29% for 40× magnification, which slightly outperforms the accuracy of shallow and deep methods. The difference in accuracy between the proposed method and the second-best method (CNN) is about 0.29% for

Table 2.9 Average accuracy (%) of classification algorithms with the BiT descriptor on the BreakHis dataset

Classification Algorithms	Magnification			
	40×	100×	200×	400×
XGBCB	94.01	94.03	92.08	91.10
HistoB	93.85	93.82	91.75	90.83
LightB	94.96	94.89	93.68	92.81
SuperL	96.18	95.95	94.63	93.81
SVM	**97.29**	**96.62**	**95.62**	**95.19**

Table 2.10 More evaluation performance for BiT+SVM on the BreakHis dataset

Magnification	Specificity	Sensitivity	Kappa
40×	95.12	94.82	95.42
100×	95.35	94.43	95.21
200×	94.05	94.18	93.22
400×	95.12	95.24	93.72

40× magnification. Notwithstanding, the best CNN method outperforms BiT+SVM for 100×, 200×, and 400× magnification with a difference of 0.88%, 1.58%, and 2.01%, respectively. Moreover, Table 2.11 presents the results achieved by Spanhol *et al.* (2016a), which also used LBP, GLCM, and other texture descriptors with monolithic classifiers and ensemble methods. For instance, the results achieved by BiT+SVM outperform their GLCM approach by 22.59%, 19.82%, 12.22% and 13.49% for 40×, 100×, 200× and 400×, respectively.

Table 2.11 Average accuracy (%) of shallow and deep approaches on the BreakHis dataset. All these works used the same data partition for training and test

Reference	Method	Magnification			
		40×	100×	200×	400×
Spanhol *et al.* (2016a)*	Shallow	75.60	73.00	72.90	71.20
Spanhol *et al.* (2016a)+	Shallow	74.70	76.80	83.40	81.70
Erfankhah, Yazdi, Babaie & Tizhoosh (2019)	Shallow	88.30	88.30	87.10	83.40
BiT+SVM	Shallow	**97.29**	96.62	95.62	95.19
Alom, Yakopcic, Nasrin, Taha & Asari (2019)	CNN	97.00	**97.50**	**97.20**	**97.20**
Han, Wei, Zheng, Yin, Li & Li (2017)	CNN	92.80	93.90	93.70	92.90
Bayramoglu, Kannala & Heikkilä (2016)	CNN	83.00	83.10	84.60	82.10
Spanhol, Oliveira, Petitjean & Heutte (2016b)	CNN	90.00	88.40	84.60	86.10

*: LBP; +: GLCM.

Though CNNs have overcome shallow methods for several classification tasks, their advantages on texture images are not so high. Pre-trained CNN architectures designed for object classification still require fine-tuning some convolutional and fully connected layers on a large amount of data to achieve good performance, including tiny architectures such as T-CNN and T-CNN Inception (Ataky *et al.*, 2020; de Matos *et al.*, 2019), compact architectures such as EfficientNets and MobileNets[11]. Besides, CNNs still have explainability and interpretability issues. In contrast, the amount of data and the computational effort to calculate the BiT descriptor are relatively low. Furthermore, the proposed BiT descriptor is generic. As the experiments over different datasets have shown, it does not require retraining or hyperparameter configuration while providing state-of-the-art performance.

2.1.8 Final Considerations

This section has presented an essential contribution to texture characterization using biodiversity measurements and taxonomic distinctiveness. We have proposed a bio-inspired texture descriptor named BiT, based on abstract modeling of an ecosystem as a gray-level image where image pixels correspond to a community of organisms. We have revisited several biodiversity measurements and taxonomic distinctiveness to compute features based on species richness, species abundance, and taxonomic indices. The combination of species richness, species abundance, and taxonomic indices takes advantage of the invariance characteristics of ecological patterns such as reflection, rotation, and scale.

These bio-inspired features form a robust and invariant texture descriptor that can be used with machine learning algorithms to build powerful classification models. Furthermore, experimental results on texture and HI datasets have shown that the proposed texture descriptor can train different classification algorithms that outperform traditional texture descriptors and achieve very competitive results compared to deep methods. Therefore, the proposed texture descriptor

[11] Number of parameters: T-CNN: 11,900; T-CNN Inception: 1.2M; MobileNetV2: 3.5M; EfficientNetB0: 5.3M

is promising for mainly dealing with texture analysis and characterization problems. The results demonstrate the promising performance of such a bio-inspired texture descriptor presented.

As an extension of the BiT in Section 2.2, we propose integrating a few sets of diversity and evenness measures widely used in ecology, aiming to resemble the completeness of alpha diversity as a generalization of biodiversity analysis.

2.2 E-BiT: Extended bio-inspired texture descriptor for texture analysis and characterization

In ecology, diversity is defined as describing the variety and abundance of species in a specific unit of study. It is a measure often used to describe the complexity of a community (Magurran, 2004b). Evenness measures the relative abundance of the different species in the same area and is a way to measure how the species are evenly distributed in a community. In other words, it refers to the equitability of the taxa frequencies in a community (Wagner, Grunwald, Zerbe, Mikulich-Gilbertson, Robertson, Zemanick & Harris, 2018). Notwithstanding the concept of diversity is rather precise, Wagner *et al.* (2018) presented a few reasons that may restrain its application, and some of them are: (i) the existence of numerous commonly used diversity indices that can yield different results; (ii) partitioning diversity into components, such as richness and evenness, may be useful, but varies depending on the diversity measure; and (iii) the terminology currently in use to describe diversity is complex and confusing. For instance, Shannon and Simpson, both indexes of diversity work differently. The former equally weights richness and evenness, whereas the latter provides more weight to evenness, and such differences in weighting clarify differences often perceived in results from each measure. What if we integrate and combine more diversity indexes? In this section, we state that by incorporating more diversity and evenness indexes, considering the variations in the mathematical properties of each of them, it is possible to build a descriptor even more robust than the existing BiT descriptor.

The main contribution of this section is an extension of the BiT descriptor that integrates a few sets of diversity and evenness measures widely used in ecology to resemble the completeness of

alpha diversity to build a robust representation for texture characterization and classification. More specifically, the contributions are:

- An extended version of the BiT descriptor (henceforth E-BiT) combining species diversity, evenness, richness, and taxonomic indexes to resemble the completeness of alpha diversity as a generalization of biodiversity analysis.
- A descriptor that captures the all-inclusive behavior of texture image patterns (both local and global features).
- The E-BiT descriptor is invariant to scale, translation and permutation; (iv) the E-BiT descriptor is easy to compute and has low computational complexity.
- The E-BiT descriptor is a generic texture descriptor that performs well on different image categories, such as natural textures and medical images.

2.2.1 Methodology

The approach proposed in this work is inspired by the previous one presented in Section 2.1, where we relied on the biodiversity measures and taxonomic indices to build a robust and generic texture descriptor. Some important properties of such a descriptor are: (i) the advantage of ecological patterns' invariance characteristics to build a permutation, rotation, and translation invariant descriptor; (ii) its ability to capture the all-inclusive behavior of texture image patterns regardless of the latter forming a non-deterministic complex system, which allows the characterization of the intrinsic properties of the whole image.

The BiT descriptor consists of fourteen ecological diversity indices comprising biodiversity measures and taxonomic indices. Nonetheless, in ecology, numerous commonly used diversity indices can yield different results, consequently bringing more information related to diversity characterization. What if we integrate and combine more diversity indexes? We extend the BiT descriptor by integrating more diversity and evenness indices as a baseline. With this, we investigate new diversity and evenness indices not used in our previous Section by considering the variations in the mathematical properties of each of them.

The combination of fourteen indices used to build the BiT descriptor and ten other indices give rise to the Extended BiT (E-BiT) descriptor, which provides a representation that leads to classification accuracy superior to the existing BiT descriptor.

2.2.2 Extended Bio-Inspired Texture (E-BiT) Descriptor

Biodiversity is measured as a combination of richness and evenness across species (Rousseau *et al.*, 1999a). The former component is also referred to as species richness. It stands for the number of groups of functionally related individuals, and the latter denotes the proportions of species or functional groups present in an ecosystem or community. Besides these components, another type of index is taxonomic, which considers the taxonomic relationships between different organisms in an ecosystem. Moreover, taxonomic diversity reflects the average taxonomic distance between any two organisms, randomly chosen from a sample. Such a distance represents the length of the path connecting these two organisms along the branches of a phylogenetic tree (Sohier, 2019).

The multitude of indices proposed in biodiversity is one of the reasons why it is a bit difficult to quantify it. Ecological diversity indexes are mathematical measures of species diversity and rely on richness (number of species) and abundance (number of individuals per species). The majority of indices used in ecology usually measure proportional abundances. Nonetheless, two main types of indices are dominance indices (e.g., the Simpson index) and information statistic indices (e.g., the Shannon-Weiner index).

Multiple diversity indices of richness can be measured. However, there is no clear consensus about which indices are more appropriate and informative (Morris, Caruso, Buscot, Fischer, Hancock, Maier, Meiners, Müller, Obermaier, Prati et al., 2014). Notwithstanding, richness is the most straightforward metric and most commonly applied to represent diversity. Whereas evenness describes the degree to which individuals are divided among species, the individuals' distribution pattern.

Lower values indicate that one or a few species dominate. Higher values suggest that relatively comparable numbers of individuals belong to each species and, consequently, a more useful measure of evenness in multiple contexts.

Therefore, the question is, "which one to choose to quantify the diversity?" Both richness and evenness represent two (among many) of biodiversity's facets. Accordingly, no single number can incorporate them both without information loss. However, taken jointly, they provide a robust suite of methods to explore and analyze the structure of a community. Likewise, Morris *et al.* (2014) stated that in simple analyses, standard diversity indices might appear interchangeable. Nonetheless, by considering complex interactions, the choice of the index can significantly alter the interpretation of the outcomes. Thus, data mining should be avoided to identify the index yielding the most effective results. However, the author suggested that simultaneously considering analyses using multiple indices can provide greater insight into the interactions in a system. In this section, we integrate ten more diversity and evenness indices to gain greater insight into the interactions in a system (a textural image). Such diversity measures (Magurran, 2004b; SDR-IV, 2020) are described underneath. It is worthy of note that the terminology and the notation used in this section are similar to those of Section 2.1.

2.2.2.1 Diversity Indices

The **Brillouin index** (d_{HB}) is defined as:

$$d_{HB} = \frac{\ln N! - \sum_{i=1}^{S} \ln n_i!}{N} \tag{2.16}$$

where N is defined as the total number of pixels in the sample, S is the number of gray levels, and n_i is defined as the number of pixels in the i-th gray level.

The **Strong's dominance index** (d_{Dw}) is defined as:

$$d_{Dw} = \max_i [(\frac{b_i}{N}) - \frac{i}{S}] \tag{2.17}$$

where b_i is the sequential cumulative totaling of the i-th distinct gray levels' values ranked from largest to smallest. The expression in brackets is computed for all gray levels, and max_i denotes the maximum value in brackets for any gray level.

The **Simpson's index** (d_C) and the **Enspie index** (d_{ENS}) are defined as:

$$d_C = 1 - \sum_{i=1}^{S} p_i^2 \tag{2.18}$$

$$d_{ENS} = \frac{1}{\sum_{i=1}^{S} p_i^2} \tag{2.19}$$

where p_i is the proportion of the community represented by gray level i.

The **McIntosh dominance diversity index** (d_{McInt}) (McIntosh, 1967) is defined as:

$$d_{McInt} = \frac{N - U}{N - \sqrt{N}} \tag{2.20}$$

where N represents the total number of pixels in the image, $U = \sqrt{\sum {n_i}^2}$, and n_i represents the number of pixels in the i-th species.

2.2.2.2 Evenness Indices

The **Chao1 richness estimator** (e_{CR}) (Chao, 1984; Eren *et al.*, 2012) uses only the number of singletons (F_1) and doubletons (F_2) and the observed richness (S_{obs}) to write the following estimator for the class richness:

$$e_{CR} = S_{obs} + \frac{F_1^2}{2F_2} \tag{2.21}$$

where F_1 and F_2 are the count of singletons and doubletons gray levels in the image, respectively. According to Chao *et al.* (2006), in the presence of many class abundance distributions, this

estimator, originally derived as an estimate of minimum possible richness (number of pixels pertaining to that specific gray level), is much sharp if the reference sample size is large enough. This corroborates the reason for its use as a valid estimator for large number of gray levels.

The **Gini coefficient** (e_{GC}) (Gini, 1912) is defined as:

$$e_{GC} = \frac{2}{mS^2}\left(\sum_{i=1}^{n}(S+1-i)x_i\right) - \frac{1}{S} \tag{2.22}$$

where x_i is the number of pixels of the i-th gray level ranked from least to most abundant, $i \in [1, S]$ and m is the mean abundance of a gray level – the mean of the x_i values. The Gini coefficient measures income inequality, but can also be used to measure any form of uneven distribution. It ranges between 0 and 1, where 0 denotes a perfect inequality and 1 denotes a perfect equality, where each gray level has the same number of pixels.

The **Heip's evenness** (e_{HE}) is defined as:

$$e_{HE} = \frac{e^H - 1}{S - 1} \tag{2.23}$$

where H is the Shannon-Wiener entropy of counts using logarithm base e.

The **Pielous evenness** $(e_{J'})$ is defined as:

$$e_{J'} = \frac{H}{\ln S} \tag{2.24}$$

Finally, the **Simpsons evenness** e_{SE} is defined as follows, where D is dominance and S_{obs} is the number of observed gray levels:

$$e_{SE} = \frac{\frac{1}{D}}{S_{obs}} \tag{2.25}$$

2.2.3 Experimental Results

We used four texture datasets (Salzburg, Outex, KTH, and CRC) to assess the performance of the proposed E-BiT descriptor, which includes histopathological images (HIs) and natural texture images. These datasets have previously been utilized to evaluate other texture descriptors, including Haralick, GLCM, and LBP (Simon & Uma, 2018a), and were used to evaluate the BiT likewise. HIs usually have structures such as nuclei (shape) and tissue variations (colors) within the same class, which make them more challenging than pure texture images (de Matos, Ataky, de Souza Britto, Soares de Oliveira & Lameiras Koerich, 2021b).

We compare the E-BiT descriptor's performance with the previous BiT texture descriptor, which already achieved state-of-the-art performance on the datasets described above. Our main contribution lies in integrating ten more diversity and evenness indices to build a more discriminant descriptor capable of efficiently classifying textures due to its ability to capture the all-inclusive behavior of texture patterns regardless of the latter forming a non-deterministic complex system. Moreover, like the previous BiT descriptor, the E-BiT descriptor is also permutation, rotation, and reflection invariant. For feature extraction, we followed the same scheme used for the BiT descriptor (Figure 2.5).

2.2.4 Results on Texture and HI Datasets

To demonstrate the benefits of the newly integrated indices, Table 2.12 compares the average accuracy achieved by the E-BiT descriptor and different classification algorithms with the accuracy achieved by BiT, GLCM, LBP, and Haralick descriptors with the same classification algorithms on the Salzburg, Outex and KTH-TIPS datasets.

On the Salzburg dataset, the proposed E-BiT descriptor achieved the best result with SuperL (95.79%). It outperformed not only the best BiT descriptor (94.23%), with a difference of 1.56%, but also all other texture descriptors. The accuracy differences between E-BiT+SuperL and Haralick+SVM and GLCM+k-NN are nearly 8% and 20%, respectively. The E-BiT descriptor also provided the best accuracy on the Outex dataset. The accuracy differences between the

E-BiT and the other descriptors range from 0.12% to 16.79% for BiT+SVM and LBP+SuperL, respectively. Finally, the E-BiT descriptor outperformed all other descriptors on the KTH-TIPS dataset. The E-BiT+SVM achieved the best accuracy of 98.92%, which is 1.05%, 4.03%, and 12.14% higher than the accuracy achieved by BiT+SVM, Haralick+SVM, and GLCM+SuperL, respectively.

Table 2.12 Average accuracy (%) on the test set of KTH-TIPS, Outex, and Salzburg datasets. The best result for each texture descriptor is in boldface

Dataset	Texture Descriptor	Classification Algorithm				
		HistoB	LightB	SuperL	k-NN	SVM
Salzburg	LBP	56.89	57.21	**64.12**	32.23	62.12
	GLCM	68.33	71.77	75.04	**75.38**	63.68
	Haralick	82.27	84.27	86.88	82.54	**87.99**
	BiT	88.20	90.17	**94.23**	88.65	92.33
	Proposed E-BiT	89.32	91.96	**95.79**	89.76	94.20
Outex	LBP	57.80	57.35	**83.21**	48.20	82.41
	GLCM	94.13	**95.52**	94.29	94.37	94.06
	Haralick	96.69	96.53	95.5	**96.92**	96.71
	BiT	99.68	98.53	99.53	99.83	**99.88**
	Proposed E-BiT	**100.0**	99.72	**100.0**	99.69	**100.0**
KTH-TIPS	LBP	59.51	57.18	**64.83**	58.26	61.78
	GLCM	83.12	86.00	**86.83**	74.89	79.83
	Haralick	88.83	90.94	93.00	89.71	**94.89**
	BiT	92.59	94.65	95.41	95.49	**97.87**
	Proposed E-BiT	94.77	94.50	96.45	96.51	**98.92**

k = 3, 5, and 1 for k-NN on Salzburg, Outex, and KTH-TIPS datasets, respectively. Linear kernel and c = 2.0, 1.3, and 1.7 for SVM on Salzburg, Outex, and KTH-TIPS datasets, respectively.

A direct comparison of the results presented in Table 2.12 with other works that used similar datasets may not be reasonable due to differences in the experimental protocols.Most of the results on the Salzburg dataset omit the subclasses used in the experiments and the samples used in the test set. Mehta & Egiazarian (2016) presented a method based on a rotation-invariant LBP that achieved an accuracy of 96.26% with a k-NN on the Outex dataset. Nonetheless, it has the disadvantage of not exploiting color information and global features. Du *et al.* (2016) proposed a rotation-invariant, impulse noise-resistant, and illumination-invariant method based on a local spiking pattern with an accuracy of 86.1% using a neural network also on the Outex dataset. Nonetheless, it does not broaden to color textures and has a large number of

hyperparameters. Hazgui *et al.* (2021) defined a method based on genetic programming and combining HOG and LBP features. With a k-NN, such an approach achieved 91.20% of accuracy on the KTH-TIPS dataset. Nonetheless, color information and global features are not taken into account. Nguyen *et al.* (2016) proposed rotational and noise invariant statistical binary patterns. This method achieved an accuracy of 97.73% on the KTH-TIPS dataset, which is approximately 1.19% lower than that achieved by the E-BiT+SVM. However, this method is resolution sensitive and has high computational complexity. Qi *et al.* (2013) used LBP and Shannon entropy to encode cross-channel texture correlation and investigated the relative variance of texture patterns between different color channels. They proposed a multi-scale rotation-invariant cross-channel LBP (CCLBP). They compute the LBP descriptors in each channel and for three scales, as well as co-occurrence statistics. Such an approach achieved 99.01% accuracy on the KTH-TIPS dataset with an SVM, which is nearly equal to the accuracy achieved by the E-BiT+SVM. Nonetheless, this method is not scaling invariant.

Table 2.13 presents and compares the performance of BiT and E-BiT descriptors. The accuracy differs in 3.69%, 3.13%, 3.04%, 1.90%, 1.07%, and 0.20% for k-NN, SVM, LightB, XGBCB, SuperL, and HistoB classifiers, respectively, The E-BiT descriptor wins in all the cases. Another remark is that, unlike the BiT, the E-BiT obtained its best accuracy with a monolithic classifier (SVM), even though E-BiT+SuperL also outperformed BiT+SuperL.

Table 2.13 Average accuracy (%) of ensemble methods and monolithic classifiers with the BiT and E-BiT descriptors on the CRC dataset

Texture Descriptor	Classification Algorithm					
	XGBCB	HistoB	LightB	SuperL	k-NN	SVM
BiT	90.50	90.71	90.23	**92.52**	83.70	90.98
Proposed E-BiT	91.59	90.91	93.27	93.59	87.39	**94.11**

Finally, Table 2.14 compares the best result of Table 2.13 achieved with the E-BiT+SVM to the state-of-the-art for the CRC dataset. The E-BiT outperforms almost all other methods in terms of accuracy. For example, considering an 8-class classification task and a 10-fold CV, the difference in accuracy between the second-best method (shallow) is 1.15% and 1.71% for the

third-best method (CNN). In addition, the E-BiT descriptor slightly outperformed the second-best method (CNN) for 5-fold CV, with a difference of 0.67%. The results highlight the E-BiT descriptor's advantages over shallow and deep methods. Furthermore, CNNs require massive labeled datasets, pre-trained models, and data augmentation strategies to learn high-quality representations. Nonetheless, this may not always be possible in medical imaging.

Table 2.14 Average accuracy (%) of state-of-the-art of deep and shallow approaches on the CRC dataset

Reference	Approach	10-fold CV	5-fold CV
Ataky & Lameiras Koerich (2022)	Shallow	92.96	–
Ribeiro *et al.* (2019)	Shallow	97.60*	–
Sarkar & Acton (2017)	Shallow	73.60	–
Kather *et al.* (2016b)	Shallow	96.90*	–
Kather *et al.* (2016b)	Shallow	87.40	–
E-BiT+SVM	Shallow	**94.11**	**93.27**
Wang *et al.* (2017)	CNN	–	92.60
Pham (2017)	CNN	–	84.00
Raczkowski *et al.* (2019)	CNN	92.40	92.20

*Used 2-classes classification instead (malignant and benign)

2.2.5 Invariance of the E-BiT Descriptor

We have also verified the invariant properties of new aggregated diversity indices. We applied different geometric transformations to each image (Figure 2.9) and computed all newly integrated features. The non-normalized feature values are shown in the Tables 2.15 and 2.16 for a texture image from KTH and CRC datasets, respectively.

The values of the E-BiT descriptors presented in Tables 2.15 and 2.16 show that: (i) all measurements are reflection and rotation invariant, as they have the same values of original texture images and HIs. This also substantiates the fact that the E-BiT descriptors capture the all-inclusive behaviors of patterns within an image; (ii) Simpson index (d_c), Gini coefficient (e_{GC}), Heip's evenness (e_{HG}), Pielous evenness ($e_{j'}$), McIntosh dominance diversity (d_{McInt}), and Simpson evenness (e_{SE}) are scale-invariant as they provided values of the order of original images. On the other hand, most diversity measures based on abundance and richness exhibit some scale dependence. By rescaling the original image, we affect the proportion of both

factors, which affects the resulting values either inversely or directly. However, normalizing such measurements by the total number of pixels can mitigate this effect. On the other hand, indexes based on evenness are based on the equitability of taxa frequencies in a community. As a result, they are unaffected by the change in scaling because evenness is determined by the intrinsic properties of the ecosystem (image).

Table 2.15 Non-normalized feature values computed from different image transformations applied to a texture image

E-BiT Descriptors	Original	Transformations				
		Rotation		Reflection		Rescaling
		90^o	180^o	Horizontal	Vertical	50%
d_{HB}	9.6646	9.6646	9.6646	9.6646	9.6646	8.2780
e_{Dw}	0.0364	0.0364	0.0364	0.0364	0.0364	0.0372
d_C	0.9999	0.9999	0.9999	0.9999	0.9999	0.9998
d_{ENS}	16263.1	16263.1	16263.1	16263.1	16263.1	4064.3
e_{CR}	16384.0	16384.0	16384.0	16384.0	16384.0	4096.0
e_{GC}	0.0493	0.0493	0.0493	0.0493	0.0493	0.0503
d_{McInt}	0.9999	0.9999	0.9999	0.9999	0.9999	0.9997
e_{HE}	0.9963	0.9963	0.9963	0.9963	0.9963	0.9962
$e_{J'}$	0.9996	0.9996	0.9996	0.9996	0.9996	0.9995
e_{SE}	0.9926	0.9926	0.9926	0.9926	0.9926	0.9923

Table 2.16 Non-normalized feature values computed from different image transformations applied to an HI

E-BiT Descriptors	Original	Transformations				
		Rotation		Reflection		Rescaling
		90^o	180^o	Horizontal	Vertical	50%
d_{HB}	10.3017	10.3017	10.3017	10.3017	10.3017	8.9275
e_{Dw}	0.0509	0.0509	0.0509	0.0509	0.0509	0.0477
d_C	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999
d_{ENS}	30136.5	30136.5	30136.5	30136.5	30136.5	7631.3
e_{CR}	30625.0	30625.0	30625.0	30625.0	30625.0	7744.0
e_{GC}	0.0703	0.0703	0.0703	0.0703	0.0703	0.0663
d_{McInt}	1.0001	1.0001	1.0001	1.0001	1.0001	1.0001
e_{HE}	0.9921	0.9921	0.9921	0.9921	0.9921	0.9928
$e_{J'}$	0.9992	0.9992	0.9992	0.9992	0.9992	0.9992
e_{SE}	0.9840	0.9840	0.9840	0.9840	0.9840	0.9850

2.2.6 Final Considerations

This section proposed an extended version of the bio-inspired texture descriptor (BiT) to characterize texture in images. This extension, named E-BiT, combines global ecological concepts of species diversity, evenness, richness, and taxonomic indexes to resemble the completeness of diversity as a generalization of biodiversity analysis. That allows the development of a descriptor that captures the all-inclusive behavior of texture image patterns, both at local and global levels. Furthermore, the E-BiT descriptor is insensitive to scale, translation, or permutation.

Compared to related methods for texture characterization and its baseline version, the E-BiT descriptor emerges as a promising tool for texture characterization, achieving state-of-the-art performance for texture classification. In addition, its generic nature is notable, as it performs well in both natural and histopathologic images.

This research could be extended to examine how the E-BiT descriptor behaves at different spatial scales and resolutions. That may allow for the most effective extraction of texture properties while maintaining or even improving performance.

CHAPTER 3

MULTI-SCALE AND MULTI-RESOLUTION TEXTURE ANALYSIS

In the context of vegetation measurements and summarizing values to describe site communities to a great extent, it is, by and large, desirable to compare two communities of species and determine how similar they may be. Such similarity measures may examine differences between two site units, similar sites under different management practices, changes that may have occurred because of disturbance (e.g., the similarity between arid and humid vegetation), and variation between different study times on the same site (e.g., determine how similar the communities are to what it was a decade ago), and comparing a site to a desirable state (described as a referent). The latter entails gathering data from the exact location throughout time to make a credible comparison to its original (reference) condition.

In restoration ecology, a site's hydrology, soil, climate, and biology may change over time, making summing values to strongly define site communities or spatial comparisons of biodiversity therefrom ineffective if such changes in primary resources are not taken into account (Dong, Hou, Xu, He & Liu, 2018). Nevertheless, taking into account differentiation should reveal the temporal trends of the biodiversity conservation capacity. For example, such a concept evaluates ecological environments and river health. Thus, in essence, channel split and color spaces, and multi-scale and multi-resolution analysis, such as wavelet decomposition and pyramids, are employed to supply a characteristic value that can reflect the background condition and summarize values that describe texture from an image to a great extent under distinct differentiations.

In this chapter, we leverage the multi-scale and multi-resolution analysis for texture analysis of natural and histopathologic images in the following sections.

3.1 Multi-resolution Texture Analysis of Histopathologic Images Using Ecological Diversity Measures

Histopathology studies how a specific disease affects cells (tissue). Usually, a biopsy study is done using a microscope and dyes. It can also be done during surgery or an autopsy (death investigation). Histopathologic images (HIs) are medical imaging obtained via microscopy of tissues from biopsies, which allow specialists to observe tissue characteristics on a cell basis. This process consists of tissue processing by chemical fixation or frozen section slides (creating the slide with the staining process). Afterward, the slides undergo staining with one or several pigments to envision the tissue through a microscope, aiming to reveal cellular components; counterstains are used to provide contrast; and finally, the pathologist analysis. The stain, by and large, used in histopathology is a combination of hematoxylin and eosin (H&E). The former is employed to stain nuclei (blue), while the latter stains cytoplasm and the extracellular connective tissue matrix (pink). However, the staining process may give rise to a variance in the analysis process because H&E is prone to produce different color intensities being conditioned by the temperature, storage conditions, and brand. Nevertheless, as stated by de Matos *et al.* (2021a), HIs continue to be the gold standard for evaluating several types of tumors for cancer diagnosis.

Computer-aided detection (CAD) and computer-assisted diagnostic (CADx) systems are continuously developed to assist in medical image analysis. Clinicians are heavily reliant upon CAD for cancer detection and monitoring. However, given the reliance on CAD and CADx for cancer detection, there is always an extra focus and need for developing systems that improve pathologists' productivity and ameliorate the reliance on outcomes by adding consistency to the diagnosis process and reducing observer subjectivity. Machine learning (ML) approaches are increasingly used in CAD and HI analysis to diagnose cancer in various tissues or organs, such as the breast, prostate, skin, brain, bones, liver, etc. Furthermore, when used in HI analysis, ML approaches reveal potential benefits. As a result, they have seen much use in tasks like feature extraction, classification, and segmentation. The visual properties of macro vision images used in machine learning applications, such as scene reconstruction, and object and face recognition, differ from those of HIs, which contain complex textures and rich geometric structures.

Into the bargain, computational tools have been proposed to help the specialist interpret the breast HI exam, providing features for detecting and diagnosing tumors and cancerous cells. Detecting tumors with a high sensitivity rate while reducing false positives remains challenging. Texture descriptors have become popular in HI analysis due to the variability of texture found in such images and tissue appearance due to irregularities in the staining process. Variability in staining protocol, such as fixation, inconsistency in staining condition, and reagents, may exist between laboratories or within the same laboratory. Moreover, textural feature extraction for discriminantly quantifying HI information is challenging because the distribution of intrinsic properties of such images is complex and nonlinear.

This section presents a method for characterizing texture across HIs with a reasonable success rate. We formulate that it is possible to quantify the intrinsic properties of HIs to the maximum extent by combining biodiversity, taxonomic measures, and discrete wavelet transform. Thus, the main highlights are:

- An information-theoretical measure of ecological diversity for HI texture characterization.
- The exploitation of independent wavelet subband coefficients' non-linear interactions across time.
- The mixture of wavelet features and statistical properties of taxonomic indexes represents an unexplored method.
- Such a mixture characterizes HIs, so intrinsic properties have provided promising performance for real-world HI datasets such as the CRC and the BreakHis datasets.

3.1.1 Ecological Modeling of Wavelet Subbands

The proposed method, illustrated in Figure 3.1, is employed and integrated for texture classification as follows: (1) image channel splitting; (2′) wavelet subband decomposition; (2″) computation of biodiversity measures and information theory from each channel (R, G, and B) to form a features vector; (3) computation of taxonomic indexes and information theory from each wavelet subband to form a features vector; (4) feature vectors concatenation; and (5) classification and performance evaluation.

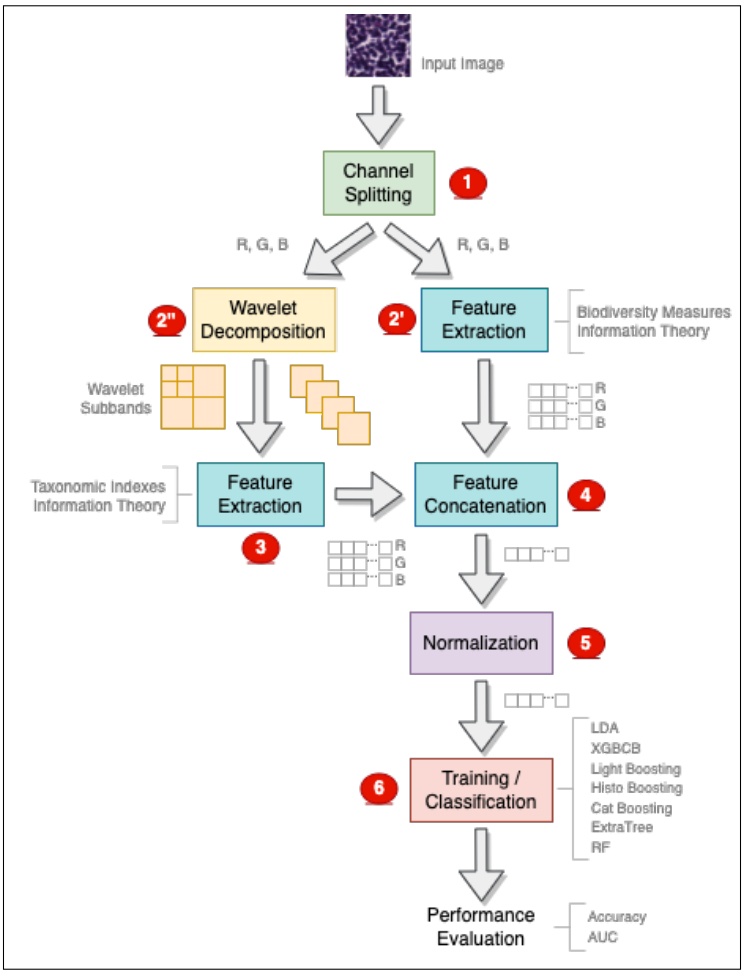

Figure 3.1 General overview of the proposed scheme

To elaborate, Algorithm 3.1 illustrates a step by step pseudocode of the proposed architecture. In regard of the discrete wavelet decomposition, numerous wavelets transform, such as Haar, Daubechies, and Symlets are available. Nonetheless, the decision to use one of them may be influenced by how well it fits in a given situation. In this work, we used Daubechies. In future work, we will verify how different orthogonal wavelets and other families of wavelets will perform.

1. **Channel Splitting and Feature Extraction**

   We applied the integrative method to make each image channel (R, G, B) a separate input. The motivation is to exploit color information and summarize values that describe texture from an image to a great extent under distinct differentiation, as mentioned above. Therefore,

we represent and characterize an input image using a set of local descriptors derived from a pixel's interaction with its neighbors in a particular channel (R, G, or B) and wavelet subbands thereof.

After the channel splitting stage, each image channel undergoes feature extraction. Intrinsic properties and discriminant characteristics within each input channel are extracted as follows. First, for each channel (R, B, and G), we perform a multi-resolution analysis of a texture employing a wavelet transform, generating four subbands: *a, h, v*, and *d*. Since we used three levels, in the following scale, the subband *a* is used for DWT computation. After, we compute taxonomic measures, Shannon entropy, and total information, resulting in a 9-dimensional vector for each subband. This sequence can be seen in Figure 3.1 (steps 2′ and 3). Taxonomic measures comprise taxonomic diversity, taxonomic distinctness, the sum of phylogenetic distances, average distance from the nearest neighbor, extensive quadratic entropy, intensive quadratic entropy, and total taxonomic distinctness (Section 2.1). Because we have split an input image into three channels and chosen a 3-level wavelet decomposition (leading to 10 subbands), step 3 will produce a 270-dimensional feature vector (9×3×10). In parallel, we compute biodiversity measures and Shannon entropy, and total information directly from each original image channel resulting in a 9-dimensional vector for each channel. Biodiversity measures from step 2″ comprise Margalef's and Menhinick's diversity indexes, Berger-Parker dominance, Fisher's alpha diversity metric, Kempton-Taylor index of alpha diversity, McIntosh's evenness measure, and Shannon-Wiener diversity index (Section 2.1). The reason for not employing DWT in this stage is that low-pass decomposition filters may present negative coefficients in the subband *a* depending on filter coefficients used for decomposition. Nevertheless, biodiversity measures such as abundance and richness are non-negative. Because we have split an input image into three channels, step 2″ will produce a 27-dimensional feature vector (9×3). Finally, feature vectors resulting from steps 2″ and 3 are concatenated to form the final feature vector (step 4 in Figure 3.1). We named it the BiTW descriptor because it results from the concatenation of biodiversity measures, information theory and taxonomic indexes extracted from wavelet subbands.

The BiTW is a 297-dimensional feature vector, which may leave a whole path to a possible feature selection step.

2. **Normalization**

   The feature vectors are split into training and test sets before the training step. Then, normalization and scaling occur independently on each feature in the training set, where values are normalized to the range $[0, 1]$ using the min-max normalization. Minimum and maximum are then stored to be used on feature normalization over the testing data. The same procedure is used for the k-fold cross-validation (CV), where feature vectors are split into k folds and computing the min-max pairs in the merged training folds. The min-max pairs calculated from the trained data normalize the training and test folds. This procedure is repeated during the CV stage for each new training/test fold.

3. **Training and Classification**

   After feature extraction, the resulting feature vectors are taken through the classification process utilizing seven classifiers: a histogram-based algorithm for building gradient boosting ensembles of decision trees (HistoB), light gradient boosting decision trees (LightB), fast, scalable, high-performance gradient boosting decision trees (CatBoost), extra trees (ExtraT), random forest (RF), gradient boosting decision trees (GB), and linear discriminant analysis (LDA). A performance analysis is carried out using accuracy and the area under the curve (AUC).

Algorithm 3.1 FeatureExtractionProcedure

Result: final feature descriptor
0. Load a RGB image file ;
1. Channel Splitting R, G, B;
2. For each channel ;
  2.1 Extract Biodiversity measures (equations 1- 7) and Shannon Entropy, append to V1 ;
3. For each channel ;
  3.1 Generate three wavelet levels decomposition ;
  3.2 For each subband of each decomposition level (from step 3.1) ;
    3.2.1 Compute taxonomic indices and Shannon entropy, append to V2 ;
4. Concatenate V1 and V2 into a single vector (297-dimensions);
5. Repeat steps 1 to 5 for all images of the dataset

3.1.2 Experiments and Results

Two medical datasets were used in the experiments, **CRC** (Kather *et al.*, 2016a)(Figure 2.7) and **BreakHis** (Spanhol *et al.*, 2016a)) (Figure 2.8) to evaluate the performance of the BiTW descriptor, as well as the experimental protocol used to assess the properties of the BiTW descriptor and its performance on classification tasks using various classification and ensemble methods. We contrast the performance of the BiTW descriptor with that of shallow and deep state-of-the-art approaches.

3.1.2.1 Experiments on CRC

The accuracy of both monolithic classifiers and ensemble methods trained with feature vectors from BiTW on the CRC dataset is presented in Table 3.1. GB produced the best results for both train-test split and k-fold CV in the group of all classification algorithms. Similarly, we computed the area under the ROC curve (AUC), another essential metric commonly used in medical images that compares the true-positive rate to the false-positive rate at various threshold levels. The BiTW descriptor with GB achieved an AUC of 0.99. AUCs of 0.7 to 0.8 are considered acceptable, 0.8 to 0.9 are regarded as excellent, and greater than 0.9 is deemed exceptional.

Table 3.1 Accuracy (%) and AUC of monolithic classifiers and ensemble methods with the BiTW descriptor on the CRC dataset for train-test split and 10-fold CV

Experimental Protocol (Metric)	Classification Algorithms						
	HistoB	LightB	LDA	CatB	ExtraT	RF	GB
70/30 (Acc)	92.42	92.20	89.27	92.20	92.42	92.42	**93.28**
70/30 (AUC)	0.991	0.992	0.988	0.993	0.993	0.991	0.991
10-fold (Acc)	91.12	91.81	89.45	91.33	91.53	90.59	**93.73**
10-fold (SD)	±0.05	±0.01	±0.08	±0.02	±0.03	±0.02	±0.02
10-fold (AUC)	0.990	0.992	0.989	0.993	0.993	0.991	0.994

Acc: Accuracy; SD: Standard Deviation.

Table 3.2 compares the results achieved by BiTW + GB with the state-of-the-art for the CRC dataset. The BiTW slightly outperforms the accuracy achieved by almost all other methods. For

instance, the difference in accuracy to the second-best method (shallow) is 0.77%, and with the third-best method (CNN) is 1.33%, considering an 8-class classification task and 10-fold CV. For 5-fold CV, BiTW slightly outperformed the second-best method (CNN) with a difference of 1.0%. Regarding the works that lay on shallow methods, Kather *et al.* (2016b) considered six distinct sets of descriptors to describe the texture of histological images: lower-order and higher-order histogram features, Local binary patterns (LBP), Gray-level co-occurrence matrix (GLCM), Gabor filters, Perception-like features, and Combined feature sets. Before computing the texture features, all images were converted to greyscale. Classifiers such as 1-nearest neighbor, linear SVM, radial-basis function SVM, and decision trees were used for evaluation purposes. The BiT (Ataky & Lameiras Koerich, 2022) and DWT used to build the proposed method share characteristics with LBP, GLCM, and Gabor. In their study, Sarkar & Acton (2017) present a novel method of saliency-based sparse coding and vocabulary learning for computing image similarity between two images. The system extracts prominent image features by utilizing a salient object recognition technique. The developed dictionary learning technique takes advantage of saliency to create sparse codes and a set of basis functions for picture representation so that the more salient image regions significantly impact dictionary learning. As in this work, the descriptor used in the proposed method allows extracting local and global information from the image to the maximum extent, as wavelet transform generates a signal representation in both the time and frequency domains, enabling quick access to the signal's localized information.

It is noteworthy to highlight that CNNs generally require massive labeled datasets and, when not possible, may need pre-trained models and data augmentation to learn high-quality representations. In the medical field, however, this is not always possible. The shallow approaches using the BiTW descriptor, unlike the latter, did not require any data augmentation on the CRC dataset and have proven to be promising relative to CNNs as well as other shallow methods, despite HIs having other structures than textures.

Elaborating, in the context of histopathological images, the future of feature extraction methods are moving away from morphological information and toward high-level representations, owing to the challenge presented by the availability and capability of capturing images with high

resolutions, such as the WSI. Among the high-level characteristics, deep learning methods as extractors are gaining popularity. They extensively use parallelism methods, providing good performance for their application in large images and allowing knowledge transfer. Another upper hand of such methods is that they can be trained and adapted to specific characteristics of the problem to be solved, resulting in representation learning. Notwithstanding deep learning methods (deep learning-based feature extractors), such as convolutional neural networks have recently gained popularity because they are end-to-end with good performance for image classification (for capturing more shape features), it is also put forth that using CNNs directly (from scratch) on texture datasets results in only moderate accuracy in texture classification (Hafemann, Oliveira & Cavalin, 2014). Improving the performance of CNNs requires millions of parameters and a large amount of data (which is not always possible in the medical domain), bringing high computational costs. Such an improvement, that is, a non-direct use of CNNs, are found in the works of Wang *et al.* (2017), Pham (2017), and Raczkowski *et al.* (2019).

Table 3.2 Average accuracy (%) of shallow and deep approaches on the CRC dataset for 5-fold CV, 10-fold CV, and AUC

Reference	Approach	Accuracy (%)		AUC
		10-fold	5-fold	
Ribeiro *et al.* (2019)	Shallow	97.60*	–	**0.994**
Sarkar & Acton (2017)	Shallow	73.60	–	–
Kather *et al.* (2016a)	Shallow	96.90*	–	–
Kather *et al.* (2016b)	Shallow	87.40	–	–
Naiyar, Asim & Shahid (2015)	Shallow	–	–	0.960
Rathore, Iftikhar, Hussain & Jalil (2013)	Shallow	–	–	0.970
Kalkan, Nap, Duin & Loog (2012)	Shallow	–	–	0.950
Masood & Rajpoot (2009)	Shallow	–	–	0.900
Ataky & Lameiras Koerich (2022)	Shallow	92.96	–	–
BiTW+GB	Shallow	**93.73**	**93.60 ±0.2**	**0.994**
Wang *et al.* (2017)	CNN	–	92.60	–
Pham (2017)	CNN	–	84.00	–
Raczkowski *et al.* (2019)	CNN	92.40	92.20	–

*Used 2-classes classification instead (malignant and benign).

Raczkowski *et al.* (2019) created and trained a CNN model whose architecture, ARA-CNN, was inspired by various cutting-edge systems, such as Microsoft ResNet and DarkNet 19. They utilized the renowned batch normalisation approach for normalization and to avoid overfitting. In ARA-CNN, dropout is also used to reduce overfitting. Pham (2017) advocated texture scaling

in CRC histology as a way to reduce the computational burden of network training. Only two hidden layers were built for deep learning; according to the author, a more significant number of hidden layers would be expected to demonstrate the superior performance of deep neural networks regarding classification accuracy and machine-learning time. Wang *et al.* (2017) introduced a novel BCNN-based method for the classification of histopathological pictures, which first decomposes the images into hematoxylin and eosin stain components and then applies BCNN to the decomposed images to fuse and improve the feature representation performance. The bilinear CNN (BCNN) is a new CNN model for classification at the granular level. At each spatial point, the convolutional-layer outputs of two CNNs are multiplied with the outer product to form the BCNN.

3.1.2.2 Experiments on BreakHis

Table 3.3 and 3.4 present the accuracy of both the monolithic classifiers and ensemble methods trained with BiTW feature vectors on the BreakHis dataset with train-test split and k-fold cross-validation, respectively. By employing the train-test split, the LightB achieved the best accuracy of 99.26% and 98.62% for 40× and 200× magnifications, respectively. The ExtraT achieved 98.50% and HistoB 98.38% accuracy for 100× and 400× magnifications, respectively. Furthermore, the AUC is nearly 0.98 or above, regardless of the classifier or ensemble method. Considering the k-fold CV, ExtraT yielded the best accuracy of 98.75% and 98.63% for 40× and 100× magnifications, respectively. The LightB, in turn, achieved an accuracy of 98.72% for 200× magnification. Finally, HistoB yielded an accuracy of 98.38% for 400× magnification. Furthermore, we carried out experiments with k-fold cross-validation to guarantee that each sample from the dataset does have a chance of being selected in the training and test sets, which is a best practice when working with limited data. Table 3.4 presents the average accuracy with the BiTW descriptor on the BreakHis dataset at the image level with a 10-fold CV. The results for both types of dataset splitting are very similar for nearly all the classifiers and ensemble methods.

Table 3.5 contrasts the results achieved by the proposed approach with the state-of-the-art for the BreakHis dataset. Though different classifiers outperformed others for different magnifications, for the sake of fairness, we choose a classifier to present for comparison with related work based on the average performance of the four magnifications. Computing the average we obtain 98.58%, 98.54%, 93.21%, 97.95%, 98.50%, 97.65% and 97.53%, for HistoB, LightB, LDA, CatB, ExtraT, RF and GB. Thus, HistoB is chosen for comparison purposes because it presented the highest average accuracy. The BiTW descriptor with HistoB achieved a substantial accuracy of 98.97%, 98.43%, 98.55% and 98.38% for 40×, 100×, 200×, and 400× magnifications, respectively. It is important to note that, regardless of magnification, the proposed method surpasses the accuracy of deep and shallow methods. The differences in accuracy between the proposed method and the second- and third-best methods are 1.47% (Shallow) and 1.97% (CNN), 0.93% (CNN) and 1.63% (Shallow), 1.35% (CNN) and 2.75% (Shallow), 1.18% and 3.18% (Shallow) for 40×, 100×, 200×, and 400× magnifications, respectively.

Analyzing the CNN results of related research leads us to believe that a more significant number of hidden layers, fine-tuning of hyperparameters, and a large amount of training data would enable CNNs methods to achieve even more accurate classification over the method proposed in this section. Notwithstanding, CNNs continue to have challenges with explainability and interpretability. In comparison, the proposed method requires a relatively small quantity of data and processing effort to compute. Moreover, it is generic (since it worked well on various image types, such as natural images) and does not require retraining or hyperparameter setup while achieving state-of-the-art performance, as demonstrated by experiments on HIs datasets. Moreover, most of the descriptors used to extract features in the related works that lay on shallow methods, which were shown to be outperformed by the BiT (Ataky & Lameiras Koerich, 2022) descriptor proposed in this thesis.

Table 3.3 Accuracy (%) and AUC of monolithic classifiers and ensemble methods with the BiTW descriptor on the BreakHis dataset at image level with train-test split

Image Magnification	Classification Algorithm							
	Metric	HistoB	LightB	LDA	CatB	ExtraT	RF	GB
40×	Acc	98.97	**99.26**	**99.26**	98.25	98.97	98.25	98.18
	AUC	0.992	0.995	0.993	0.981	0.985	0.988	0.987
100×	Acc	98.43	98.30	92.41	97.19	**98.50**	97.19	96.64
	AUC	0.991	0.990	0.982	0.987	0.992	0.987	0.980
200×	Acc	98.55	**98.62**	90.88	**98.62**	98.55	98.00	97.58
	AUC	0.989	0.993	0.963	0.994	0.992	0.989	0.988
400×	Acc	**98.38**	97.98	90.29	97.74	97.98	97.19	97.74
	AUC	0.993	0.990	0.983	0.987	0.989	0.982	0.986

Table 3.4 Average accuracy (%) of monolithic classifiers and ensemble methods with the BiTW descriptor on the BreakHis dataset at image level with 10-fold CV

Image Magnification	Classification Algorithm						
	HistoB	LightB	LDA	CatB	ExtraT	RF	GB
40×	98.62	98.61	98.61	98.06	**98.75**	98.04	98.12
100×	98.48	98.45	91.60	97.74	**98.63**	97.87	97.76
200×	98.54	**98.72**	91.55	97.98	98.65	97.98	98.53
400×	**98.76**	98.74	91.00	98.08	98.33	97.58	97.77

Table 3.5 Average accuracy (%) on the BreakHis dataset of shallow and deep approaches. For training and testing, all of these works used the same data partitions

Reference	Method	Image Magnification			
		40×	100×	200×	400×
Alom *et al.* (2019)	CNN	97.00	97.50	97.20	97.20
Han *et al.* (2017)	CNN	92.80	93.90	93.70	92.90
Bayramoglu *et al.* (2016)	CNN	83.00	83.10	84.60	82.10
Spanhol *et al.* (2016b)	CNN	90.00	88.40	84.60	86.10
Gandomkar, Brennan & Mello-Thoms (2018)	CNN	94.10	93.20	94.70	93.50
Bardou, Zhang & Ahmad (2018)	CNN	88.23	84.64	83.31	8.98
Nawaz, Ahmed, Tahir & Khan (2018)	CNN	–	95.00	–	–
Spanhol *et al.* (2016a)*	Shallow	75.60	73.00	72.90	71.20
Spanhol *et al.* (2016a)+	Shallow	74.70	76.80	83.40	81.70
Erfankhah *et al.* (2019)*	Shallow	88.30	88.30	87.10	83.40
Ataky & Lameiras Koerich (2022)†	Shallow	97.50	96.80	95.80	95.20
BiTW + HistoB	Shallow	**98.97**	**98.43**	**98.55**	**98.38**

*LBP descriptor; +GLCM descriptor;†BiT descriptor.

3.1.3 Discussion

The proposed approach was assessed with two HI datasets, both with eight classes. The experiment protocol employed a train-test split (70/30) and a k-fold CV. For either experimental protocol, the results led to the following findings:

- Exploiting information-theoretical measures of ecological diversity indices in conjunction with non-linear interactions of single and independent wavelet subband coefficients throughout time yielded promising results.
- Although HIs contain other structures than texture, it was possible to characterize texture and achieve a reasonable discriminating capability by employing biodiversity measures and taxonomic indexes together with multi-resolution analysis through DWT.
- Such a mixture allowed the characterization of HIs to such an extent that intrinsic properties have provided a promising performance for the real-world datasets classification, reaching 93.73% accuracy and 0.994 AUC for the CRC dataset. Regarding the results on the BreakHis dataset, the accuracy were 99.26%, 98.50%, 98.62%, and 98.76% for 40×, 100×, 200× and 400× magnification, respectively. The AUC of all the magnifications was above 0.98.
- By and large, results for HIS are better than for the CRC dataset.

Overall, the proposed approach outperformed state-of-the-art shallow and deep works on CRC and BreakHis datasets, regardless of the non-textural information that HIs may contain.

3.1.4 Final Considerations

The current research leveraged the information-theoretical measure of ecological diversity indices together with a discrete wavelet transform to characterize texture across HIs. We explored the interactions of individual wavelet subband coefficients over time and modeled each as an ecosystem from which measures of biodiversity and statistical properties of taxonomic indexes are extracted to represent HI texture effectively. We stated that by combining measurements of biodiversity from each HI channel and taxonomic indexes extracted from different wavelet subbands, it should be possible to quantify the intrinsic properties of such images to the maximum

extent. The mixture of wavelet features and statistical properties of ecology diversity indexes represent a novel method and a promising tool for the quantification of the intrinsic properties of texture across HIs. Wherefore, the experimental results have shown an increase in terms of texture discrimination over both HI datasets. Moreover, the proposed method outperformed several shallow and deep state-of-the-art methods.

Withal, it is noteworthy that the proposed method presents the following limitations: (i) Specifying the expected dynamic ranges of the biodiversity and taxonomic indices is not possible. The ranges of the biodiversity and taxonomic indices are necessarily related to the species richness, abundance, and their relationship within an image and a wavelet subband of an image, either directly or inversely. The dynamic range is, therefore, not known beforehand. Thereby, feature normalization may be necessary; (ii) Images are usually corrupted with noise introducing random variations to gray levels so that homogeneous regions present a gray-level dispersion depending on noise power, decreasing the performance. So far, the proposed method does not provide, to some extent, a tolerance parameter to cope with such noise-produced variations that helps to decide if a species (gray-level) or individual (pixel) should be considered or not. Thus, we suggest considering the preprocessing before the feature extraction phase.

In future work, we intend to further improve the classification accuracy by exploring various color spaces and fractal features for dealing with multi-resolution aspects of the textures that will bring on discriminative information for characterizing HI categories. Furthermore, because an increasing number of DWT-level decompositions can result in high-dimensional feature vectors, we want to look into appropriate feature selection approaches or dimensionality reduction techniques to find feature subspaces that not only reduce texture features dimensionality but also improve the proposed method's class discrimination capability. Furthermore, we will also focus on finding a tolerance parameter to minimize performance issues when we apply our method in the presence of noise.

The following section presents how texture classification can be improved by leveraging the multi-scale / multi-resolution analysis and combining different descriptors.

3.2 Multi-scale Analysis for Improving Texture Classification

Information from an image occurs over multiple and distinct spatial scales. Image pyramid multi-resolution representations are a proper data structure for image analysis and manipulation over a spectrum of spatial scales. Likewise, even though the descriptors presented in Chapter 2 have shown a significant discriminative ability, combining them with other descriptors may be a promising strategy to provide a representation based on several intrinsic textural properties. With this in mind, we evaluate the combination of texture descriptors such as the BiT (Ataky & Lameiras Koerich, 2022), information-theoretic measures (Shannon, 1948), GLCM (Haralick *et al.*, 1973), and Haralick features (Haralick, 1979) to improve the performance of texture classification. To this end, individually, the feature vector representing an image is formed by concatenating features provided by different descriptors. The rationale is to compensate for the possible loss when using a single technique to give the textural description. This process is done over a spectrum of spatial scales representation using Gaussian-Laplacian (GLP), a proper data structure for image analysis and manipulation.

The highlights of this section are threefold, as follows:

- The combination of BiT descriptor, information theory, Haralick, and GLCM features for texture characterization.
- A better discriminating ability while using color and gray-scale features on different categories of images.
- A method for texture classification that represents the state-of-the-art on challenging datasets.

3.2.1 Proposed Approach

This section describes how the proposed method integrates multi-resolution analysis and multiple texture descriptors for texture classification. To this end, we put forward an architecture structured in five main stages as follows. Figure 3.2 shows an overview of the proposed scheme. Algorithm 3.2 shows the steps of the first three stages.

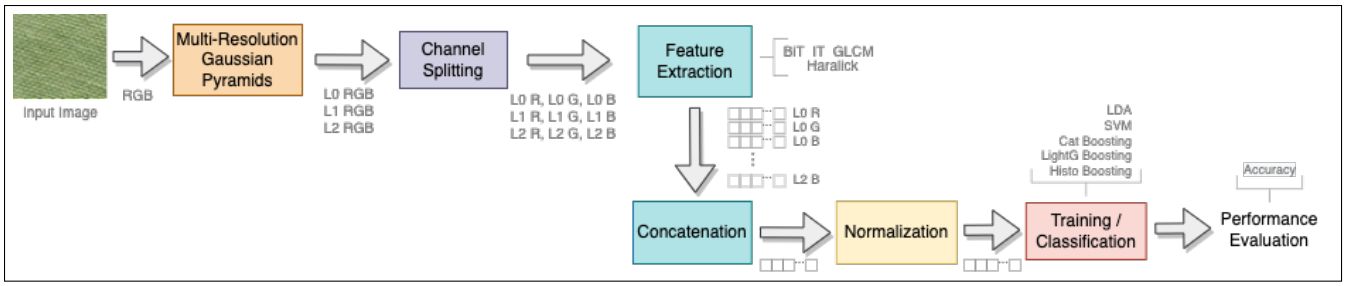

Figure 3.2 An overview of the proposed scheme

Multi-resolution representation: in this stage, for an input image, we generate three others corresponding to the levels of the Gaussian pyramid (L0: original image, L1: first level of the pyramid, L2: second level of the pyramid). The purpose is to represent an input image in different resolutions to capture intrinsic details.

Channel Splitting: in this phase, each image channel (R, G, B) is considered as a separate input. The key reason behind the splitting channels is to exploit color information. Thus, we represent and characterize an input image in a given resolution by a set of local descriptors generated from the interaction of a pixel with its neighborhood inside a given channel (R, G, or B).

Feature Extraction: after the channel splitting step, the images undergo feature extraction, which looks for intrinsic properties and discriminative characteristics within. For each level of GP of an image, we extract: BiT (Ataky & Lameiras Koerich, 2022) (14-dimensional), Shannon entropy and multi-information, aka total information (3-dimensional), Haralick (Haralick *et al.*, 1973) (13-dimensional), and GLCM (Haralick *et al.*, 1973) (6-dimensional) from each channel. Images are then represented by the concatenation of different measurements organized as feature vectors (Algorithm 2.1). For simplicity, we name the resulting descriptor *Three-in-One* (TiO). It is worth mentioning that images were converted in gray-scale for GLCM and Haralick measurement to be extracted but remained color for extracting the bio-inspired indices. After feature extraction and before concatenation, the feature vectors have 126, 117, 54, and 18 dimensions for BiT, Haralick, GLCM, and Shannon entropy and multi-information, respectively. Thereby, after concatenation, we have a 315-dimensional feature vector.

Normalization: because test points simulate real-world data, we split the data into train and test sets and perform min-max normalization on the training data. Subsequently, we used the same min-max values to normalize the test set.

Training/Classification: we have a normalized feature vector comprising the concatenation of features extracted from each resolution and color channel from the previous stage. This texture representation is used to train one monolithic and three ensemble-based models. The linear discriminant analysis (LDA) is used for the monolithic classification model.

In contrast, the ensemble models are created using the histogram-based algorithm for building gradient boosting ensembles of decision trees (HistoB), the light gradient boosting decision trees (LightB), and the CatBoost[1], which is an efficient implementation of the gradient boosting algorithm.

Algorithm 3.2 FeatureExtractionProcedure

Result: feature descriptor
1 **1. Read** a RGB image file ;
2 **2. Generate 3 multi-resolution levels with Gaussian pyramids L0, L1, L2**;
3 **3. For each level (L0, L1, L2) of Gaussian pyramids ;**
4 **3.1. Separate the image in channels R, G, B;**
5 **3.2. Convert R, G, and B to grayscale images (for GLCM and Haralick only);**
6 **3.3. Compute Bio-inspired indices, Shannon entropy and multi-information, Haralick, and GLCM of R, G, B ;**
7 **3.4. Concatenate these values into a single vector (315-dimensional);**
8 **4. Repeat steps 1 to 4 for all images of the dataset**

3.2.2 Experimental Results and Discussion

The used experimental protocol considers five datasets, three of which are texture image collections, and two are composed of medical images. In addition to the KTH-TIPS, Outex, CRC, and BreakHis datasets presented in Section 2.1, we also used the Amsterdam Library of Textures (ALOT) (Napoletano, 2017), which is a color image collection of 250 (classes) rough textures.

[1] https://catboost.ai/

The authors systematically varied the viewing angle, illumination angle, and illumination color for each material so that they could capture the sensory variation in object recordings.

We split each dataset into training (70%) and testing (30%) sets to evaluate all the classifiers on the same experimental condition. In the experiments, the full feature vector generated by concatenating all descriptors is used for training a monolithic classifier using the LDA algorithm and ensembles using the HistoB, LightB, and CatBoost algorithms.

3.2.2.1 Experiments with Texture Datasets

Table 3.6 presents the results of the proposed method (TiO), BiT, GLCM, and Haralick over Outex, KTH-TIPS, and ALOT datasets. For each dataset, we present the results obtained by combining all the descriptors (TiO) and the results obtained with each descriptor individually. The primary purpose is not only to verify the effectiveness of TiO but also the complementary of the descriptors in the study. The proposed method achieved the best average accuracy of 100% on the Outex dataset with LDA and CatBoost classifiers. On the ALOT dataset, the proposed method achieved an average accuracy of 98.48% with the LDA classifier. Again, TiO outperformed all individual descriptors. Finally, on the KTH-TIPS dataset, the proposed method achieved the best average accuracy of 100% with the LDA classifier and outperformed all individual descriptors.

Different works have used the Outex dataset for texture classification. For instance, Mehta & Egiazarian (2016) introduced an approach based on a rotation-invariant LBP, achieving an accuracy of 96.26% with a k-NN classifier. Du *et al.* (2016) presented a rotation-invariant, impulse noise resistant, and illumination-invariant approach based on a local spiking pattern. This approach achieved an accuracy of 86.12% with a neural network but is not extended for color textures, and it requires several input parameters. In Section 2.1, we introduced a bio-inspired texture descriptor based on biodiversity and taxonomic measures (Ataky & Lameiras Koerich, 2022). Such a texture descriptor achieved an accuracy of 99.88% with an SVM. Table 3.7 compares the best results of the proposed method with a few works that also used the Outex dataset.

Table 3.6 Average accuracy (%) on the test set of Outex, ALOT and KTH-TIPS datasets. The best accuracy for each dataset is shown in boldface. The best result for each texture descriptor is marked with *

Dataset	Texture Descriptors	Classification Algorithms			
		HistoB[1]	LightB[2]	LDA	CatBoost[3]
Outex	TiO	**100.0**	**100.0**	**100.0**	**100.0**
	BiT	99.92	99.69	100.0*	99.84
	GLCM	98.84	97.99	95.91	99.30*
	Haralick	99.69	99.53	99.38	99.84*
ALOT	TiO	41.50	43.07	**98.48**	93.47
	BiT	10.71	11.52	68.14	74.90*
	GLCM	11.28	14.52	66.61*	64.23
	Haralick	18.52	20.47	86.57*	83.23
KTH-TIPS	TiO	98.53	97.20	**100.0**	98.91
	BiT	96.29	96.70	99.58*	97.53
	GLCM	89.71	90.12	87.65	92.01*
	Haralick	95.88	96.29	98.35*	96.11

[1]max_bins=10, max_iter=1500. [2,3]n_estimators=1500.

Table 3.7 Average accuracy (%) of the proposed method with related works on Outex dataset. The best result is marked with *

Reference	Approach	Accuracy
Ataky & Lameiras Koerich (2022)	GLCM	95.52
	Haralick	96.92
	BiT	99.88
Mehta & Egiazarian (2016)	Shallow	96.23
Du *et al.* (2016)	Shallow	86.12
TiO+LDA or TiO+CatBoost	Shallow	100.0*

Table 3.8 presents a few works that also used the ALOT dataset. There was an improvement in the accuracy of nearly 1% and 0.1% compared with shallow and deep methods. Notwithstanding, the difference is not significantly high; it is worth mentioning that the success of CNNs places reliance on the ability to leverage massive labeled datasets to learn high-quality representations. Nonetheless, data availability for a few domains may be restricted, and therefore CNNs become restrained from several fields. Moreover, some works evaluated small-scale CNN architectures, such as T-CNN and T-CNN Inception, with 11,900 and 1.5M parameters. Despite the reduced number of parameters and lower computational cost for training, both still required a large amount of training data to perform satisfactorily.

Even some small architectures of comparable performance, such as MobileNets, EfficientNets, and sparse architectures resulting from pruning, still have many training parameters. For instance, the number of parameters of a MobileNet range between 3.5M and 4.2M, while EfficientNets range between 5.3M and 66.6M. GoogleNet, ResNet, and VGG generally need extensive training, and the number of hyperparameters and the computational cost is high.

Table 3.8 Accuracy (%) of shallow and deep approaches on the ALOT dataset. The best result is marked with *

Reference	**Approach**	**Accuracy**
Armi, Abbasi & Zarepour-Ahmadabadi (2021)	Shallow	97.56
Dubey, Singh & Singh (2016)	Shallow	63.04
He, Zhang, Ren & Sun (2016)	ResNet50	75.68
	ResNet101	75.60
Napoletano (2017)	ResNet101	98.13
	ResNet50	98.35
	VGG VeryDeep19	94.93
	VGG M128	85.56
	GoogleNet	92.65
	VGG M1024	92.58
	VGG M2048	93.30
Alpaslan & Hanbay (2020)	Shallow	97.20
TiO+LDA	Shallow	98.48*

Likewise, the KTH-TIPS dataset has been used to evaluate texture characterization and classification approaches. Hazgui *et al.* (2021) introduced an approach that integrates the genetic programming and the fusion of HOG and LBP features, which achieved an accuracy of 91.20% with a k-NN classifier. Such an approach does not use color information and global features. Nguyen *et al.* (2016) put forth rotational and noise invariant statistical binary patterns, which reached an accuracy of 97.73%, which is lower than the accuracy achieved by the proposed method of about 2.3%. This approach is resolution sensitive and presents a high computational complexity. Qi *et al.* (2013) proposed a rotation-invariant multiscale cross-channel LBP (CCLBP) that encodes the cross-channel texture correlation. The CCLBP computes the LBP descriptors in each channel and three scales and computes co-occurrence statistics before concatenating the extracted features. Such an approach achieved an accuracy of 99.01% for three scales with an SVM. Nevertheless, this method is not invariant to scale. Table 3.9 shows that the proposed approach outperforms other works that also used the KTH-TIPS.

Table 3.9 Average accuracy (%) of the proposed method with related works on KTH-TIPS dataset. The best result is marked with *

Reference	**Approach**	**Accuracy**
Ataky & Lameiras Koerich (2022)	GLCM	86.83
	Haralick	94.89
	BiT	97.87
Hazgui *et al.* (2021)	Shallow	91.20
Nguyen *et al.* (2016)	Shallow	97.73
Qi *et al.* (2013)	Shallow	99.01
TiO+LDA	Shallow	100.0*

3.2.2.2 Experiments with HI Datasets

Table 3.10 presents the accuracy achieved by monolithic classifiers and ensemble methods, both trained with the proposed method on the CRC dataset with TiO and BiT, GLCM, and Haralick, individually. The proposed approach provided its best accuracy of 94.71%, with HistoB and LightB with all the features. Into the bargain, the accuracy difference between TiO and the first best-related work is nearly 2.00%, which corroborates the discriminating ability of the proposed method. Additionally, compared to each descriptor that TiO is made up of, TiO outperformed each of them when employed individually.

Table 3.11 compares the results of the proposed approach with some state-of-the-art works to assess its effectiveness. The results achieved by TiO on the CRC dataset have shown that the proposed approach works well on images with other structures apart from textures and with no need for data augmentation. Besides, such CNNs need to be trained with a large amount of data which may be prohibitive in fields such as medical imaging.

Table 3.12 shows the accuracy achieved by monolithic classifiers and ensemble methods trained with all the features of the proposed approach. For the BreakHis dataset, the HistoB classifier achieved the best accuracy for 40×, 100×, and 400× magnifications, and LightB for 200×. Table 3.13 shows and compares the results achieved by the proposed approach using all the features with the state-of-the-art for the BreakHis dataset. One can note that the proposed approach achieved a considerable accuracy of 98.64% with all the features for 40× magnification,

which slightly outperforms the accuracy of both shallow and deep methods. The difference in accuracy between the proposed approach and the second-best method (CNN) is nearly 1% for 40× magnification. Likewise, the proposed method achieved a considerable accuracy of 97.85%, 98.76% and 98.22% for 100×, 200×, and 400×, respectively, which slightly outperformed the second-best method with difference of 0.9% for 100×, 1.05% for 200×, and 1.0% for 400×.

We also carried out the experiments individually with BiT, GLCM, and Haralick. For all the magnifications, TiO outperformed each descriptor with the maximum accuracy difference of 1.25%, 0.48%, 1.01%, 1.5% for 40×, 100×, 200×, and 400×, respectively. Thus, the combination of the descriptors mentioned above increased nearly 1% in accuracy.

Table 3.10 Accuracy (%) of monolithic classifiers and ensemble methods with TiO and each descriptor employed individually on the CRC dataset

Texture Descriptors	**Classification Algorithms**			
	HistoB	**LightB**	**LDA**	**CatBoost**
TiO	**94.71**	**94.71**	94.37	93.80
BiT	93.22	92.61	86.67	92.92
GLCM	85.98	86.97	81.79	87.20
Haralick	91.62	91.31	92.38	91.39

Table 3.11 Average accuracy (%) of shallow and deep approaches on the CRC dataset. The best results are marked with *

Reference	**Approach**	**10-fold**	**5-fold**
Kather *et al.* (2016b)	Shallow	87.40	–
Sarkar & Acton (2017)	Shallow	73.60	–
Ataky & Lameiras Koerich (2022)	Shallow	92.96	–
TiO+HistoB	Shallow	94.71*	92.77*
Wang *et al.* (2017)	CNN	–	92.60
Pham (2017)	CNN	–	84.00
Raczkowski *et al.* (2019)	CNN	92.40	92.20

Table 3.12 Accuracy (%) of monolithic classifiers and ensemble methods with TiO descriptor on balanced 8-classes image-level BreakHis dataset. The best result for each magnification is marked with *

Magnification	Classification Algorithms			
	HistoB	**LightB**	**LDA**	**CatBoost**
40×	98.64*	98.12	97.83	98.04
100×	97.85*	97.50	96.81	96.67
200×	97.62	98.76*	94.63	97.28
400×	98.22*	98.06	94.92	96.98

Table 3.13 Accuracy (%) of shallow and deep approaches on the BreakHis dataset. All these works used the same data partitions for training and test. The best results are indicated with *

Reference	Method	Magnification			
		40×	**100×**	**200×**	**400×**
Spanhol *et al.* (2016a) (LBP)	Shallow	75.60	73.00	72.90	71.20
Spanhol *et al.* (2016a) (GLCM)	Shallow	74.70	76.80	83.40	81.70
Erfankhah *et al.* (2019)	Shallow	88.30	88.30	87.10	83.40
Ataky & Lameiras Koerich (2022)	Shallow	97.50	96.80	95.80	95.20
Alom *et al.* (2019)	CNN	97.00	97.50	97.20	97.20
Han *et al.* (2017)	CNN	92.80	93.90	93.70	92.90
Bayramoglu *et al.* (2016)	CNN	83.00	83.10	84.60	82.10
Spanhol *et al.* (2016b)	CNN	90.00	88.40	84.60	86.10
TiO+HistoB	Shallow	98.64*	97.85*	98.76*	98.22*

3.2.3 Final Considerations

This section provided an essential study regarding image analysis and manipulation over a spectrum of spatial scales and the complementarity of feature descriptors for texture classification. We stated that individually employing each descriptor may overlook relevant textural information, reducing the classification performance. Moreover, we exploited the pyramids' multi-resolution representation as a proper data structure for analyzing and capturing intrinsic details from texture over a spectrum of spatial scales. To produce a feature vector from an image, we combined several descriptors that have proven to be discriminating for the classification, namely, the BiT, information-theoretic measures, GLCM, and Haralick descriptors to extract gray-level and color features in different resolutions. Such a combination aimed to bring features that characterize

the texture to the maximum extent and with some advantages such as rotation, permutation, scale-invariant, reduced noise sensitivity, and generic and high generalization ability, as it has provided effective performance for the real-world datasets.

The proposed approach outperformed state-of-the-art shallow and deep methods. Moreover, the descriptors employed herein have proven complementary as their combination resulted in a better performance than using each individually. Regardless, some features may be redundant after the concatenation into a single feature vector, given the different resolutions, channels, and descriptors. This may cause a downfall in the classification performance. Likewise, such a concatenation may also lead to the Hughes phenomenon. This also explains why we have not included other descriptors such as LBP, HOG, etc. However, we will consider including other descriptors in future studies.

To circumvent the possible feature redundancy and cope with the increase in dimensionality, in future works, we intend to investigate the impact of incorporating a decision marker multi-objective feature selection.

CONCLUSION AND RECOMMENDATIONS

Texture analysis in images is fundamental in computer vision. Several applications in remote sensing, medicine, agriculture, image analysis, microscopy, etc., require the study of the surface properties of an object, the understanding of how humans discriminate between different textures for, afterward, modeling techniques that can perform this task, by quantifying a suitable similarity measure between different objects in a dataset.

In addition to natural images, texture analysis is a task of high complexity and significant relevance for detecting and diagnosing tumors and cancerous cells since, in histopathological images, for example, the cells often have very similar textures, even though they belong to different classes, which can cause errors in classification. Because detecting tumors with a high sensitivity rate and reducing the false positives is still challenging, texture descriptors have been quite popular in medical image analysis, particularly in histopathologic images (HI), due to the variability of the texture found in such images and the tissue appearance caused by irregularity ascribable to the staining process. Such variability may exist depending on differences in staining protocol such as fixation, inconsistency in the staining condition, and reagents, either between laboratories or in the same laboratory. Moreover, such images contain not solely texture, for most of them present shape, making them challenging to characterize with only texture. Into the bargain, texture feature extraction for quantifying HI information in a discriminant way is challenging, given that the distribution of intrinsic properties of such images forms a non-deterministic complex system.

In summary, this thesis presented methods for classifying natural and medical images, such as histopathological images, through image processing and pattern recognition techniques, leveraging texture properties. In this context, the theoretical foundation of the subjects related to the area of study was presented, which are necessary for understanding the work developed

and the techniques addressed in the methods. Also, some related works are presented in the characterization and classification of texture images and those that do not have only texture.

Experiments with natural image datasets (Salzburg, Outex, KTH-TIPS, and ALOT) and histopathological image datasets (CRC and BreakHis) were used to evaluate the proposed approaches. At this point, the adaptation and application of biological concepts, such as biodiversity measures and taxonomy indices, to the computing area enabled the creation of a new descriptor for texture image characterization and classification. As a result, this is the main contribution of the present doctoral research.

The promising performance in the classification task of such images through ecological modeling and mathematical representation of an image demonstrated that the proposed approach is valid. It produces statistically significant results and is competitive with state-of-the-art works in the field.

Overall, the contributions of this thesis can be summarized as follows:

We developed methods for image classification leveraging texture analysis, in particular:

1. We adapted and developed strategies that enable analysis with as much relevant information as possible;
2. We leveraged ecological concepts for texture characterization for both natural and histopathologic images by using:
   - Biodiversity indices based on species richness and abundance;
   - Diversity indices based on species evenness;
   - Diversity indices based on evolutionary relationships between pairs of species;
   - Diversity indices based on phylogenetic distance; and
   - Ecological patterns' invariance characteristics to permutation, rotation, and scale.

3. We proposed an image representation scheme that allows the extraction of mathematical properties therefrom through ecological concepts;
4. We developed several descriptors that are generic, independent of macro-level variations in terms of contrast, invariant to in-plane rotations of the image, lend themselves to fast computation, explainable and interpretable based on biology concepts;

In general, the methods outlined in this thesis could form a reliable instrument for application in real-world classification and diagnosis scenarios. It is also essential to look into their sturdiness when they are put to the evaluation in databases with enormous amounts of data and cases that require complex processing.

Moreover, it is noteworthy that the methods proposed in this thesis have limitations in that the dynamic ranges of the biodiversity and taxonomic indices are not specified beforehand. Feature normalization may be required as the biodiversity, and taxonomic indices range depend on the species richness, abundance, and relationship within an image. In addition, images are typically corrupted by noise, which introduces random variations to grey levels, causing homogeneous regions to display a gray-level dispersion proportional to the noise power, thereby reducing performance. Although the proposed descriptor performed well in the presence of noise, it does not provide a tolerance parameter to deal with such noise-induced variations that aid in determining whether a species (gray-level) or individual (pixel) should be considered or not, suggesting that preprocessing be performed prior to the feature extraction phase.

In addition to the scientific contributions, we have also developed several open-source python packages that implement all the methods proposed in this thesis, which may be used by everyone considering incorporating the proposed descriptors into their applications (Ataky & L. Koerich, 2022).

4.1 Future Works

As presented in this thesis, the interest in computational research that assists in the characterization of texture in various areas of computational vision is extremely relevant, given a large amount of existing research. Therefore, for future works, some suggestions were listed:

- In real-world applications, images are usually corrupted with noise introducing random variations to gray levels so that homogeneous regions present a gray-level dispersion depending on noise power. Accordingly, one will study a tolerance parameter to cope with such noise-produced variations that help to decide whether a species or a region should be taken into consideration or not before computing ecological diversity indices.
- Alternative levels of Gaussian pyramids could be used to improve the performance of the suggested descriptor.
- The exploitation of fractal features for dealing with multi-resolution aspects of the textures that could bring on discriminative information for characterizing texture images.
- The exploitation of the curvelet transforms (an extension of the wavelet concept) for a multi-scale directional transform that allows an almost optimal nonadaptive sparse representation of objects with edges, in which case it could be possible to employ ecological diversity measures to characterize images containing the most edges.
- Because 1-D transform processing of intrinsic geometrical structures, such as curve smoothness, is limited in one direction, more powerful representations in higher dimensions are necessary. Thus, the contourlet transform stands as an interesting two-dimensional transform method for image representations. Multiresolution, localization, directionality, critical sampling, and anisotropy are all features of the contourlet transform. As such, we will study the contourlet transforms to effectively capture the contours of original images, for example, cells contours, which are the significant features in both natural and histopathologic images, using only a few coefficients.

- Investigation of how to incorporate a decision-maker to select relevant features and scales and resolutions from wavelet decompositions and curvelets that provide the most relevant intrinsic properties for texture characterization.
- Considering an image as an ecosystem, there will be many different patches of gray-level x and gray-level y. Alpha diversity is the species diversity within each gray-level x or gray-level y patch of the image. Beta diversity is represented by the species diversity between any two patches and their communities. Gamma diversity of the images is the species diversity along with the entire range of the dataset or class to which the image belongs. Thus, we aim to leverage such concepts to develop an ecological representation-based descriptor-classifier.

4.2 Publications

The current doctoral project resulted in the publication/submission of several scholarly articles in related conferences and journals. The primarily published works, as well as those submitted for publication, are listed in Table 4.1.

Table 4.1 Published and submitted articles related to this research proposal

Place	Paper	Status
Journal	Ataky, STM., Lameiras Koerich, A. A novel bio-inspired texture descriptor based on biodiversity and taxonomic measures. Pattern Recognition. Volume 123, March 2022, 108382	Published
Journal	Ataky, STM., Lameiras Koerich, A. E-BiT: Extended bio-inspired texture descriptor for texture analysis and characterization. Image and Vision Computing.	Submitted
Journal	Ataky, STM., Lameiras Koerich, A. Multiresolution Texture Analysis of Histopathologic Images Using Ecological Diversity Measures. Expert Systems With Applications.	Minor review
Journal	Ataky, STM., Saqui,D., Lameiras Koerich, A. Improving Texture Classification by Multi-objective Feature Selection. Expert Systems With Applications.	Under review
Conference	Ataky, STM.,Lameiras Koerich, A. Multiscale Analysis for Improving Texture Classification. Applied Science (Journal)	Submitted

Table 4.2 Published articles not directly related to this research proposal

Place	Paper	Status
Conference	STM Ataky, J de Matos, AS Britto, LES Oliveira, AL Koerich Data Augmentation for Histopathological Images Based on Gaussian-Laplacian Pyramid Blending. IEEE International Joint Conference on Neural Networks, June 2020	Published
Journal	J de Matos, STM Ataky, A de Souza Britto, LE Soares de Oliveira, AL Koerich Machine Learning Methods for Histopathological Image Analysis: A Review. Electronics, 2021	Published

Table 4.2 lists the scientific articles published throughout the doctoral path, but not directly related to the objectives of the proposal herein.

REFERENCES

Adelson, E. H., Anderson, C. H., Bergen, J. R., Burt, P. J. & Ogden, J. M. (1984). Pyramid methods in image processing. *RCA engineer*, 29(6), 33–41.

Alom, M. Z., Yakopcic, C., Nasrin, M. S., Taha, T. M. & Asari, V. K. (2019). Breast cancer classification from histopathological images with inception recurrent residual convolutional neural network. *Journal of digital imaging*, 32(4), 605–617.

Alpaslan, N. & Hanbay, K. (2020). Multi-resolution intrinsic texture geometry-based local binary pattern for texture classification. *IEEE Access*, 8, 54415–54430.

Anderson, C. H., Burt, P. J. & Van Der Wal, G. S. (1985). Change detection and tracking using pyramid transform techniques. *Intelligent Robots and Computer Vision IV*, 579, 72–78.

Andrearczyk, V. & Whelan, P. F. (2016). Using filter banks in Convolutional Neural Networks for texture classification. *Pattern Recognition Letters*, 84, 63-69. doi: 10.1016/j.patrec.2016.08.016.

Arivazhagan, S. & Ganesan, L. (2003). Texture classification using wavelet transform. *Pattern recognition letters*, 24(9-10), 1513–1521.

Armi, L., Abbasi, E. & Zarepour-Ahmadabadi, J. (2021). Texture images classification using improved local quinary pattern and mixture of ELM-based experts. *Neural Comp App*, 1–24.

Ataky, S. T. M. & Lameiras Koerich, A. (2022). A novel bio-inspired texture descriptor based on biodiversity and taxonomic measures. *Patt Recogn*, 123, 108382. doi: https://doi.org/10.1016/j.patcog.2021.108382.

Ataky, S. T. M., de Matos, J., de Souza Britto Jr., A., Oliveira, L. E. S. & Koerich, A. L. (2020). Data Augmentation for Histopathological Images Based on Gaussian-Laplacian Pyramid Blending. *Intl Joint Conf on Neural Networks, (IJCNN), Glasgow, UK*, pp. 1–8.

Ataky, S. & L. Koerich, A. (2022). Bio-Inspired texture descriptors. GitHub.

Ayed, N. G. B., Larousi, M. G. & Masmoudi, A. D. (2014). Multi-scale Gray Level and Local Difference for texture classification. *2014 World Congress on Computer Applications and Information Systems (WCCAIS)*, pp. 1–5.

Backes, A. R., Casanova, D. & Bruno, O. M. (2013). Texture analysis and classification: A complex network-based approach. *Information Sciences*, 219, 168–180.

Bardou, D., Zhang, K. & Ahmad, S. M. (2018). Classification of breast cancer based on histology images using convolutional neural networks. *Ieee Access*, 6, 24680–24693.

Barrow, D. K. & Crone, S. F. (2016). Cross-validation aggregation for combining autoregressive neural network forecasts. *International Journal of Forecasting*, 32(4), 1120–1137.

Bashier, H. K., Hoe, L. S., Hui, L. T., Azli, M. F. & Han, P. Y. (2016). Texture classification via extended local graph structure. *Optik*, 127(2), 638–643.

Bayramoglu, N., Kannala, J. & Heikkilä, J. (2016). Deep learning for magnification independent breast cancer histopathology image classification. *23rd Intl. Conf. on pattern recognition (ICPR)*, pp. 2440–2445.

Bergen, J. R. & Julesz, B. (1983). Rapid discrimination of visual patterns. *IEEE Transactions on Systems, Man, and Cybernetics*, (5), 857–863.

Bharati, M. H., Liu, J. J. & MacGregor, J. F. (2004). Image texture analysis: methods and comparisons. *Chemometrics and intelligent laboratory systems*, 72(1), 57–71.

Blakemore, C. & Campbell, F. W. (1969). On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. *The Journal of physiology*, 203(1), 237–260.

Burt, P. & Adelson, E. (1983). The Laplacian pyramid as a compact image code. *IEEE Transactions on communications*, 31(4), 532–540.

Burt, P. J. (1981). Fast filter transform for image processing. *Computer graphics and image processing*, 16(1), 20–51.

Burt, P. J. (1983). Fast algorithms for estimating local image properties. *Computer Vision, Graphics, and Image Processing*, 21(3), 368–382.

Caicedo, J. C., González, F. A. & Romero, E. (2011). Content-based histopathology image retrieval using a kernel-based semantic annotation framework. *Journal of Biomedical Informatics*, 44(4), 519–528. doi: https://doi.org/10.1016/j.jbi.2011.01.011.

Castleman, K. R. (1996). *Digital image processing*. Pearson.

Cernadas, E., Fernández-Delgado, M., González-Rufino, E. & Carrión, P. (2017). Influence of normalization and color space to color texture classification. *Pattern Recognition*, 61, 120–138.

Chao, A. (1984). Nonparametric estimation of the number of classes in a population. *Scandinavian Journal of statistics*, 265–270.

Chao, A., Shen, T.-J. & Hwang, W.-H. (2006). Application of Laplace's boundary-mode approximations to estimate species and shared species richness. *Australian & New Zealand Journal of Statistics*, 48(2), 117–128.

Clarke, K. R. & Warwick, R. M. (1998a). A taxonomic distinctness index and its statistical properties. *Journal of applied ecology*, 35(4), 523–531.

Clarke, K. R., Gorley, R., Somerfield, P. J. & Warwick, R. (2014). *Change in marine communities: an approach to statistical analysis and interpretation*. Primer-E Ltd.

Clarke, K. & Warwick, R. (1998b). A taxonomic distinctness index and its statistical properties. *Journal of applied ecology*, 35(4), 523–531.

Clifford, H. T., Stephenson, W. et al. (1975). *An introduction to numerical classification*. Academic Press New York.

Cox, M. A. & Cox, T. F. (2008). Multidimensional scaling. In *Handbook of data visualization* (pp. 315–347). Springer.

Crimmins, T. R. (1985). Geometric filter for speckle reduction. *Applied optics*, 24(10), 1438–1443.

Cross, G. R. & Jain, A. K. (1983). Markov random field texture models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (1), 25–39.

Crowley, J. L. (1981). *A Representation for Visual Information.*

Crowley, J. L. & Riff, O. (2003). Fast computation of scale normalised gaussian receptive fields. *International Conference on Scale-Space Theories in Computer Vision*, pp. 584–598.

Cruz-Roa, A., Caicedo, J. C. & González, F. A. (2011). Visual pattern mining in histology image collections using bag of features. *Artificial Intelligence in Medicine*, 52(2), 91–106. doi: https://doi.org/10.1016/j.artmed.2011.04.010.

Da Silva, I. A. & Batalha, M. A. (2006). Taxonomic distinctness and diversity of a hyperseasonal savanna in central Brazil. *Diversity and Distributions*, 12(6), 725–730.

Daly, A. J., Baetens, J. M. & De Baets, B. (2018). Ecological diversity: measuring the unmeasurable. *Mathematics*, 6(7), 119.

Das, D. K., Mitra, P., Chakraborty, C., Chatterjee, S., Maiti, A. K. & Bose, S. (2017). Computational approach for mitotic cell detection and its application in oral squamous cell carcinoma. *Multidimensional Systems and Signal Processing*, 28(3, SI), 1031–1050. doi: 10.1007/s11045-017-0488-6.

de Matos, J., de Souza Britto Jr., A., de Oliveira, L. E. S. & Koerich, A. L. (2019). Texture CNN for Histopathological Image Classification. *32nd IEEE Int'l Symp on Computer-Based Medical Systems*, pp. 580–583. doi: 10.1109/CBMS.2019.00120.

de Matos, J., Ataky, S. T. M., de Souza Britto, A., Soares de Oliveira, L. E. & Lameiras Koerich, A. (2021a). Machine learning methods for histopathological image analysis: A review. *Electronics*, 10(5), 562.

de Matos, J., Ataky, S. T. M., de Souza Britto, A., Soares de Oliveira, L. E. & Lameiras Koerich, A. (2021b). Machine Learning Methods for Histopathological Image Analysis: A Review. *Electronics*, 10(5). doi: 10.3390/electronics10050562.

Dong, K., Hou, G., Xu, D., He, H. & Liu, Z. (2018). A Method to Compare the Biodiversity Conservation Effectiveness between Regions based on a Reference Condition. *Sustainability*, 10(10), 3694.

Dougherty, G. (2009). *Digital image processing for medical applications*. Cambridge University Press.

Du, S., Yan, Y. & Ma, Y. (2016). Local spiking pattern and its application to rotation-and illumination-invariant texture classification. *Optik*, 127(16), 6583–6589.

Dubey, S. R., Singh, S. K. & Singh, R. K. (2016). Multichannel decoded local binary patterns for content-based image retrieval. *IEEE Trans Image Process*, 25(9), 4018–4032.

Dunteman, G. H. (1989). *Principal components analysis*. Sage.

Eren, M. I., Chao, A., Hwang, W.-H. & Colwell, R. K. (2012). Estimating the richness of a population when the maximum number of classes is fixed: a nonparametric solution to an archaeological problem. *PLoS One*, 7(5), e34179.

Erfankhah, H., Yazdi, M., Babaie, M. & Tizhoosh, H. R. (2019). Heterogeneity-aware local binary patterns for retrieval of histopathology images. *IEEE Access*, 7, 18354–18367.

Faith, D. P. (1992). Conservation evaluation and phylogenetic diversity. *Biological conservation*, 61(1), 1–10.

Faith, D. P. (1994). Phylogenetic pattern and the quantification of organismal biodiversity. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 345(1311), 45–58.

Feng, J., Jiao, L., Zhang, X. & Sun, T. (2013). Hyperspectral band selection based on trivariate mutual information and clonal selection. *IEEE Transactions on Geoscience and Remote Sensing*, 52(7), 4092–4105.

Feng, J., Jiao, L., Liu, F., Sun, T. & Zhang, X. (2014). Mutual-information-based semi-supervised hyperspectral band selection with high discrimination, high information, and low redundancy. *IEEE Transactions on Geoscience and Remote Sensing*, 53(5), 2956–2969.

Feng, J., Jiao, L., Liu, F., Sun, T. & Zhang, X. (2016). Unsupervised feature selection based on maximum information and minimum redundancy for hyperspectral images. *Pattern Recognition*, 51, 295–309.

Fernández-Carrobles, M. M., Bueno, G., Déniz, O., Salido, J., García-Rojo, M. & Gonzández-López, L. (2015). Frequential versus spatial colour textons for breast {TMA} classification. *Computerized Medical Imaging and Graphics*, 42, 25-37. doi: https://doi.org/10.1016/j.compmedimag.2014.11.009.

Fisher, R. A., Corbet, A. S. & Williams, C. B. (1943a). The relation between the number of species and the number of individuals in a random sample of an animal population. *The Journal of Animal Ecology*, 42–58.

Fisher, R. A., Corbet, A. S. & Williams, C. B. (1943b). The relation between the number of species and the number of individuals in a random sample of an animal population. *The Journal of Animal Ecology*, 42–58.

Fogel, I. & Sagi, D. (1989). Gabor Filters as Texture Discriminator. *Biological Cybernetics*, 61, 103–113.

Frank, S. A. & Bascompte, J. (2019). Invariance in ecological pattern. *F1000Research*, 8.

Fujieda, S., Takayama, K. & Hachisuka, T. (2017). Wavelet convolutional neural networks for texture classification. *arXiv preprint arXiv:1707.07394*.

Gandomkar, Z., Brennan, P. C. & Mello-Thoms, C. (2018). MuDeRN: Multi-category classification of breast histopathological image using deep residual networks. *Artif Intell Med*, 88, 14–24.

Ghalati, M. K., Nunes, A., Ferreira, H., Serranho, P. & Bernardes, R. (2021). Texture analysis and its applications in biomedical imaging: A survey. *IEEE Reviews in Biomedical Engineering*.

Gibson, R., Barnes, M. & Atkinson, R. (2001). Practical measures of marine biodiversity based on relatedness of species. *Oceanography and Marine Biology*, 39, 207–231.

Gilpin, L. H., Bau, D., Yuan, B. Z., Bajwa, A., Specter, M. & Kagal, L. (2018). Explaining explanations: An overview of interpretability of machine learning. *2018 IEEE 5th International Conference on data science and advanced analytics (DSAA)*, pp. 80–89.

Gini, C. (1912). Variabilità e mutabilità. *Reprinted in Memorie di metodologica statistica (Ed. Pizetti E.*

Gonzalez, R. C. & Woods, R. C. (2009). *Processamento digital de imagens* . Pearson Educación.

Goutsias, J. & Heijmans, H. (1998). Multiresolution signal decomposition schemes. Part 1: Linear and morphological pyramids. *CWI (Centre for Mathematics and Computer Science) Amsterdam, The Netherlands, Tech. Rep*.

Guo, Z., Zhang, L. & Zhang, D. (2010). A completed modeling of local binary pattern operator for texture classification. *IEEE transactions on image processing*, 19(6), 1657–1663.

Habermann, M. (2018). *Band selection in hyperspectral images using artificial neural networks*. (Ph.D. thesis, Compiègne).

Habermann, M., Fremont, V. & Shiguemori, E. H. (2019). Supervised band selection in hyperspectral images using single-layer neural networks. *International journal of remote sensing*, 40(10), 3900–3926.

Hafemann, L. G., Oliveira, L. S. & Cavalin, P. (2014). Forest species recognition using deep convolutional neural networks. *2014 22nd International Conference on Pattern Recognition*, pp. 1103–1107.

Han, Z., Wei, B., Zheng, Y., Yin, Y., Li, K. & Li, S. (2017). Breast cancer multi-classification from histopathological images with structured deep learning model. *Scientific reports*, 7(1), 1–10.

Hao, H., Wang, Q., Li, P. & Zhang, L. (2016). Evaluation of ground distances and features in EMD-based GMM matching for texture classification. *Pattern Recognition*, 57, 152–163.

Haralick, R. M. (1979). Statistical and structural approaches to texture. *Proceedings of the IEEE*, 67(5), 786–804.

Haralick, R. M., Shanmugam, K. & Dinstein, I. H. (1973). Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics*, (6), 610–621.

Hazgui, M., Ghazouani, H. & Barhoumi, W. (2021). Genetic programming-based fusion of HOG and LBP features for fully automated texture classification. *The Visual Computer*, 37, 1–20. doi: 10.1007/s00371-020-02028-8.

He, K., Zhang, X., Ren, S. & Sun, J. (2016). Deep residual learning for image recognition. *IEEE Conf Comp Vis Patt Recognit*, pp. 770–778.

Heip, C. & Engels, P. (1974). Comparing species diversity and evenness indices. *Journal of the Marine Biological Association of the United Kingdom*, 54(3), 559–563.

I, I. F. & Sagi, D. (1989). Biological Cybernetics O Springer-Verlag 1989 Gabor Filters as Texture Discriminator.

Izsák, J. & Papp, L. (2000). A link between ecological diversity indices and measures of biodiversity. *Ecological Modelling*, 130(1-3), 151–156.

Izsáki, J. & Papp, L. (1995). Application of the quadratic entropy indices for diversity studies of drosophilid assemblages. *Environmental and Ecological Statistics*, 2(3), 213–224.

Jain, A. K. & Farrokhnia, F. (1990). Unsupervised texture segmentation using Gabor filters. *1990 IEEE international conference on systems, man, and cybernetics conference proceedings*, pp. 14–19.

Jain, A. K., Duin, R. P. W. & Mao, J. (2000). Statistical pattern recognition: A review. *IEEE Transactions on pattern analysis and machine intelligence*, 22(1), 4–37.

Jolion, J.-M. & Rosenfeld, A. (2012). *A pyramid framework for early vision: multiresolutional computer vision*. Springer Science & Business Media.

Jost, L. (2007). Partitioning diversity into independent alpha and beta components. *Ecology*, 88(10), 2427–2439.

Julesz, B. (1975). Experiments in the visual perception of texture. *Scientific American*, 232(4), 34–43.

Julesz, B. (1981). Textons, the elements of texture perception, and their interactions. *Nature*, 290(5802), 91–97.

Junior, J. J. d. M. S. & Backes, A. R. (2016). ELM based signature for texture classification. *Pattern Recognition*, 51, 395–401.

Kalkan, H., Nap, M., Duin, R. P. & Loog, M. (2012). Automated classification of local patches in colon histopathology. *21st Int'l Conf Patt Recogn (ICPR2012)*, pp. 61–64.

Kannala, J. & Rahtu, E. (2012). Bsif: Binarized statistical image features. *21st Intl. Conf. on pattern recognition (ICPR2012)*, pp. 1363–1366.

Kaplan, L. M. (1999). Extended fractal analysis for texture classification and segmentation. *IEEE Transactions on Image Processing*, 8(11), 1572–1585.

Karu, K., Jain, A. K. & Bolle, R. M. (1996). Is there any texture in the image? *Pattern Recognition*, 29(9), 1437–1446.

Kather, J. N., Weis, C.-A., Bianconi, F., Melchers, S. M., Schad, L. R., Gaiser, T., Marx, A. & Zöllner, F. G. (2016a). Multi-class texture analysis in colorectal cancer histology. *Scientific Reports*, 6(1), 27988. doi: 10.1038/srep27988.

Kather, J. N., Weis, C.-A., Bianconi, F., Melchers, S. M., Schad, L. R., Gaiser, T., Marx, A. & Zöllner, F. G. (2016b). Multi-class texture analysis in colorectal cancer histology. *Scientific reports*, 6, 27988.

Kempton, R. & Taylor, L. (1976). Models and statistics for species diversity. *Nature*, 262(5571), 818–820.

Kohavi, R., John, G. H. et al. (1997). Wrappers for feature subset selection. *Artificial intelligence*, 97(1-2), 273–324.

Kumar, V. & Minz, S. (2014). Feature selection: a literature review. *SmartCR*, 4(3), 211–229.

Kuse, M., Sharma, T. & Gupta, S. (2010). A classification scheme for lymphocyte segmentation in H&E stained histology images. *Lecture Notes in Computer Science*, 6388 LNCS, 235–243. doi: 10.1007/978-3-642-17711-8_24.

Larkin, L. I. & Burt, P. J. (1983). Multi-resolution texture energy measures. *Proceedings in IEEE conference on computer vision and pattern recognition*.

Leo, P., Lee, G., Shih, N. N. C., Elliott, R., Feldman, M. D. & Madabhushi, A. (2016). Evaluating stability of histomorphometric features across scanner and staining variations: prostate cancer diagnosis from whole slide images. *Journal of Medical Imaging*, 3(4). doi: 10.1117/1.JMI.3.4.047502.

Levine, M. D. (1985). *Vision in man and machine*. McGraw-Hill College.

Lindeberg, T. (2013). *Scale-space theory in computer vision*. Springer Science & Business Media.

Liu, L. & Fieguth, P. (2012). Texture classification from random features. *IEEE transactions on pattern analysis and machine intelligence*, 34(3), 574–586.

Liu, L., Chen, J., Fieguth, P., Zhao, G., Chellappa, R. & Pietikäinen, M. (2019). From BoW to CNN: Two decades of texture representation for texture classification. *International Journal of Computer Vision*, 127(1), 74–109.

Loew, M. H. (2000). Feature Extraction, chapter 5. *SPIE, Belligham, WA, m. sonka and j. michael fitzpatrick edition*.

Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2), 91–110.

Mäenpää, T. & Pietikäinen, M. (2004). Classification with color and texture: jointly or separately? *Pattern Recognition*, 37(8), 1629–1640.

Magurran, A. (2004a). Measuring Biological Diversity Blackwell Publishing. *Malden, MA*.

Magurran, A. E. (1988). A variety of diversities. In *Ecological Diversity and Its Measurement* (pp. 81–99). Dordrecht: Springer Netherlands. doi: 10.1007/978-94-015-7358-0_5.

Magurran, A. E. (2004b). *Measuring biological diversity*. John Wiley & Sons.

Magurran, A. E. (2013). *Measuring biological diversity*. John Wiley & Sons.

Mallat, S. (1999). *A wavelet tour of signal processing*. Elsevier.

Mandelbrot, B. (1967). How long is the coast of Britain? Statistical self-similarity and fractional dimension. *science*, 156(3775), 636–638.

Masood, K. & Rajpoot, N. (2009). Texture based classification of hyperspectral colon biopsy samples using CLBP. *IEEE Int'l Symp Biom Imag: From Nano to Macro*, pp. 1011-1014. doi: 10.1109/ISBI.2009.5193226.

Materka, A., Strzelecki, M. et al. (1998). Texture analysis methods–a review. *Technical university of lodz, institute of electronics, COST B11 report, Brussels*, 10(1.97), 4968.

May, R., Cody, M. & Diamond, J. (1975). Ecology of Species and Communities.

McIntosh, R. P. (1967). An index of diversity and the relation of certain concepts to diversity. *Ecology*, 48(3), 392–404.

Mehta, R. & Egiazarian, K. (2014). Texture classification using dense micro-block difference (DMD). *Asian Conference on Computer Vision*, pp. 643–658.

Mehta, R. & Egiazarian, K. (2016). Dominant rotated local binary patterns (DRLBP) for texture classification. *Pattern Recognition Letters*, 71, 16–22.

Mitchell, T. M. (1982). Generalization as search. *Artificial intelligence*, 18(2), 203–226.

Molina, L. C., Belanche, L. & Nebot, À. (2002). Feature selection algorithms: A survey and experimental evaluation. *2002 IEEE International Conference on Data Mining, 2002. Proceedings.*, pp. 306–313.

Morris, E. K., Caruso, T., Buscot, F., Fischer, M., Hancock, C., Maier, T. S., Meiners, T., Müller, C., Obermaier, E., Prati, D. et al. (2014). Choosing and using diversity indices: insights for ecological applications from the German Biodiversity Exploratories. *Ecology and evolution*, 4(18), 3514–3524.

Nadler, M. & Smith, E. P. (1993). *Pattern recognition engineering*. Wiley-interscience.

Nagasubramanian, K., Jones, S., Sarkar, S., Singh, A. K., Singh, A. & Ganapathysubramanian, B. (2018). Hyperspectral band selection using genetic algorithm and support vector machines for early identification of charcoal rot disease in soybean stems. *Plant methods*, 14(1), 86.

Naiyar, M., Asim, Y. & Shahid, A. (2015). Automated colon cancer detection using structural and morphological features. *2015 13th Int'l Conf Frontiers of Information Technology (FIT)*, pp. 240–245.

Napoletano, P. (2017). Hand-crafted vs learned descriptors for color texture classification. *Intl Workshop Comp Color Imaging*, pp. 259–271.

Nawaz, W., Ahmed, S., Tahir, A. & Khan, H. A. (2018). Classification of breast cancer histology images using alexnet. *Int'l Conf image analysis and recognition*, pp. 869–876.

Nguyen, T. P., Vu, N.-S. & Manzanera, A. (2016). Statistical binary patterns for rotational invariant texture classification. *Neurocomputing*, 173, 1565–1577.

Nsimba, C. B. & Levada, A. (2019). An information-theoretic wavelet-based texture descriptor using Gaussian Markov random field models. *Multimedia Tools and Applications*. doi: 10.1007/s11042-019-07916-3.

Nsimba, C. B. & Levada, A. L. (2020a). Exploring Information Theory and Gaussian Markov Random Fields for Color Texture Classification. *International Conference on Image Analysis and Recognition*, pp. 130–143.

Nsimba, C. B. & Levada, A. L. (2020b). Exploring Information Theory and Gaussian Markov Random Fields for Color Texture Classification. *International Conference on Image Analysis and Recognition*, pp. 130–143.

Ojansivu, V. & Heikkilä, J. (2008). Blur insensitive texture classification using local phase quantization. *Intl. Conf. on image and signal processing*, pp. 236–243.

Oliveira, H. (2007). Análise de fourier e wavelets: sinais estacionários e não estacionários. *Editora Universitária da UFPE*.

Pavoine, S., Ollier, S. & Dufour, A.-B. (2005). Is the originality of a species measurable? *Ecology Letters*, 8(6), 579–586.

Petrou, M. M. & Sevilla, P. G. (2006). *Image processing: dealing with texture*. Wiley.

Pham, T. D. (2017). Scaling of texture in training autoencoders for classification of histological images of colorectal cancer. *International Symposium on Neural Networks*, pp. 524–532.

Pianka, E. R. (2011). *Evolutionary ecology*. Eric R. Pianka.

Pietikäinen, M., Hadid, A., Zhao, G. & Ahonen, T. (2011). *Computer vision using local binary patterns*. Springer Science & Business Media.

Polesel, A., Ramponi, G. & Mathews, V. J. (2000). Image enhancement via adaptive unsharp masking. *IEEE transactions on image processing*, 9(3), 505–510.

Qi, X., Qiao, Y., Li, C.-G. & Guo, J. (2013). Exploring cross-channel texture correlation for color texture classification.

Qiu, Q., Thompson, A. & Calderbank, R. (2015). Data Representation Using the Weyl Transform. *IEEE Transactions on Signal Processing*, 64(7), 1844–1853.

Quan, Y., Xu, Y. & Sun, Y. (2014). A distinct and compact texture descriptor. *Image and Vision Computing*, 32(4), 250–259.

Raczkowski, Ł., Możejko, M., Zambonelli, J. & Szczurek, E. (2019). ARA: accurate, reliable and active histopathological image classification framework with Bayesian deep learning. *Scientific reports*, 9(1), 1–12.

Rao, A. R. (2012). *A taxonomy for texture description and identification*. Springer Science & Business Media.

Rao, C. R. (1982). Diversity and dissimilarity coefficients: a unified approach. *Theoretical population biology*, 21(1), 24–43.

Rathore, S., Iftikhar, M. A., Hussain, M. & Jalil, A. (2013). Classification of colon biopsy images based on novel structural features. *2013 IEEE 9th Int'l Conf Emerging Technologies (ICET)*, pp. 1–6.

Reis, S., Gazinska, P., Hipwell, J. H., Mertzanidou, T., Naidoo, K., Williams, N., Pinder, S. & Hawkes, D. J. (2017). Automated Classification of Breast Cancer Stroma Maturity from Histological Images. *IEEE Transactions on Biomedical Engineering*, 64(10), 2344–2352. doi: 10.1109/TBME.2017.2665602.

Remeseiro, B. & Bolon-Canedo, V. (2019). A review of feature selection methods in medical applications. *Computers in biology and medicine*, 112, 103375.

Ribeiro, M. G., Neves, L. A., do Nascimento, M. Z., Roberto, G. F., Martins, A. S. & Tosta, T. A. A. (2019). Classification of colorectal cancer based on the association of multidimensional and multiresolution features. *Expert Systems with Applications*, 120, 262–278.

Ricotta, C. (2002). Bridging the gap between ecological diversity indices and measures of biodiversity with Shannon's entropy: comment to Izsák and Papp. *Ecological Modelling*, 152(1), 1–3.

Ricotta, C. (2004). A parametric diversity measure combining the relative abundances and taxonomic distinctiveness of species. *Diversity and Distributions*, 10(2), 143–146.

Rogers, S. I., Clarke, K. R. & Reynolds, J. D. (1999). The taxonomic distinctness of coastal bottom-dwelling fish communities of the North-east Atlantic. *Journal of Animal Ecology*, 68(4), 769–782.

Rousseau, R., Van Hecke, P., NIjssen, D. & Bogaert, J. (1999a). The relationship between diversity profiles, evenness and species richness based on partial ordering. *Environmental and Ecological Statistics*, 6(2), 211–223.

Rousseau, R., Van Hecke, P., NIjssen, D. & Bogaert, J. (1999b). The relationship between diversity profiles, evenness and species richness based on partial ordering. *Environmental and Ecological Statistics*, 6(2), 211–223.

Saqui, D., Saito, J. H., De Lima, D. C., Cura, L. M. D. V. & Ataky, S. T. M. (2019). Incorporated Decision-maker-based Multiobjective Band Selection for Pixel Classification of Hyperspectral Images. *Advances in Electrical and Computer Engineering*, 19(4), 21–28.

Sarkar, R. & Acton, S. T. (2017). Sdl: Saliency-based dictionary learning framework for image similarity. *IEEE Transactions on Image Processing*, 27(2), 749–763.

Saxena, J., Teckchandani, K., Pandey, P., Dutta, M. K., Travieso, C. M., Alonso-Hernández, J. B. et al. (2015). Palm vein recognition using local tetra patterns. *2015 4th International Work Conference on Bioinspired Intelligence (IWOBI)*, pp. 151–156.

SDR-IV. (2020). Species Diversity and Richness 4.

Shahana, A. & Preeja, V. (2016). Survey on feature subset selection for high dimensional data. *2016 International Conference on Circuit, Power and Computing Technologies (ICCPCT)*, pp. 1–4.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell system technical journal*, 27(3), 379–423.

Shimadzu, H., Dornelas, M., Henderson, P. A. & Magurran, A. E. (2013). Diversity is maintained by seasonal variation in species abundance. *BMC Biology*, 11(1), 98.

Simon, P. & Uma, V. (2018a). Review of texture descriptors for texture classification. In *Data Engineering and Intelligent Computing* (pp. 159–176). Springer.

Simon, P. & Uma, V. (2018b). Review of texture descriptors for texture classification. In *Data Engineering and Intelligent Computing* (pp. 159–176). Springer.

Singh, A., Sunkaria, R. K. & Kaur, A. (2022). A Review on Local Binary Pattern Variants. *Proceedings of First International Conference on Computational Electronics for Wireless Communications*, pp. 545–552.

Sohier, C. (2019). Measurements of biodiversity.

Solow, A., Polasky, S. & Broadus, J. (1993). On the measurement of biological diversity. *Journal of Environmental Economics and Management*, 24(1), 60–68.

Solow, A. R. & Polasky, S. (1994). Measuring biological diversity. *Environmental and Ecological Statistics*, 1(2), 95–103.

Sonka, M., Hlavac, V. & Boyle, R. (2014). *Image processing, analysis, and machine vision*. Cengage Learning.

Spanhol, F. A., Oliveira, L. S., Petitjean, C. & Heutte, L. (2016a). A Dataset for Breast Cancer Histopathological Image Classification. *IEEE Transactions on Biomedical Engineering*, 63(7), 1455–1462. doi: 10.1109/TBME.2015.2496264.

Spanhol, F. A., Oliveira, L. S., Petitjean, C. & Heutte, L. (2016b). Breast cancer histopathological image classification using convolutional neural networks. *2016 international joint conference on neural networks (IJCNN)*, pp. 2560–2567.

Studenỳ, M. & Vejnarová, J. (1998). The multiinformation function as a tool for measuring stochastic dependence. In *Learning in graphical models* (pp. 261–297). Springer.

Tobias, O. J., Seara, R., Soares, F. A. & Bermudez, J. C. (1995). Automatic visual inspection using the co-occurrence approach. *38th Midwest Symposium on Circuits and Systems. Proceedings*, 1, 154–157.

Traina, A. J. M. (2001). *Suporte à visualização de consultas por similaridade em imagens médicas através de estrutura de indexação métrica*. (Ph.D. thesis, ICMC/USP).

Tuceryan, M. & Jain, A. K. (1993). Texture analysis. In *Handbook of pattern recognition and computer vision* (pp. 235–276). World Scientific.

Tuv, E., Borisov, A., Runger, G. & Torkkola, K. (2009). Feature selection with ensembles, artificial variables, and redundancy elimination. *The Journal of Machine Learning Research*, 10, 1341–1366.

van der Laan, M. J., Polley, E. C. & Hubbard, A. E. (2007). Super Learner. *Statistical Applications in Genetics and Molecular Biology*, 6(1). doi: doi:10.2202/1544-6115.1309.

Vandamme, P., Pot, B., Gillis, M., De Vos, P., Kersters, K. & Swings, J. (1996). Polyphasic taxonomy, a consensus approach to bacterial systematics. *Microbiological reviews*, 60(2), 407–438.

Vane-Wright, R. I., Humphries, C. J. & Williams, P. H. (1991). What to protect?—Systematics and the agony of choice. *Biological conservation*, 55(3), 235–254.

Vellend, M., Cornwell, W. K., Magnuson-Ford, K. & Mooers, A. Ø. (2011). Measuring phylogenetic biodiversity. *Biological diversity: frontiers in measurement and assessment. Oxford University Press, Oxford*, 194–207.

Venkatesh, B. & Anuradha, J. (2019). A review of feature selection and its methods. *Cybernetics and information technologies*, 19(1), 3–26.

Vriesman, D., Britto Junior, A. S., Zimmer, A. & Koerich, A. L. (2019). Texture CNN for Thermoelectric Metal Pipe Image Classification. *IEEE 31st Int'l Conf on Tools with Artificial Intelligence*, pp. 569-574. doi: 10.1109/ICTAI.2019.00085.

Wagner, B. D., Grunwald, G. K., Zerbe, G. O., Mikulich-Gilbertson, S. K., Robertson, C. E., Zemanick, E. T. & Harris, J. K. (2018). On the Use of Diversity Measures in Longitudinal Sequencing Studies of Microbial Communities. *Frontiers in Microbiology*, 9. doi: 10.3389/fmicb.2018.01037.

Wang, C., Shi, J., Zhang, Q. & Ying, S. (2017). Histopathological image classification with bilinear convolutional neural networks. *39th Annual Intl. Conf. of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 4050–4053.

Wang, H., Li, Z., Li, Y., Gupta, B. & Choi, C. (2020). Visual saliency guided complex image retrieval. *Pattern Recognition Letters*, 130, 64–72.

Wang, K., Bichot, C.-E., Zhu, C. & Li, B. (2013). Pixel to patch sampling structure and local neighboring intensity relationship patterns for texture classification. *IEEE Signal Processing Letters*, 20(9), 853–856.

Wang, Y., Zhao, Y., Cai, Q., Li, H. & Yan, H. (2016). A varied local edge pattern descriptor and its application to texture classification. *Journal of visual communication and image representation*, 34, 108–117.

Watanabe, S. (1960). Information theoretical analysis of multivariate correlation. *IBM Journal of research and development*, 4(1), 66–82.

Whittaker, R. H. (1972). Evolution and measurement of species diversity. *Taxon*, 21(2-3), 213–251.

Xie, F., Li, F., Lei, C. & Ke, L. (2018). Representative band selection for hyperspectral image classification. *ISPRS International Journal of Geo-Information*, 7(9), 338.

Xu, X., Shi, Z. & Pan, B. (2017). A new unsupervised hyperspectral band selection method based on multiobjective optimization. *IEEE Geoscience and Remote Sensing Letters*, 14(11), 2112–2116.

Xu, Y., Yang, X., Ling, H. & Ji, H. (2010). A new texture descriptor using multifractal analysis in multi-orientation wavelet pyramid. *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 161–168.

Zhang, J. & Tan, T. (2002). Brief review of invariant texture analysis methods. *Pattern recognition*, 35(3), 735–747.

Zhang, J., Zhao, H. & Liang, J. (2013). Continuous rotation invariant local descriptors for texton dictionary-based texture classification. *Computer Vision and Image Understanding*, 117(1), 56–75.

Zhang, L., Zhou, Z. & Li, H. (2012). Binary gabor pattern: An efficient and robust descriptor for texture classification. *2012 19Th IEEE international conference on image processing*, pp. 81–84.

Zhang, S., Wang, H., Huang, W. & Zhang, C. (2018). Combining modified LBP and weighted SRC for palmprint recognition. *Signal, Image and Video Processing*, 12(6), 1035–1042.

Zhao, Y., Zhang, L., Li, P. & Huang, B. (2007). Classification of high spatial resolution imagery using improved Gaussian Markov random-field-based texture features. *IEEE Transactions on Geoscience and Remote Sensing*, 45(5), 1458–1468.

Zhuo, L., Zheng, J., Li, X., Wang, F., Ai, B. & Qian, J. (2008). A genetic algorithm based wrapper feature selection method for classification of hyperspectral images using support vector machine. *Geoinformatics 2008 and Joint Conference on GIS and Built Environment: Classification of Remote Sensing Images*, 7147, 71471J.