

**UNIVERSIDADE ESTADUAL DE PONTA GROSSA**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM COMPUTAÇÃO APLICADA**

**VONCARLOS MARCELO DE ARAÚJO**

**IDENTIFICAÇÃO DE ESPÉCIES DE PLANTAS UTILIZANDO COMBINAÇÃO DE  
CLASSIFICADORES**

**PONTA GROSSA**

**2016**

**VONCARLOS MARCELO DE ARAÚJO**

**IDENTIFICAÇÃO DE ESPÉCIES DE PLANTAS UTILIZANDO COMBINAÇÃO DE  
CLASSIFICADORES**

Dissertação apresentada ao Programa de Pós-Graduação em Computação Aplicada da Universidade Estadual de Ponta Grossa, como requisito parcial para obtenção do título de Mestre.

Orientador: Prof Dr. Alceu de Souza Britto Jr  
Coorientação: Prof<sup>a</sup>. Dr<sup>a</sup>. Rosane Falate

**PONTA GROSSA**

**2016**

**Ficha Catalográfica**  
**Elaborada pelo Setor de Tratamento da Informação BICEN/UEPG**

A663 Araújo, Voncarlos Marcelo de  
Identificação de espécies de plantas  
utilizando combinação de classificadores/  
Voncarlos Marcelo de Araújo. Ponta Grossa,  
2016.  
95f.

Dissertação (Mestrado em Computação  
Aplicada - Área de Concentração:  
Computação para Tecnologias em  
Agricultura), Universidade Estadual de  
Ponta Grossa.

Orientador: Prof. Dr. Alceu de Souza  
Britto Jr.

Coorientadora: Prof<sup>a</sup> Dr<sup>a</sup> Rosane Falate.

1. Identificação de plantas. 2. Extração  
de características. 3. Combinação de  
classificadores. 4. Reconhecimento de  
padrões. I. Britto Jr, Alceu de Souza. II.  
Falate, Rosane. III. Universidade Estadual  
de Ponta Grossa. Mestrado em Computação  
Aplicada. IV. T.

CDD: 006.3

## TERMO DE APROVAÇÃO

Voncarlos Marcelo de Araujo

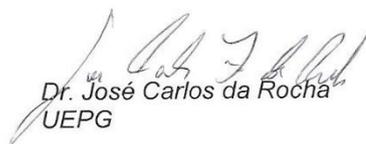
“IDENTIFICAÇÃO DE ESPÉCIES DE PLANTAS UTILIZANDO COMBINAÇÃO DE CLASSIFICADORES”.

Dissertação aprovada como requisito parcial para obtenção do grau de Mestre no Programa de Pós-Graduação em Computação Aplicada da Universidade Estadual de Ponta Grossa, pela seguinte banca examinadora:

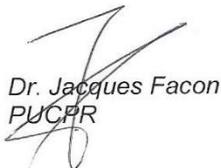
Orientador:



Dr. Alceu de Souza Britto Junior  
UEPG



Dr. José Carlos da Rocha  
UEPG



Dr. Jacques Facon  
PUCPR

Ponta Grossa, 04 de março de 2016.

## AGRADECIMENTOS

Agradeço primeiramente a Deus pela minha vida, que me concedeu sabedoria e força para estar concluindo este curso.

A minha família, Vera e Francielli que nunca me deixaram desistir dos meus sonhos, e me apoiaram do início ao fim do mestrado, também a minha segunda família, Sonia, Oscar (*in memorium*), Marcelo e Leandro Pilarski, que me receberam de braços abertos em sua casa em Ponta Grossa, a vocês agradeço imensuravelmente por toda a gratidão. A minha parceira Taina e sua família, por sempre me ajudar, para que fosse possível seguir em frente no mestrado.

Ao meu orientador, Prof. Dr Alceu de Souza Britto Jr, que de forma muito profissional e paciente me ajudou na condução deste trabalho.

À minha coorientadora, Profa. Dr. Rosane Falate, que me ajudou a ampliar minha visão sobre as aplicações de técnicas de PDI, juntamente com o seu grupo de pesquisa CAFIRA.

À Prof Dra Marluce Gonçalves Cortez, Coordenadora do Curso de Agronomia, que deu dicas essenciais da área agrônômica, para a definição do escopo do trabalho.

Aos professores e colegas que colaboraram significativamente no progresso da minha formação, em especial os meus colegas do mestrado, Alisson Prestes, Luiz, Teruo, Sergio Ribeiro, Thiago, Wilian, pelas contribuições, conversas e conselhos durante esse tempo.

À UEPG, por disponibilizar toda a sua estrutura, tanto para o programa de mestrado quanto para esta pesquisa, bem como pela qualidade dos professores, comprovada nas aulas.

À CAPES, pela ajuda, com ela foi possível conceder dedicação exclusiva ao mestrado.

E finalmente, meu agradecimento a todos que contribuíram direta ou indiretamente para a realização desse trabalho.

## RESUMO

A biodiversidade das espécies de plantas desempenha um papel fundamental na ecologia da Terra, fornecendo alimento, abrigo e mantendo uma atmosfera respirável saudável para todos os seres vivos. As plantas também têm propriedades medicinais e são utilizadas para fontes alternativas de energia, como o biocombustível. No entanto, o número de plantas em risco de extinção tem aumentado gradativamente e as dificuldades presentes no processo manual de reconhecimento de plantas, torna esta tarefa muito complexa e morosa. Uma metodologia viável para a identificação das plantas, ou para fornecer uma categorização de plantas, é a aquisição da imagem da planta e o uso técnicas de reconhecimento de padrões. Dessa forma, o uso da computação, apesar de ainda ter pequena contribuição na área, pode prover informações importantes sobre a taxonomia das plantas, além de poder servir como base para sistemas que executem tarefas como a de seleção de determinado tipo de plantas ou que guiem o especialista para possíveis tomadas de decisões. Neste trabalho é proposto um método para classificação de plantas baseado em imagens colaborativas de especialistas do mundo inteiro. Esse método é capaz de lidar com algumas complexidades impostas durante a captura das imagens, como a presença de ruídos (luminosidade, sombras e objetos indesejáveis) e variações de posições das plantas. Para cumprir essa tarefa são utilizados descritores de textura baseados em SIFT, SURF e HOG, que têm mostrado excelentes resultados em diversos trabalhos. Para possibilitar os testes do método proposto, foi empregada uma base de imagens disponibilizada pela tarefa mundial de reconhecimento de plantas em 2011, ImageCLEF, que contém cerca de 2.586 amostras de plantas composta por 41 espécies divididas em duas categorias distintas: a primeira com 13 espécies e imagens com presença de ruídos, e a segunda com 28 espécies e imagens de folhas plotadas em um fundo branco. Os resultados dos experimentos mostram que os classificadores treinados com descritores de textura são capazes de atingir boas taxas de acertos, próximas a 70%, dada a complexidade do problema. Métodos de combinação de classificadores também foram utilizados e se mostraram capazes de melhorar o desempenho dos classificadores, principalmente nos testes com imagens que tem a presença de ruídos.

**Palavras-Chave:** Identificação de plantas, Extração de características, Combinação de classificadores, Reconhecimento de padrões

## ABSTRACT

The biodiversity of plant species plays a key role in the Earth's ecology, providing food, shelter and maintaining a healthy breathable atmosphere for all living beings. The plants also have medicinal properties and are used for alternative energy sources, such as biofuel. However, the number of plants endangered has gradually increased and the difficulties in the plants manual recognition process, does become a complex and slow task. A viable method for the identification of plants, or to provide a categorization of the plant, is the plant image acquisition and use pattern recognition techniques. In this way, the use of computers, despite having little contribution in the area, can provide important information on the taxonomy of plants, and can serve as a basis for systems that perform tasks such as the selection of certain plants or to guide the specialist for possible decision-making. This paper proposes a method for classification of plants based on collaborative images of the world experts. This method is able to deal with some complexities imposed during the capture of images, as the presence of noise (lighting, shadows and undesirable objects) and plants position variations. To accomplish this task are used texture descriptors based on SIFT, SURF and HOG, which have shown excellent results in several works. To enable testing of the proposed method, we used an image provided by the global task basis for recognition of plants in 2011, ImageCLEF, containing about 2,586 plant samples composed by 41 species divided into two distinct categories: the first one with 13 species and images with presence of noise, and with the second species and 28 sheets of images plotted on a white background. The results of the experiments show that the classifiers trained with texture descriptors are able to achieve good hit rates close to 70%, given the complexity of the problem. Classifiers combination methods have also been used and have been shown capable to improve the performance of classifiers, especially in the test with images that has the presence of noises.

**Keywords:** Plant identification, Extraction of features, Combination of classifiers, Pattern recognition

## LISTA DE FIGURAS

FIGURA 1.1: EXEMPLOS DE VARIAÇÕES SOBRE A MESMA ESPÉCIE DE PLANTA ( <i>INTRA-ESPÉCIES</i> ).....	13
FIGURA 1.2: EXEMPLOS DE SEMELHANÇAS ENTRE DIFERENTES ESPÉCIES DE PLANTAS ( <i>INTER-ESPÉCIES</i> ): (A) <i>ADOXACEAE VIBURNUM TINUS</i> , (B) <i>LAURACEAE LAURUS NOBILIS</i> E (C) <i>ERICACEAE ARBUTUS UNEDO</i> .....	13
FIGURA 2.1: ETAPAS ENVOLVIDAS NA CONSTRUÇÃO DE UM CLASSIFICADOR.....	17
FIGURA 2.2: CRIAÇÃO DO DESCRITOR SIFT DE PONTOS CHAVES: (A) OBTENÇÃO DOS GRADIENTES PARA CADA DIREÇÃO NA REGIÃO SOBREPOSTA; (B) AGLOMERAÇÃO DOS GRADIENTES PARA CADA SUB-REGIÃO. ....	21
FIGURA 2.3: EXEMPLO DE USO DE IMAGEM ÍNTEGRA: ILUSTRAÇÃO DO CÁLCULO DA SOMA DOS PIXELS CONTIDOS EM UMA SUB-REGIÃO DBCA DA IMAGEM COM APENAS TRÊS OPERAÇÕES DE ADIÇÃO. ....	23
FIGURA 2.4: JANELA DESLIZANTE: A VARIAÇÃO DOS PIXELS DE UMA REGIÃO EM VOLTA DE UM PONTO DE INTERESSE .....	23
FIGURA 2.5: DISTRIBUIÇÕES DE HAAR WAVELETS: (A) APLICADO A MÁSCARA 8X8 COM AS ORIENTAÇÕES IDENTIFICADAS E OBTENÇÃO DO GRADIENTE PARA CADA DIREÇÃO NA REGIÃO SOBREPOSTA DA IMAGEM; (B) ACUMULO DO GRADIENTE PARA CADA REGIÃO.....	24
FIGURA 2.6: UMA VISÃO GLOBAL DO CÁLCULO DO DESCRITOR HOG. A IMAGEM INTEIRA É PREENCHIDA COM UMA GRADE DE BLOCOS SOBREPOSTOS. CADA BLOCO É FORMADO POR UM CONJUNTO DE CÉLULAS, NAS QUAIS SÃO CALCULADOS HISTOGRAMAS DA ORIENTAÇÃO DO GRADIENTE DE CADA PIXEL PRESENTE NAS MESMAS. OS HISTOGRAMAS SÃO NORMALIZADOS LOCALMENTE EM RELAÇÃO AOS BLOCOS E CONCATENADOS PARA A CRIAÇÃO DO VETOR DE CARACTERÍSTICAS. ....	26
FIGURA 2.7: MÉTODOS DE TREINAMENTO SVM: DECOMPOSIÇÃO ( <i>CHUNKING</i> ) E SMO. ....	27
FIGURA 2.8: FUNCIONAMENTO DE UM REDE NEURAL COM DUAS CAMADAS INTERMEDIÁRIAS.....	28
FIGURA 2.9: FLUXO DO PROCESSAMENTO DO ALGORITMO <i>BACKPROPAGATION</i> .....	29
FIGURA 2.10: EXEMPLO DO FUNCIONAMENTO DO ALGORITMO KNN .....	30
FIGURA 2.11: EXEMPLO DO FUNCIONAMENTO DE UMA ÁRVORE ALEATÓRIA .....	31
FIGURA 2.12: DIFERENTES ABORDAGENS PARA A CONSTRUÇÃO DE CONJUNTOS EM COMBINAÇÃO DE CLASSIFICADORES .....	33
FIGURA 2.13: ARQUITETURA DE FUSÃO DAS RESPOSTAS DE DIVERSOS CLASSIFICADORES .....	34
FIGURA 3.1: (A) IMAGEM DA FOLHA EM RGB ORIGINAL, (B) IMAGEM CONVERTIDA PARA HSV, E (C) SEGMENTAÇÃO USANDO EM. ....	41
FIGURA 3.2: METODOLOGIA DO TRABALHO PROPOSTO POR ARUN, EMMANUEL E DURAIRAJ. ....	44
FIGURA 3.3: (A) FOLHA DE FEIJÃO DE SOJA. (B) SEGMENTAÇÃO DA IMAGEM PARA (A). (C) FOLHA DE FEIJÃO BRANCO. (D) SEGMENTAÇÃO DA IMAGEM PARA (C). (E) FOLHA DE FEIJÃO VERMELHO. (F) SEGMENTAÇÃO DA IMAGEM PARA (E). ....	45
FIGURA 3.4: ARQUITETURA DO TRABALHO PROPOSTO POR MZOGUI.....	47
FIGURA 3.5: ARQUITETURA DO TRABALHO PROPOSTO POR CERUTTI.....	51
FIGURA 4.1: FLUXOGRAMA PARA CLASSIFICAÇÃO DE ESPÉCIES DE PLANTAS .....	60
FIGURA 4.2: RUÍDOS NA CATEGORIA SCAN-LIKE. (A) VARIAÇÃO DE ROTAÇÃO DAS PLANTAS DE MESMA ESPÉCIE GINKGO BILOBA, (B) PRESENÇA DE SOMBRAS E LUMINOSIDADE DA ESPÉCIE QUERCUS ILEX, (C) RUÍDOS NA ESPÉCIE NERIUM OLEANDER.....	63
FIGURA 4.3: RUÍDOS NA CATEGORIA SCAN. (A) QUERCUS ILEX: MANCHAS AO REDOR DA FOLHA E PEDAÇO DE TERRA, (B) JUGLANS REGIA: VARIAÇÕES DA MORFOLOGIA NA MESMA ESPÉCIE DE PLANTA. ....	64
FIGURA 4.4: EXEMPLO DE DESCRITOR XML DE UMA PLANTA. ....	65
FIGURA 4.5: DELIMITAÇÃO DO HISTOGRAMA DA IMAGEM PARA A SEPARAÇÃO DA SOMBRA CONTIDA NA IMAGEM. ....	66
FIGURA 4.6: (A) DIOSPYROS KAKI: SEPARAÇÃO DA SOMBRA COMPLETA, (B) PLATANUS X HISPANICA: SEPARAÇÃO INCOMPLETA DA SOMBRA.....	66
FIGURA 4.7: ACER CAMPESTRE: REMOÇÃO APENAS DO FUNDO PELO ALGORITMO DE OTSU. ....	67
FIGURA 4.8: QUERCUS ILEX: REMOÇÃO DE OBJETOS INDESEJADOS. ....	68
FIGURA 4.9: ESQUEMA DE TREINAMENTO DE CLASSIFICADOR POR CATEGORIA. ....	72
FIGURA 4.10: EXEMPLOS DE IMAGEM DE CADA ESPÉCIE PARA A CATEGORIA SCAN-LIKE.....	75
FIGURA 4.11: EXEMPLOS DE IMAGEM DE CADA ESPÉCIE PARA A CATEGORIA SCAN. ....	75
FIGURA 5.1: LUMINOSIDADE NA CLASSE <i>ILEX AQUIFOLIM</i> , OBJETOS PROJETANDO PADRÕES DE TEXTURA NA CLASSE <i>LAURUS NOBILIS</i> E O NÃO PREENCHIMENTO DE ESPAÇOS VAZIOS NA CLASSE <i>CERCIS SILIQUASTRUM</i> .....	78
FIGURA 5.2: SEMELHANÇA ENTRE ESPÉCIES DISTINTAS. ....	79
FIGURA 5.3: LUMINOSIDADE EXCESSIVA NA ESPÉCIE <i>ILEX AQUIFOLIM</i> . ....	79
FIGURA 5.4: ROTAÇÃO NAS AMOSTRAS DA ESPÉCIE PLATANUS X HISPANICA. ....	80

## LISTA DE TABELAS

TABELA 2.1: MATRIZ DE CONFUSÃO PARA PROBLEMAS DE DUAS CLASSES .....	37
TABELA 2.2: MATRIZ DE CONFUSÃO PARA PROBLEMAS DE MÚLTIPLAS CLASSES.....	37
TABELA 3.1: DESCRITORES UTILIZADOS NO TRABALHO DE (YANIKOGLU; APTOULA; TIRKAZ, 2011). .....	52
TABELA 3.2: RESULTADO GERADOS ATRAVÉS DA VALIDAÇÃO CRUZADA NO TRABALHO DE (YANIKOGLU; APTOULA; TIRKAZ, 2011).....	53
TABELA 3.3: SUMARIZAÇÃO DAS METODOLOGIAS.....	9
TABELA 4.1: DETALHAMENTO DA BASE DE DADOS .....	63
TABELA 4.2: EXTRAÇÃO DE CARACTERÍSTICAS.....	69
TABELA 4.3: PARÂMETROS UTILIZADOS NA CLASSIFICAÇÃO.....	71
TABELA 4.4: SEPARAÇÃO ENTRE A BASE DE TREINAMENTO E TESTES. ....	74
TABELA 5.1: CATEGORIA SCAN-LIKE: MELHORES CLASSIFICADORES INDIVIDUAIS ORDENADOS PELA TAXA DE ACERTO. ....	76
TABELA 5.2: CATEGORIA SCAN-LIKE: MATRIZ DE CONFUSÃO DO MELHOR CLASSIFICADOR INDIVIDUAL HOG+RNA.....	77
TABELA 5.3: CATEGORIAS SCAN-LIKE: MATRIZ DE CONFUSÃO DO CLASSIFICADOR INDIVIDUAL SURF+RF. ....	78
TABELA 5.4: CATEGORIA SCAN-LIKE: MATRIZ DE CONFUSÃO DO CLASSIFICADOR INDIVIDUAL SIFT+RNA. ....	80
TABELA 5.5: CATEGORIA SCAN: MELHORES CLASSIFICADORES INDIVIDUAIS ORDENADOS PELA TAXA DE ACERTO. .....	81
TABELA 5.6: CATEGORIA SCAN: MATRIZ DE CONFUSÃO DO MELHOR CLASSIFICADOR INDIVIDUAL SIFT+RNA. ..	82
TABELA 5.7: CATEGORIA SCAN: MATRIZ DE CONFUSÃO DO CLASSIFICADOR INDIVIDUAL SURF+RNA. ....	82
TABELA 5.8: CATEGORIA SCAN: MATRIZ DE CONFUSÃO DO CLASSIFICADOR INDIVIDUAL HOG+RNA. ....	83
TABELA 5.9: CATEGORIA SCAN: MELHORES RESULTADOS DE COMBINAÇÃO COM A REGRA PRODUTO. ....	84
TABELA 5.10: CATEGORIA SCAN: MATRIZ DE CONFUSÃO DA COMBINAÇÃO DE CLASSIFICADORES COM A REGRA PRODUTO. ....	85
TABELA 5.11: CATEGORIA SCAN-LIKE: MELHORES RESULTADOS DE COMBINAÇÃO COM A REGRA PRODUTO. ....	86
TABELA 5.12: CATEGORIA SCAN-LIKE: MATRIZ DE CONFUSÃO DA COMBINAÇÃO DE CLASSIFICADORES COM A REGRA PRODUTO .....	87
TABELA 5.13: COMPARAÇÃO DOS RESULTADOS.....	88
TABELA 8.1: AVALIAÇÃO DO <i>CODEBOOK</i> DOS DESCRITORES SIFT E SURF.....	95

## SUMÁRIO

<b>1. INTRODUÇÃO .....</b>	<b>9</b>
1.1. Motivação .....	10
1.2. Desafios .....	12
1.3. Objetivos .....	14
1.4. Contribuições.....	14
1.5. Estrutura do Documento .....	15
<b>2. FUNDAMENTAÇÃO TEÓRICA.....</b>	<b>16</b>
2.1. Reconhecimento de Padrões .....	16
2.2. Representação de Características Texturais.....	18
2.3. Classificadores.....	26
2.4. Combinação de Classificadores .....	32
2.5. Métricas de Avaliação de Classificadores .....	36
<b>3. ESTADO DA ARTE.....</b>	<b>40</b>
3.1 Sistemas de reconhecimento de plantas.....	40
3.2 ImageCLEF .....	48
3.3 Sumário do Estado da Arte.....	53
3.4 Discussão Crítica Sobre os Métodos Apresentados.....	59
<b>4. MÉTODO PROPOSTO.....</b>	<b>58</b>
4.1 Modelagem do Problema.....	58
4.2 Base de Dados .....	60
4.3 Pré-processamento e segmentação da imagem .....	65
4.4 Criação dos Classificadores.....	68
4.5 Metodologia de Avaliação do Método Proposto .....	72
<b>5. RESULTADOS EXPERIMENTAIS .....</b>	<b>76</b>
5.1 Avaliações dos descritores: melhores resultados individuais para a categoria Scan-like .....	76
5.2 Avaliações dos descritores: melhores resultados individuais para a categoria Scan .....	80
5.3 Combinação de Classificadores.....	83
5.4 Comparação dos resultados obtidos.....	88
<b>6. CONCLUSÃO .....</b>	<b>90</b>
<b>7. REFERÊNCIAS BIBLIOGRÁFICAS .....</b>	<b>91</b>
<b>8. APÊNDICE A – GERAÇÃO DO CODEBOOK PARA OS DESCRITORES SIFT E SURF .....</b>	<b>95</b>

## 1. INTRODUÇÃO

A biodiversidade de plantas é indispensável para o ecossistema terrestre, pois todos os seres vivos dependem, direta ou indiretamente, das inúmeras espécies de plantas que proporcionam diferentes formas de energia na natureza (CHAKI; PAREKH; BHATTACHARYA, 2015). As plantas são consideradas os principais fornecedores de oxigênio na Terra, uma vez que estas convertem gás carbônico em oxigênio (essencial para a maioria dos organismos vivos) por meio da fotossíntese.

A importância das plantas para o ser humano ainda vai além. As diferentes espécies de plantas são utilizadas para uma ampla gama de aplicações industriais e em diferentes setores como o da nutrição, da produção de ervas e plantas medicinais e dos biocombustíveis, usados na geração de energia sustentável e renovável (ADAM et al., 2012; PROCHNOW et al., 2009). Além disso, as plantas ajudam a regular o clima, servindo de *habitat* e comida para insetos e outros animais.

Um bom conhecimento sobre a flora é crucial para aumentar a produtividade agrícola e garantir a sustentabilidade do planeta. Por isso, tornam-se importantes abordagens que permitam reconhecer e identificar espécies novas e raras ou ainda realizar a categorização de diferentes espécies de plantas.

Nos últimos anos, técnicas de visão computacional e reconhecimento de padrões têm sido empregadas para preparar sistemas de categorização de plantas por meio de imagens digitais de maneira eficiente. Isso se torna relevante considerando que muitas espécies de plantas estão à beira da extinção devido ao constante desmatamento para pavimentar o caminho para a modernização.

Diante da importância desse tema, o presente trabalho tem como foco a construção de um método computacional para a identificação de espécies de plantas por meio do reconhecimento de imagens de sua componente folha, utilizando combinação de classificadores. O presente capítulo apresenta a motivação do trabalho, focando diversos aspectos e perspectivas quanto ao tema. Destacam-se também os desafios a serem transpostos, os objetivos que delimitam o escopo do trabalho, suas contribuições, além da estrutura geral deste relatório.

## 1.1. MOTIVAÇÃO

Reconhecer plantas, nos dias atuais, não é apenas um trabalho para botânicos e ecologistas. Com a interdisciplinaridade e o avanço da tecnologia, essa tarefa é exigida em diferentes áreas, como agricultura, arquitetura e urbanismo, paisagismo, medicina, biologia, e para o público em geral como os amantes da natureza, conservacionistas, biólogos e visitantes de parques florestais. No entanto, a tarefa de identificar plantas é dispendiosa e não trivial, dado o elevado número de espécies de plantas existentes, e a necessidade de treinamento e conhecimento profundo sobre a sistemática das plantas.

De fato, a abordagem tradicional para a identificação de espécies de plantas requer um especialista, o qual possui o conhecimento necessário para examinar uma amostra (parte de uma planta ou a planta) e atribuir um rótulo taxonômico<sup>1</sup> com base na análise de amostras de espécies de plantas normalmente disponíveis em imagens. Dadas as dificuldades mencionadas, uma abordagem promissora é o desenvolvimento de um sistema preciso, rápido e eficiente para auxiliar na identificação automática de uma vasta gama de espécies de plantas.

A facilidade de acesso às tecnologias, como as câmeras digitais e computadores, trouxe essa ideia de identificação de objetos de forma automática mais próxima da realidade (HUSIN et al., 2012; KUMAR et al., 2012; ZHAO et al., 2015). Além disso, a multidisciplinaridade envolvendo as áreas de Processamento de Imagem e da Aprendizagem de Máquina (Visão Computacional e Reconhecimento de Padrões) também está contribuindo para que essa ideia seja concretizada, o que vem despertando interessados em variados segmentos.

Dentre esses segmentos estão as indústrias que fazem o uso de plantas em seu processo de produção, pois sistemas de identificação automática de plantas possibilitam atingir altas taxas de reconhecimento de sua matéria prima, acelerando algumas partes do processo produtivo. Sistemas dotados de visão computacional também auxiliam na agricultura, pela identificação de doenças de plantas (SANKARAN et al., 2010); na detecção de deficiência nutricional da planta (ROMUALDO et al., 2014); e na seleção de plantas pela qualidade (SILVA et al., 2013).

---

<sup>1</sup> Taxonomia é a ciência que classifica os seres vivos. Também chamada de “rotulo taxonômico”, ela estabelece critérios para classificar todos os animais e plantas sobre a Terra em grupos de acordo com as características fisiológicas, evolutivas e anatômicas e ecológicas de cada grupo de plantas ou animal

Porém, essas aplicações ainda sofrem restrições devido à deficiência de pesquisas na área e à falta de um conjunto de dados robusto que abranja uma quantidade satisfatória de espécies de plantas. Por exemplo, pesquisas para identificar espécies florestais, usualmente empregam classificadores tradicionais e bases de imagens com um número reduzido de espécies florestais, diante da grande diversidade existente. Ou seja, as soluções geralmente não são adequadas em razão das plantas terem um elevado número de classes, o qual chega a milhares de espécies florestais, e, talvez, pelo uso inadequado do classificador (ASLINA et al., 2013; HERDIYENI; KADEK; WAHYUNI, 2012; LE; TRAN; PHAM, 2014).

Diante das considerações anteriormente expostas, e da inviabilidade de se construir modelos que abrangem todas as espécies florestais existentes, o presente trabalho explora as técnicas fornecidas por sistemas baseados em combinação de classificadores. Sistemas de múltiplos classificadores (*Multi-Classififer Systems* - MCS) têm sido útil para melhorar as taxas de reconhecimento em problemas de classificação. Isso é possível devido à combinação das vantagens de diferentes opiniões (classificadores), as quais podem oferecer informações complementares para aumentar a efetividade geral do processo de reconhecimento (KUNCHEVA, 2004). Tal contexto é intuitivo uma vez que um conjunto simultâneo de opiniões de diferentes classificadores pode gerar taxas de classificação melhoradas (WOODS; KEGELMEYER; BOWYER, 1997).

Combinação de classificadores é uma das principais abordagens em MCS. Nessa abordagem cada amostra é avaliada individualmente e paralelamente por cada classificador, então, cada decisão individual é combinada com as demais seguindo alguma estratégia para alcançar uma decisão global e em comum acordo (JAIN; DUIN; MAO, 2000). Dentre suas vantagens, destaca-se o pressuposto de que os erros de classificação são independentes entre si. No entanto, quando a condição de independência não se verifica para parte dos classificadores gerados, não se pode assegurar que a fusão dos classificadores irá melhorar o desempenho da classificação final (KITTLER et al., 1998).

Sabendo da inviabilidade de se construir um modelo totalmente genérico para abranger milhares de espécies florestais existentes, o presente trabalho busca construir um método robusto para a classificação de espécies florestais utilizando características texturais presentes em imagens de folhas de plantas, por meio das estratégias de combinação de classificadores

## 1.2. DESAFIOS

A tarefa de identificação de espécies de plantas, bem como sua catalogação e estudo, é um desafio para os pesquisadores frente a sua abundante biodiversidade. Também, a natureza dos vegetais, do ponto de vista de suas características e peculiaridades, faz com que o processo de reconhecimento através das folhas seja um desafio adicional para um sistema de visão computacional.

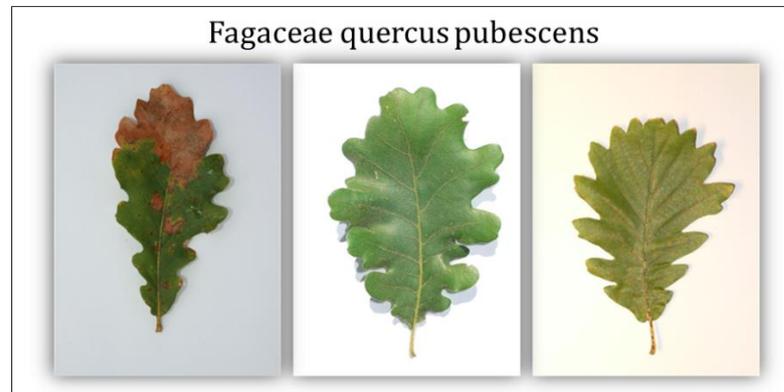
Existem alguns conjuntos de dados, tais como Flavia (WU et al., 2007) e Swedish (SÖDERKVIST, 2001) que foram construídos especialmente para avaliar o uso de Visão Computacional e Aprendizagem de Máquina. A desvantagem de tais conjuntos, no entanto, é a falta de diversidade, limitando a capacidade de generalização. Tais bases de dados foram criadas por poucas pessoas que coletam as folhas em um período limitado de tempo, em áreas restritas e usando exatamente o mesmo protocolo de aquisição. Como resultado, todas as folhas de uma dada espécie de tal conjunto de dados são mais propensas a representar algumas plantas individuais em um determinado local e período do ano.

Assim, ao usar esses dados para o treinamento de um sistema fará que a aplicação de identificação, no mundo real, tenha pouca chance de classificar corretamente as imagens que foram coletadas em distintos períodos ou lugares e/ou adquiridos de forma diferente. Portanto um dos desafios é apresentar uma base de dados que consiga satisfazer o processo de reconhecimento independente do lugar, da árvore e da pessoa que está tentando identificar uma planta, imitando o mundo real.

Outro importante desafio são as variações morfológicas encontradas em uma mesma espécie de folha. Fatores como, maturidade ou até mesmo o processo de formação das folhas (exposição ao sol, chuva etc.), fazem com que essas tenham grandes variações quanto ao tamanho, coloração, textura, formato etc. A Figura 1.1 fornece alguns exemplos dessa variabilidade visual, chamada *intra-espécies*, ao longo de vários critérios, incluindo a forma da folha global, o número e as posições relativas dos folhetos e sua coloração.

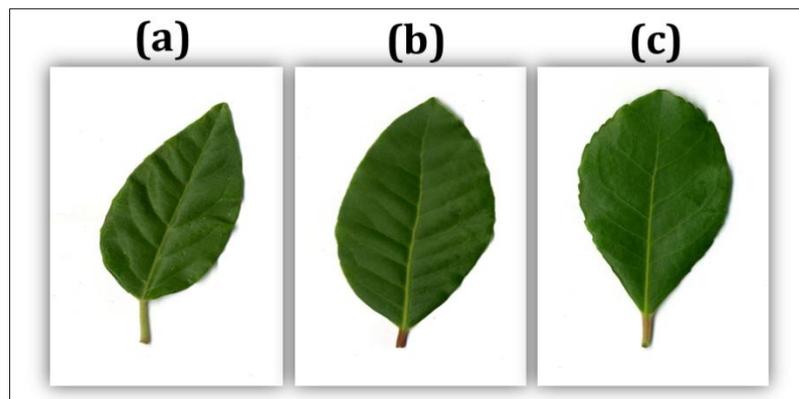
Além disso, indivíduos de diferentes espécies podem possuir diferenças insignificantes, mesmo quando situados em regiões ou ecossistemas diferentes. A Figura 1.2 mostra variações morfológicas similares entre diferentes espécies, chamada de *inter-espécies*

Figura 1.1: Exemplos de variações sobre a mesma espécie de planta (*intra-espécies*).



Fonte: O autor.

Figura 1.2: Exemplos de semelhanças entre diferentes espécies de plantas (*inter-espécies*): (a) *Adoxaceae viburnum tinus*, (b) *Lauraceae laurus nobilis* e (c) *Ericaceae arbutus unedo*.



Fonte: O autor.

O último desafio é em relação aos ruídos encontrados nas imagens que, dependendo do modo de aquisição dos diferentes usuários que contribuíram para a construção da base de dados utilizada neste trabalho, as imagens podem apresentar diversidades em termos de sombras, reflexos de luzes e objetos indesejáveis, como pedaços de terra, dedos etc. Nesse caso, o desafio está em conseguir identificar apenas o objeto principal, que é a folha, ignorando todo o conteúdo restante e que pode perturbar o processo de identificação.

### 1.3. OBJETIVOS

O objetivo geral deste trabalho é construir um sistema robusto para classificação de espécies florestais a partir de características texturais extraídas de imagens de folhas de plantas. Para isso, tal sistema teve como base estratégias de combinação de classificadores.

Para o cumprimento do objetivo geral, estabeleceram-se os seguintes objetivos específicos:

- Avaliar um método para fazer um pré-processamento na imagem, que consiste em retirar objetos indesejáveis da imagem, deixando somente o objeto a ser identificado, ou seja, a folha;
- Definir descritores de textura e classificadores que sejam complementares, para que estes classificadores possam ser usados na combinação dentre eles;
- Avaliar estratégias de combinação de classificadores com o propósito de aumentar a taxa de acertos do sistema. Para isso serão treinados diferentes classificadores, com conjuntos de características específicos, aonde a resposta final da classificação de uma amostra de planta será dada por meio de um esquema da combinação das respostas destes classificadores.

### 1.4. CONTRIBUIÇÕES

As contribuições deste trabalho podem ser relacionadas aos contextos científico e social. Sob o contexto científico, as principais contribuições esperadas para este trabalho são:

- A avaliação do desempenho de descritores baseados em textura para a distinção sobre diversas espécies de folhas de plantas.
- A avaliação de estratégias de combinação de diversos classificadores para classificação de imagens utilizando-se características baseado em textura.

As contribuições sociais esperadas para este trabalho são:

- A criação de um método para classificação e identificação de folhas de plantas que possa servir como base, por exemplo, para outros sistemas existentes em laboratórios, aonde a identificação de plantas automática pouparia tempo dos especialistas em botânica e diminuiria gastos com a adição de profissionais especializados na área.
- A categorização das plantas de forma eficaz, provendo informações sobre cada uma das plantas, para a preservação contra a extinção sobre espécies ameaçadas.

## **1.5. ESTRUTURA DO DOCUMENTO**

Este trabalho está organizado em seis capítulos. O Capítulo 2 trata dos principais conceitos de reconhecimento de padrões utilizados no método proposto. No capítulo 3 é discutido o estado da arte do reconhecimento automático de plantas voltados para imagens de folha. No Capítulo 4 é descrita a abordagem proposta para a detecção automática de folhas de plantas utilizando-se imagens e combinação de classificadores de textura, bem como metodologia de avaliação do método. No capítulo 5 são apresentados os resultados obtidos. Finalmente, no capítulo 6, são apresentados as conclusões e os direcionamentos para trabalhos futuros.

## 2. FUNDAMENTAÇÃO TEÓRICA

Neste capítulo são apresentados os principais conceitos envolvidos no desenvolvimento do método proposto. Na Seção 2.1 é dada a definição de um sistema de reconhecimento de padrões. A Seção 2.2 trata da extração e seleção de características de textura de imagens. Já na Seção 2.3 são apresentados os classificadores do tipo Otimização Sequencial Mínima (*Sequential Minimal Optimization* - SMO), Redes Neurais Artificiais (*Artificial Neural Network* - RNA), K-vizinhos mais próximos (*K-Nearest Neighbors* - KNN) e Árvores Aleatórias (*Random Forest* - RF), enquanto na Seção 2.4 são descritas as estratégias de combinação de classificadores. Por fim, na Seção 2.5 são expostas algumas métricas de avaliação de classificadores.

### 2.1. RECONHECIMENTO DE PADRÕES

Padrões são a forma em que o mundo é interpretado pelos seres vivos, e por meio desta interpretação é que são elaboradas atitudes e decisões. O ser humano consegue de forma fácil reconhecer pessoas, carros, até mesmo, estado de humor nos rostos de pessoas desconhecidas. No entanto, o grande desafio proposto aos pesquisadores da área de reconhecimento de padrões, é desenvolver sistemas computacionais que tenham capacidade de simular tais comportamentos (DUDA; HART; STORK, 2000).

Nos sistemas de reconhecimento de padrões, normalmente um padrão é formado por diversas características que podem ser descritas, por exemplo, através de suas formas, cores ou texturas. Tais características são usualmente agrupadas para então formarem vetores de características na forma  $X = [x_1, x_2, \dots, x_N]$ , contendo  $N$  valores. Esses vetores de características são utilizados pelo sistema de reconhecimento tanto na criação de um modelo quanto na etapa de classificação dos objetos (THEODORIDIS; KOUTROUMBAS, 2009).

O reconhecimento de padrões é uma das tarefas mais importantes após a etapa de extração de características (GONZALEZ; WOODS, 2002). Quando esses padrões possuem características em comum, ou seja, características que são similares dentro de um grupo de padrões, denotam-se que estes padrões pertencem a uma mesma classe. Com base nas informações disponíveis, os sistemas de reconhecimento podem ser capazes de determinar como os objetos são divididos em classes (SONKA; HLAVAC; BOYLE, 1999).

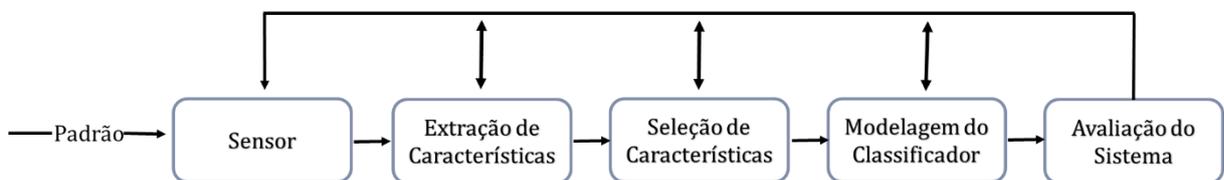
Na literatura podem-se encontrar duas abordagens principais para reconhecimento de padrões (JAIN; DUIN; MAO, 2000): a não-supervisionada e a supervisionada. Na primeira, na abordagem não-supervisionada, não se sabe previamente quais as possíveis classes de objetos existentes. Nesse caso, para um conjunto de objetos de entrada, é necessário descobrir algum tipo de relação, através das suas características, que os classifique e defina seus agrupamentos.

Já para sistemas com a abordagem supervisionada, foco deste trabalho, o padrão de entrada é identificado como membro de uma classe pré-definida, e então é utilizado na construção de um modelo de classificação, que formará a base de conhecimento para a classificação dos próximos padrões desconhecidos.

Nesses sistemas são utilizados algoritmos de classificação geralmente chamados de indutores, enquanto uma instância de um indutor gerado através de um conjunto específico de treinamento é denominada classificador (DUDA; HART; STORK, 2000).

A Figura 2.1 contempla as principais etapas envolvidas na construção de um sistema de classificação.

Figura 2.1: Etapas envolvidas na construção de um classificador.



Fonte: Adaptado de Theodoridis e Koutroumbas (2009).

Na primeira etapa, para coletar os objetos que serão identificados, é utilizado um sensor. Na segunda etapa são extraídas as características dos objetos, e na terceira busca-se a redução do número de características, por meio de algum método de seleção. Na quarta fase é realizada a modelagem do classificador através da definição de aspectos como o indutor a ser utilizado, a forma em que serão realizados os treinamentos e as possíveis estratégias de combinação de classificadores. Por fim, na última fase, é realizada a avaliação do desempenho do sistema como um todo.

Com o objetivo de maximizar o resultado geral do sistema, cada etapa ou fase apresentada pode ser remodelada dependendo dos resultados encontrado nas demais e, por conta disto, as setas de retorno, indicam que as etapas são inter-relacionadas (THEODORIDIS; KOUTROUMBAS, 2009).

Como este trabalho trata de classificação de espécies de plantas através de imagens de folhas, uma câmera é usada como sensor para fazer a coleta das imagens do banco de dados ImageCLEF (GOËAU et al., 2011). Características de textura foram extraídas dos objetos (Seção 2.2), e indutores distintos juntamente com estratégias de combinação de classificadores foram escolhidos para a modelagem do sistema de classificação (Seções 2.3 e 2.4, respectivamente). A avaliação do sistema é realizada através de métricas como taxas de acerto e matrizes de confusão (Seção 2.5).

## **2.2. REPRESENTAÇÃO DE CARACTERÍSTICAS TEXTURAIS**

Como citado na Seção 2.1, uma das etapas para criar o modelo de reconhecimento de padrões envolve a extração de características referentes aos objetos, as quais são utilizadas para treinar os indutores e para classificar objetos.

A escolha de um conjunto adequado de características para a discriminação entre as classes dos objetos que se deseja classificar está diretamente conectada ao domínio de aplicações e métricas (características) disponíveis (GUYON et al., 2006). Com isso, os vetores de características utilizados exercem grande impacto no desempenho do sistema de reconhecimento de padrões.

Alguns descritores de características muito utilizados são baseados em textura. Apesar de não haver consenso sobre a definição de textura, pode-se referenciá-la como sendo um conjunto de determinados padrões existentes nas superfícies físicas, perceptíveis ao olho humano, e que trazem grande quantidade de informações sobre a natureza da superfície, como suavidade, rugosidade e regularidade (PEDRINI; SCHWARTZ, 2007). Alguns exemplos de descritores de texturas são Padrões Locais Binários (*Local Binary Pattern* - LBP) (OJALA; PIETIKÄINEN, 1999), Quantização de Fase Local (*Local Phase Quantization* - LPQ) (OJANSIVU; HEIKKILÄ, 2008) e, mais recentemente, Características Robustas Aceleradas (*Speed-Up Robust Feature* -SURF) (BAY; TUYTELAARS; VAN GOOL, 2006).

Por se tratarem de metodologias relativamente recentes e por terem apresentados bons resultados em diversos trabalhos (ASLINA et al., 2013; JAMIL et al., 2015; JOLY et al., 2014; PHAM; GRARD, 2013; TSOLAKIDIS; KOSMOPOULOS; PAPADOURAKIS, 2014), a metodologia proposta neste documento utiliza, para a extração da textura, a Transformada de Características Invariantes à Escala (*Scale Invariant Feature Transform* - SIFT) (2.2.1),

Características Robustas Aceleradas (Speed-Up Robust Features - SURF) (2.2.2) e Histograma de Gradientes Orientados (*Histogram of Oriented Gradients* - HOG) (2.2.3).

### 2.2.1 TRANSFORMADA DE CARACTERÍSTICAS INVARIANTES À ESCALA - SIFT

A Transformada de Características Invariantes à Escala (*Scale Invariant Feature Transform* - SIFT) foi proposta por (LOWE, 1999) com o objetivo de extrair características e identificar regiões de interesse que permitam a comparação de imagens de objetos capturadas em diferentes perspectivas.

As regiões identificadas são representadas por pontos e são caracterizadas por diferenças máximas e mínimas da função Gaussiana aplicada no que o autor denomina espaço escala. Essas altas variações das regiões e escala garantem a estabilidade do descritor com relação à invariância, translação, escala e rotação, além da invariância parcial sobre as mudanças de luminosidade em uma imagem.

As principais etapas que compõe o SIFT para gerar a coleção de descritores são (LOWE, 1999):

**Detecção dos pontos de interesse:** Neste primeiro estágio é feita a procura por todas as escalas e posições de uma imagem. Esta é a parte mais custosa do algoritmo, pois é feita uma filtragem em cascata tomando uma mesma imagem sob diferentes escalas. A cada iteração são tomadas novas amostras dos pixels da imagem, as quais são geradas pela interpolação bilinear dos pontos identificadas em uma vizinhança de raio 1,5 da escala anterior; este processo garante a estabilidade do descritor, pois procura por características estáveis em todas as escalas possíveis, correlacionando uma região detectada em uma iteração com aquelas identificadas nas iterações anteriores.

O espaço escala de uma imagem é definido como uma função  $L(x, y, \sigma)$ , produzida pela convolução de uma variável escala Gaussiana  $G(x, y, \sigma)$  na imagem de entrada  $I(x, y)$ , conforme a equação (2.1)

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (2.1)$$

sendo  $*$  a operação de convolução em  $x$  e  $y$ , a gaussiana  $G(x, y, \sigma)$ , e  $\sigma$  é a escala definida conforme a equação (2.2)

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \quad (2.2)$$

Para uma eficiente detecção de posições das regiões de interesse, Lowe emprega a função diferença da Gaussiana convoluída na imagem,  $D(x, y, \sigma)$ , a qual pode ser computada pela diferença entre duas escalas próximas e separadas por uma constante multiplicativa  $k$ , conforme a equação( 2.3).

$$D(x, y, \sigma) = [G(x, y, k\sigma) - G(x, y, \sigma) * I(x, Y)] = L(x, y, k\sigma) - L(x, y, \sigma) \quad (2.3)$$

Para detectar os máximos e mínimos locais de  $D(x, y, \sigma)$ , cada ponto  $(x, y)$  é comparado com oito vizinhos na escala corrente e nove vizinhos nas escalas superior e inferior. Os pontos selecionados nesta fase obrigatoriamente são maiores (máximo) ou menores (mínimo) que todos os seus vizinhos.

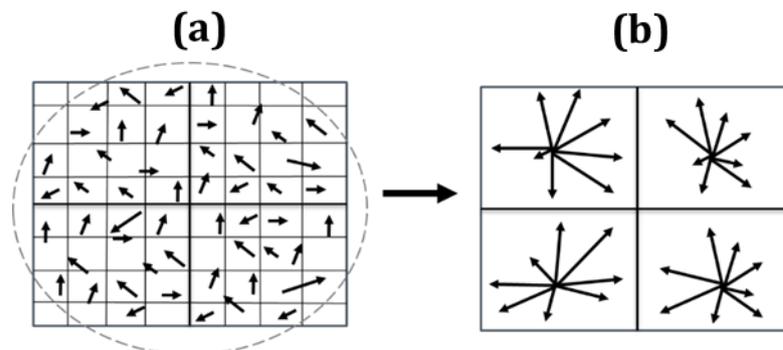
**Descrição dos pontos chaves:** A escala de um ponto selecionado é utilizada na escolha da imagem  $I(x, y)$  suavizada pela Gaussiana, a partir da qual todas as computações são realizadas de forma invariante a escala. Para cada ponto  $I(x, y)$  de uma imagem, considerando a escala já predefinida, são computados a magnitude  $m(x, y)$  e a orientação  $\theta(x, y)$  do gradiente por meio de diferenças entre pixels (Equações 2.4 e 2.5).

$$m(x, y) = \sqrt{(I(x+1, y) - I(x-1, y))^2 + (I(x, y+1) - I(x, y-1))^2} \quad (2.4)$$

$$\theta(x, y) = \tan^{-1}((I(x, y+1) - I(x, y-1))/(I(x+1, y) - I(x-1, y))) \quad (2.5)$$

A orientação do gradiente da região de interesse é obtida pela composição de um histograma de orientações dos pixels vizinhos a  $D(x, y, \sigma)$ , sendo cada pixel adicionado ao histograma é posteriormente ponderado pela magnitude do gradiente local e uma Gaussiana circular também ponderada representada por um círculo na Figura 2.2(a). A orientação final do gradiente é dada pelo maior valor do histograma.

Figura 2.2: Criação do descritor SIFT de pontos chaves: (a) obtenção dos gradientes para cada direção na região sobreposta; (b) aglomeração dos gradientes para cada sub-região.



Fonte: Adaptado de (LOWE, 1999).

Para fazer o cálculo dos descritores, os pontos  $L(x, y, \sigma)$  são ponderados pela Gaussiana identificada pelo círculo na Figura 2.2(a). Esses são então acumulados em um histograma de orientação que sumariza os conteúdos em 4x4 sub-regiões (Figura 2.2(b)), com o comprimento de cada aresta correspondente a soma das magnitudes dos gradientes que pertencem a aquela sub-região e que possuem a mesma direção.

Considerando os histogramas das 4x4 regiões com oito direções em cada, obtém-se um vetor de 128 elementos para cada ponto característico. Além disso, normalmente para cada imagem  $I$  pode haver milhares de pontos chaves, todos potencialmente candidatos a compor o conjunto de característica que descreve  $I$ .

### 2.2.2 CARACTERÍSTICAS ROBUSTAS ACELERADAS - SURF

Características Robustas Aceleradas (*Speed-Up Robust Features* - SURF), proposto por (BAY; TUYTELAARS; VAN GOOL, 2006), é um descritor baseado no SIFT, que busca categorizar uma imagem detectando e descrevendo regiões de interesse.

Uma das vantagens desse descritor está na velocidade de processamento em relação ao SIFT. Em (BAY et al., 2008) os autores afirmam que SIFT apresenta melhor desempenho e consegue ser robusto a pequenas deformações e erros de localização das regiões detectadas. No entanto SIFT tem uma alta dimensionalidade que dificulta seu uso em aplicações *on-line*, por conta de seu alto custo computacional (BAY; TUYTELAARS; VAN GOOL, 2006). Para dar suporte a essa afirmação, os autores propuseram um descritor que utiliza a metade do número

de elementos do vetor de SIFT, e é baseado em matrizes Hessiana e no conceito de imagens integrais, que apresentam boa performance em termos de tempo computacional e acurácia (BAY; TUYTELAARS; VAN GOOL, 2006; BAY et al., 2008).

Os pontos de interesse detectados são regiões que tem variações de valores nos pixels da imagem, e para detectar tais variações, é utilizado a matriz Hessiana (Equação 2.6) que por definição é dado o ponto  $(x, y)$  na imagem  $I$  sobre uma escala  $\sigma$ , que consiste na matriz Hessiana  $H(x, y, \sigma)$  juntamente a função Gaussiana de Laplaciana  $L_{xx}(x, y, \sigma)$ . Por fim, o determinante da mesma indica se há variações dos valores dos pixels em uma determinada região da imagem, indicando um ponto de interesse.

$$H(x, y, \sigma) = \begin{bmatrix} L_{xx}(x, y, \sigma) & L_{xy}(x, y, \sigma) \\ L_{xy}(x, y, \sigma) & L_{yy}(x, y, \sigma) \end{bmatrix} \quad (2.6)$$

Uma aproximação do determinante da matriz Hessiana é obtida na equação 2.7, onde  $D_{xx}, D_{yy}$ , e  $D_{xy}$  são as derivadas parciais de segunda ordem da Gaussiana em suas direções, enquanto  $w$  representa o peso usado para preservar a energia entre as diferentes escalas da Gaussiana.

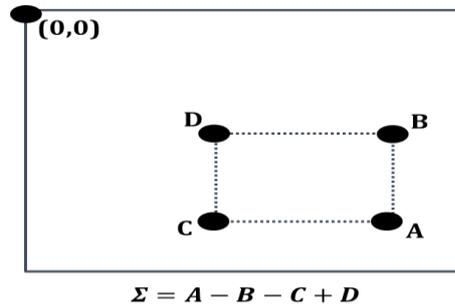
$$\det(H_{approx}) = D_{xx}D_{yy} - (wD_{xy})^2 \quad (2.7)$$

O uso do conceito de imagens integrais faz com que todos os valores dos pixels de uma imagem sejam somados formando uma matriz onde, para cada posição, está a soma de todos os valores das posições anteriores. Nesse contexto, dada uma matriz  $M$ , a posição  $M_{2,2}$  contém a soma dos valores das posições correspondentes aos pixels:  $I_{1,1}, I_{1,2}, I_{2,1}, I_{2,2}$  de uma imagem  $I$ . Portanto, o valor de uma posição é dado pela equação 2.8:

$$M_{x,y} = \sum_{i=0, j=0}^{i < x, j < y} I_{i,j} \quad (2.8)$$

Diante disso, para o cálculo da soma dos pixels contidos em uma região qualquer é necessário aplicar apenas 3 operações básicas de adição (Figura 2.3).

Figura 2.3: Exemplo de uso de imagem Integral: Ilustração do cálculo da soma dos pixels contidos em uma sub-região DBCA da imagem com apenas três operações de adição.

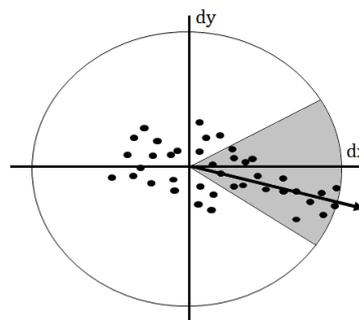


Fonte : (BAY et al., 2008).

SURF considera a derivada de primeira ordem da *wavelet* de Haar para representar as distribuições dos níveis de cinza na vizinhança dos pontos de interesse  $(x, y)$ . As respostas são representadas como pontos no espaço (Figura 2.4), após serem multiplicadas por pesos definidos pela função gaussiana centrada no ponto de interesse.

Esses pontos no espaço são acumulados considerando as direções horizontais e verticais e uma janela deslizante de orientação que abrange um ângulo  $\pi/3$ . Por fim, considerando todas as possibilidades geradas durante o deslocamento da janela deslizante, o maior valor (vetor) determina a orientação do ponto de interesse (BAY et al., 2008).

Figura 2.4: Janela deslizante: a variação dos pixels de uma região em volta de um ponto de interesse

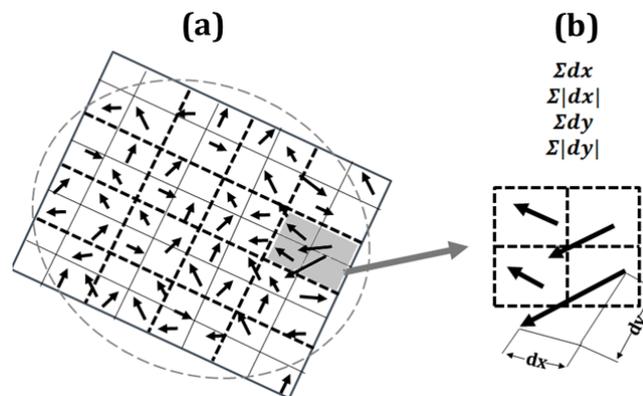


Fonte: (BAY; TUYTELAARS; VAN GOOL, 2006).

Definida uma região de interesse, o próximo passo consiste em calcular os valores dos elementos que compõe o vetor de características do descritor. Para isso, sobre o ponto de interesse define-se uma região quadrada  $(8 \times 8)$  que é alinhada com a orientação previamente

identificada (Figura 2.5). A partir de cada sub-região são obtidos 4 valores formando um total de 64 elementos que correspondem a cada posição do vetor de descrição. Para cada um dos 64 elementos, obtém-se a resposta da *wavelet* de Haar, nas direções  $x$  e  $y$ . Para cada sub-região  $2 \times 2$ ,  $d_x$ ,  $d_y$ ,  $|d_x|$  e  $|d_y|$  são acumuladas separadamente, gerando 16 conjuntos com as quatro características  $\{(\sum d_x, \sum d_y, \sum |d_x|, \sum |d_y|)\}$  em um total de 64 elementos. Todos os pontos característicos encontrados são invariantes a translação, escala e rotação, além de ser parcialmente invariante a mudanças de iluminação (BAY et al., 2008).

Figura 2.5: Distribuições de Haar Wavelets: (a) aplicado a máscara  $8 \times 8$  com as orientações identificadas e obtenção do gradiente para cada direção na região sobreposta da imagem; (b) acúmulo do gradiente para cada região.



Fonte : (BAY et al., 2008)

### 2.2.3 HISTOGRAMA DE GRADIENTES ORIENTADOS - HOG

O Histograma de Gradientes Orientados foi proposto por (DALAL; TRIGGS, 2005). Esse descritor é considerado similar ao descritor proposto por (LOWE, 1999). Entretanto, ao contrário de SIFT, que calcula os histogramas de gradientes em volta dos pontos chaves, o descritor HOG calcula os histogramas de gradientes em grades de células uniformemente espaçadas sobrepostas a imagem. A extração do descritor HOG consiste em quatro passos.

No primeiro passo é feito o cálculo de gradientes de primeira ordem da imagem. Esse cálculo é responsável pela captura de contornos, silhuetas e textura da imagem. As máscaras espaciais em  $x$  e  $y$  são dadas por  $[-1 \ 0 \ 1]$ , para o cálculo do gradiente na horizontal e por  $[-1 \ 0 \ 1]^t$ , para o cálculo do gradiente na vertical, respectivamente.

O segundo passo consiste em criar os histogramas de orientações de gradiente, que tem por objetivo reduzir e ser resistente a pequenas variações de poses e aparência. Para isso, uma imagem inteira é dividida em pequenas regiões espaciais de 8x8 pixels, chamadas de células. Cada célula então é formada por 64 pixels, e um histograma de orientações é criada para cada um desses pixels, formando um histograma por célula. Cada histograma gera 9 divisões (bins), espaçados entre 0 a 180 graus. A contribuição de cada pixel para os histogramas é ponderada pelo valor da magnitude do gradiente na posição do pixel.

O terceiro passo trata da normalização de histogramas, usando a normalização de contraste local, considerando que vão existir problemas de iluminação e sombreamento nas imagens do mundo real. Para executar a normalização, as células são aglomeradas em blocos e cada bloco é normalizado de forma separada. Cada bloco é formado por um conjunto de 2x2 ou 16x16 pixels. O descritor final consiste em um vetor com todos os componentes normalizados das células de todos os blocos. A normalização é dada usando a regra L2-Hys que limita os valores máximos do vetor a 0,2. Na equação 2.9 é calculado a regra L2-norm.

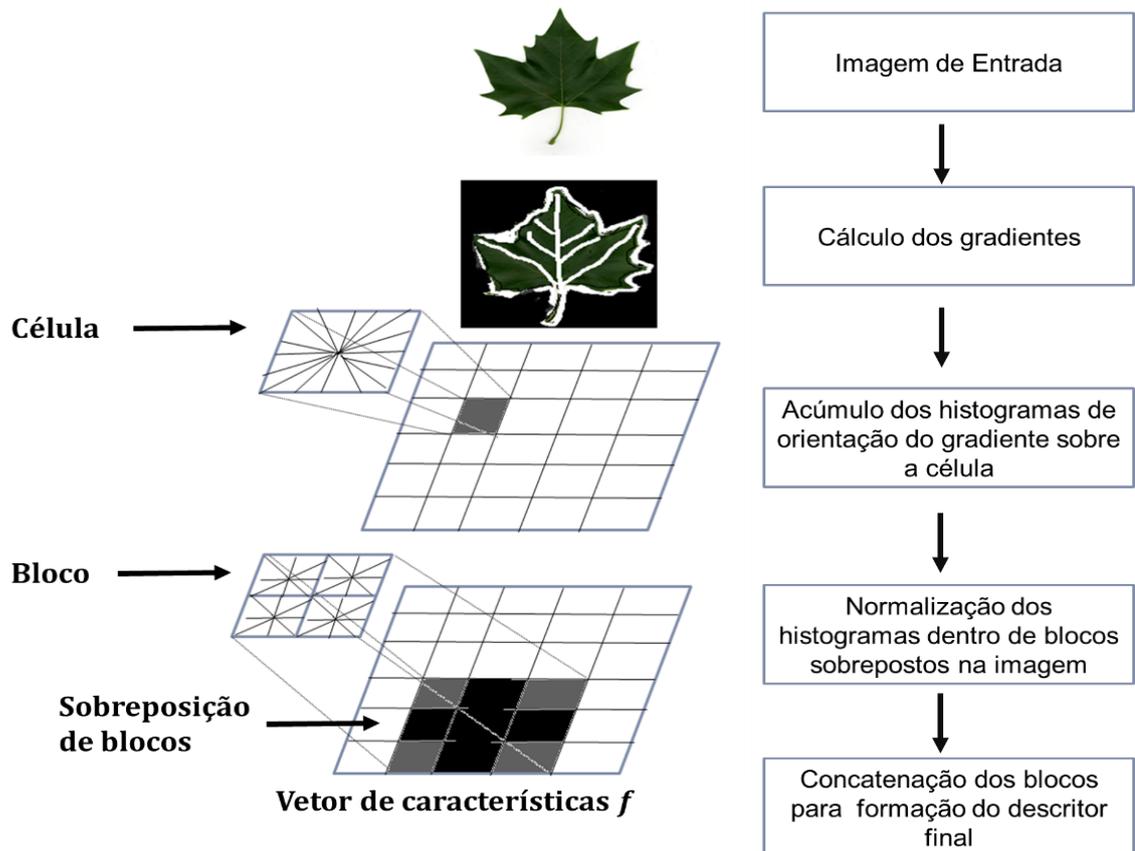
$$L2 - norm: v \rightarrow v / \sqrt{\|v\|_2^2 + \epsilon^2} \quad (2.9)$$

onde  $v$  é o descritor não normalizado,  $\|v\|$  sua k-ésima norma utilizada para  $k=1,2$ , e  $\epsilon$  é uma constante usada para evitar uma possível divisão por zero em locais da imagem onde não existe a variação da orientação do gradiente.

O passo final consiste em concatenar todos os blocos encontrados na janela de detecção em um único vetor, após sua normalização, criando assim o vetor descritor final. Na janela de detecção usada neste trabalho, de tamanho 128x64 pixels, contém 15x7 blocos. Cada bloco é formado por 2x2 células, que por sua vez possuem um histograma de 9 bins. Portanto, o descritor final utilizado no método proposto possui uma dimensão igual a  $15 \times 7 \times 2 \times 2 \times 9 = 3.780$ , ou seja, 3.780 elementos.

Foi mostrado em (DALAL; TRIGGS, 2005) que a sobreposição dos blocos, fazendo com que a mesma célula seja normalizada em relação a mais de um bloco, aumenta o desempenho do descritor, no entanto, traz a desvantagem do aumento da dimensão do mesmo. O fluxograma exemplificando o cálculo do descritor pode ser visto na Figura 2.6.

Figura 2.6: Uma visão global do cálculo do descritor HOG. A imagem inteira é preenchida com uma grade de blocos sobrepostos. Cada bloco é formado por um conjunto de células, nas quais são calculados histogramas da orientação do gradiente de cada pixel presente nas mesmas. Os histogramas são normalizados localmente em relação aos blocos e concatenados para a criação do vetor de características.



Fonte: (DALAL, 2006).

## 2.3. CLASSIFICADORES

Para classificar as espécies de plantas descritas pelos extratores de atributos foram analisados quatro tipos de classificadores: Otimização Sequencial Mínima (*Sequential Minimal Optimization* - SMO), Redes Neurais Artificiais (*Artificial Neural Network* - RNA), K-vizinhos mais próximos (*K-Nearest Neighbors* - KNN) e Árvores Aleatórias (*Random Forest* - RF).

### 2.3.1 OTIMIZAÇÃO SEQUENCIAL MÍNIMA - SMO

O método de Otimização Sequencial Mínima (SMO), proposto por (PLATT, 1999), consiste em um algoritmo de aprendizagem, que pode ser considerado uma variante do método

Máquina de Vetor de Suporte (*Support Vector Machine* - SVM). Esse algoritmo tem como diferencial fazer a decomposição do problema de minimização quadrática em vários subproblemas de programação quadrática (PQ) de tamanho fixo.

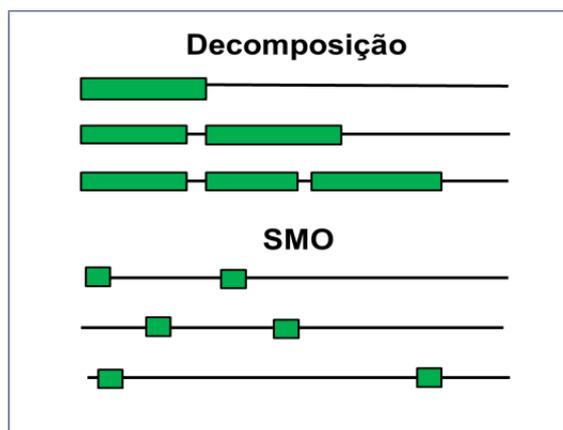
Na resolução de PQ em SVM, é introduzido o método chamado de *chunking*. O *chunking* consiste em remover os exemplos que não tem influência na solução ( $\alpha_i = 0$ ), e para isso é separado o problema PQ em subproblemas menores. A dimensão da matriz do problema de otimização é assim reduzida, passando do número de exemplos de treino para o número de vetores de suporte. No entanto, se forem usadas técnicas usuais de PQ, o *chunking* não consegue lidar com problemas de grande escala, visto que mesmo uma matriz reduzida pode não caber em memória com a grande quantidade de exemplos de treinamento.

Por conta disso, o SMO difere da técnica de SVM agilizando na resolução do menor problema de otimização possível. Assim, em cada passo, o algoritmo SMO escolhe dois exemplos de multiplicadores  $\alpha_i$  para otimizar, encontrando assim os valores ótimos para essas variáveis e atualizando depois as SVMs.

A grande vantagem está na resolução analítica do problema, evitando a otimização segundo métodos numéricos de PQ. Existem assim duas componentes principais do algoritmo SMO, o método analítico para calcular os dois alfas e a heurística de escolha dos multiplicadores a otimizar.

A Figura 2.7, ilustra os dois métodos para o treinamento de SVM citados até aqui. São ilustrados três passos para cada método. As linhas horizontais representam o conjunto de treinamento em cada etapa, enquanto que, as caixas representam os multiplicadores  $\alpha_i$  otimizados em cada iteração.

Figura 2.7: Métodos de Treinamento SVM: Decomposição (*Chunking*) e SMO.



Fonte: Adaptado de (PLATT, 1999).

### 2.3.2 REDES NEURAIS ARTIFICIAIS - RNA

Redes Neurais Artificiais são sistemas computacionais, formado por processos, denominados neurônios, que são conectados entre si (HAYKIN, 1999). Seu método é semelhante ao cérebro humano, onde o conhecimento obtido pela RNA ocorre por meio do processo de aprendizado. Fazendo uma analogia simples, o conhecimento de um cérebro humano é baseado em aprender através de exemplos, ou seja, quanto mais o ser humano vê alguma coisa, maior é a sua capacidade de aprendizado.

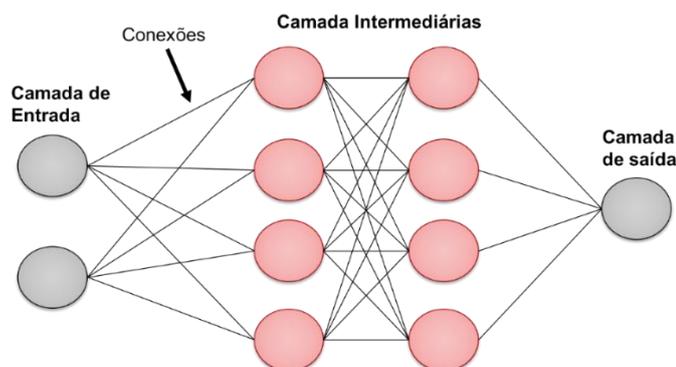
A arquitetura de uma RNA é muito importante devido ao fato de que sua composição depende do problema a ser tratado pela rede. Ainda mais, a arquitetura da rede está também relacionada ao algoritmo de aprendizagem usado para treinamento (BRAGA; CARVALHO; LUDEMIR, 2000).

Por se tratar de um problema não linear, quando considerado diversas classes a serem abordadas, neste trabalho a arquitetura de rede utilizada foi de múltiplas camadas com o algoritmo de aprendizagem de retro-propagação (*Backpropagation*).

Redes de múltiplas camadas são aquelas que apresentam pelo menos uma camada intermediária e seu principal objetivo está em tratar dados não linearmente separáveis.

Pode-se entender uma rede de múltiplas camadas sendo aquele que possui três tipos básicos de camadas: a camada de entrada, que é responsável por apresentar os padrões de entrada à rede e não realiza processamento algum; a camada oculta, que funciona como extratores de características, com a codificação interna destas características traduzindo-se no ajuste dos pesos sinápticos, que são constituídas de nós computacionais; e a camada de saída, que tem a função de construir os padrões de respostas da rede (Figura 2.8).

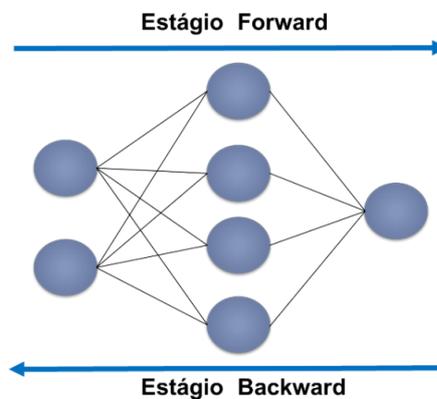
Figura 2.8: Funcionamento de um Rede neural com duas camadas intermediárias.



Fonte: Adaptado de (HAYKIN, 1999).

Para treinamento de redes de múltiplas camadas é utilizado um algoritmo que se baseia na regra de aprendizagem por correção de erro, denominado algoritmo de retro-propagação (*Backpropagation*). No algoritmo de retro-propagação o treinamento ocorre em dois estágios, e em cada estágio a rede é percorrida em diferentes sentidos. Esses dois estágios são chamadas de estágio *forward* e estágio *backward*. O estágio *forward* é utilizado para definir a saída da rede para um dado padrão de entrada e o estágio *backward* utiliza a saída desejada e a saída fornecida pela rede para atualizar os pesos de suas conexões (Figura 2.9).

Figura 2.9: Fluxo do processamento do algoritmo *Backpropagation*



Fonte: Adaptado de (HAYKIN, 1999).

### 2.3.3 K-VIZINHOS MAIS PRÓXIMOS - KNN

O algoritmo de classificação KNN é um método de aprendizado baseado em instâncias. Esse tipo de aprendizado se torna diferente da maioria dos métodos convencionais pelo fato do aprendizado consistir no armazenamento dos exemplos de treinamento.

Esse algoritmo de classificação, proposto por (MITCHELL, 1997), assume que todas as instâncias correspondem a pontos em um espaço n-dimensional. Quando uma nova amostra é exposta para ser classificada, é feita uma verificação para buscar um conjunto de amostras similares, logo após, este conjunto é recuperado do conjunto de treinamento e usado para classificar a nova amostra.

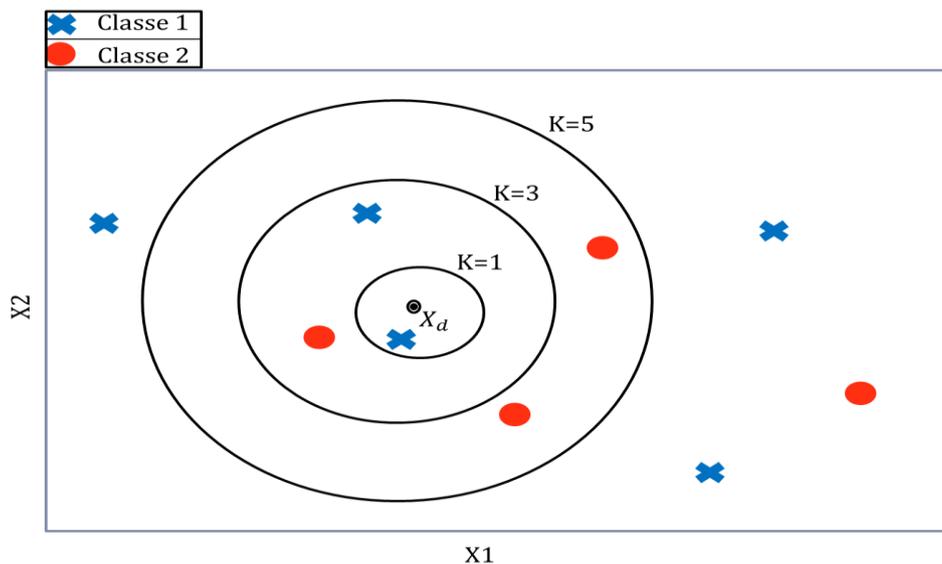
As amostras similares são as que possuem a menor distância no espaço n-dimensional e por isso este método é conhecido como “Vizinho mais próximo” (*Nearest Neighbor*). Mais especificadamente para predizer uma classe de um novo padrão X, KNN calcula os k-vizinhos

mais próximos para este novo padrão de entrada  $X$  e classifica-o como sendo da classe que aparece mais vezes dentre os seus  $k$ -vizinhos.

O KNN possui um parâmetro chamado  $K$ , que indica o número de vizinhos que se serão usados pelo algoritmo durante a fase de teste. Esse valor é definido para que se obtenha um valor mais acurado. No entanto, o valor ótimo de  $K$  varia para cada problema, o que faz com que para cada problema tratado sejam testados vários valores de  $K$  diferentes para descobrir o melhor valor de  $K$ .

Um exemplo de funcionamento do KNN pode ser visto na Figura 2.10, onde se tem um padrão desconhecido  $X_d$  entre os padrões da classe 1 e 2. O processo do KNN é predizer o padrão  $X_d$  sendo pertencente a uma dos padrões de classe exibidas no exemplo. Dependendo do número de  $K$ -vizinho,  $X_d$  será classificado como: Se  $K$  for igual a 1 ( $K=1$ ),  $X_d$  será classificado como “Classe 1”, se  $K=3$ , então  $X_d$  será classificado como “Classe 1”, por fim se  $K=5$ ,  $X_d$  será classificado como “Classe 2”.

Figura 2.10: Exemplo do funcionamento do algoritmo KNN



Fonte: O autor.

### 2.3.4 ÁRVORES ALEATÓRIAS - RF

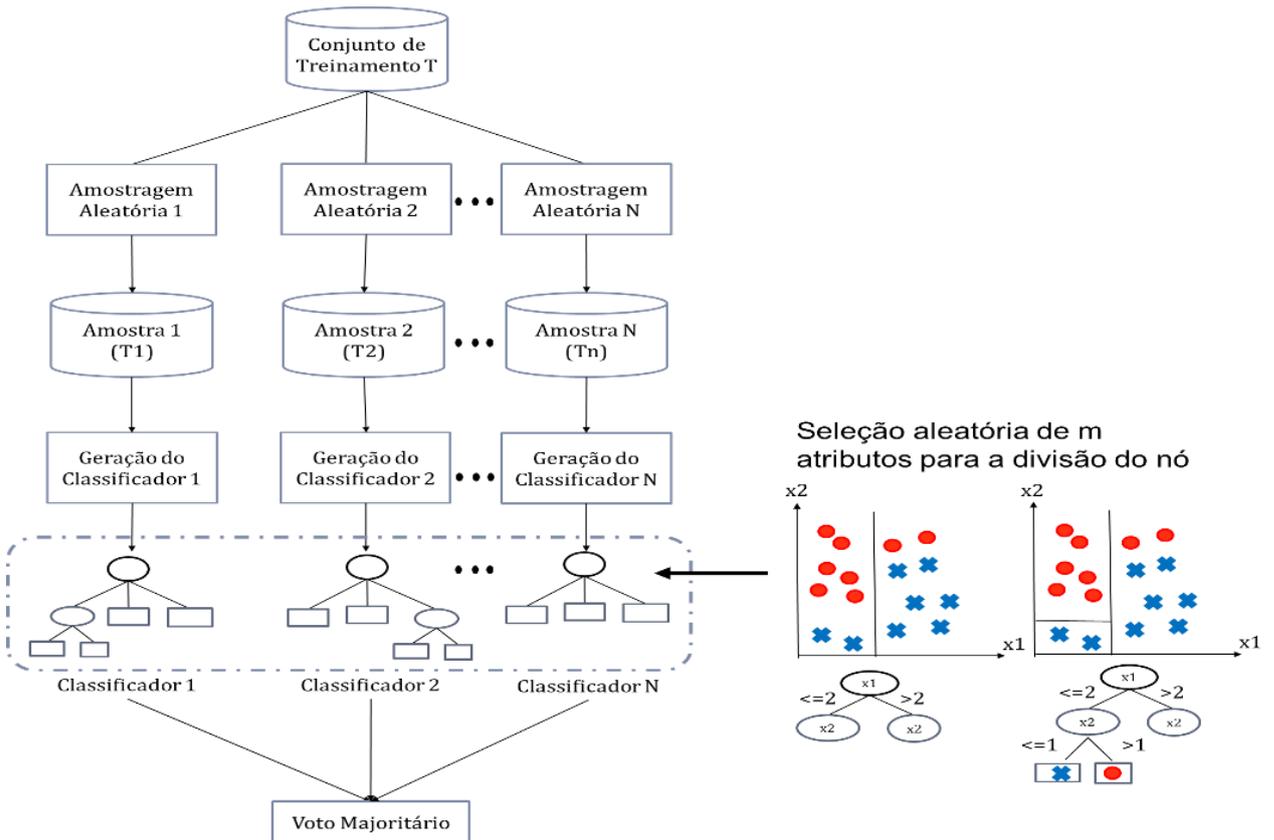
O algoritmo de árvores aleatórias (*Random Florest* – RF) consiste na geração de vários conjuntos de árvores de decisões dentro do mesmo conjunto. Cada conjunto de árvores passa por um mecanismo de votação chamado *bagging*, que elege a classificação mais votada, para produzir sub-amostras aleatórias de conjuntos de treinamento.

Mais especificamente, o *random florest* foi proposto por (BREIMAN, 2001), e consiste em um classificador construído por uma coleção de árvores  $\{h_k(x)\}, k = 1, 2, \dots, n$ , onde  $T_n$  são amostras aleatórias distribuídas e cada árvore vota na classe que mais aparece na entrada  $x$ .

A árvore é construída sempre usando um novo subconjunto e uma seleção aleatória de atributos. A cada nó da árvore, um subconjunto de  $m$  atributos é avaliado e selecionado aleatoriamente. O melhor atributo é, então, escolhido para dividir o nó. O valor de  $m$  é fixado para todos os nós e as árvores crescem sem poda.

O funcionamento da árvore de decisão é exibido na Figura 2.11. Note que a árvore aleatória difere no funcionamento do método *bagging* na parte tracejada da figura, uma vez que as árvores são construídas selecionando  $m$  atributos aleatoriamente para cada nó das árvores, o que não acontece no *bagging* em que todos os atributos são usados.

Figura 2.11: Exemplo do funcionamento de uma Árvore Aleatória



Fonte: O autor.

## 2.4. COMBINAÇÃO DE CLASSIFICADORES

A principal função de um classificador é prever uma amostra padrão a uma determinada classe. Cada classificador possui suas peculiaridades distintas de decisão, ou seja, são processados de formas diferentes, e a seleção se faz por aquele que representa a melhor generalização dado um conjunto de validação. No entanto, essa estratégia (escolha de somente um classificador) pode ocasionar perda de informações importantes daqueles que foram ignorados.

Para resolver esse problema, a predição (saída) de cada classificador pode ser combinada com os restantes classificadores, aprimorando sua capacidade de generalização sobre um determinado problema, e deixando a classificação mais estável, já que refina os pontos fortes de cada um deles.

Para algumas aplicações, apenas um único classificador consegue ter um bom desempenho ao ponto de resolver determinado problema. Porém, quando se depara com problemas de maior complexidade, por exemplo, com uma grande quantidade de classes distintas ou conjuntos de atributos com dimensionalidade e características diferentes, a escolha de um único classificador pode não resolver o problema, pois possivelmente a capacidade de reconhecimento seria insuficiente e, nestes casos, a combinação é aconselhada.

Segundo (SUROWIECKI, 2005), existem alguns critérios que devem ser seguidos para que o conjunto de tomadores de decisões seja robusto:

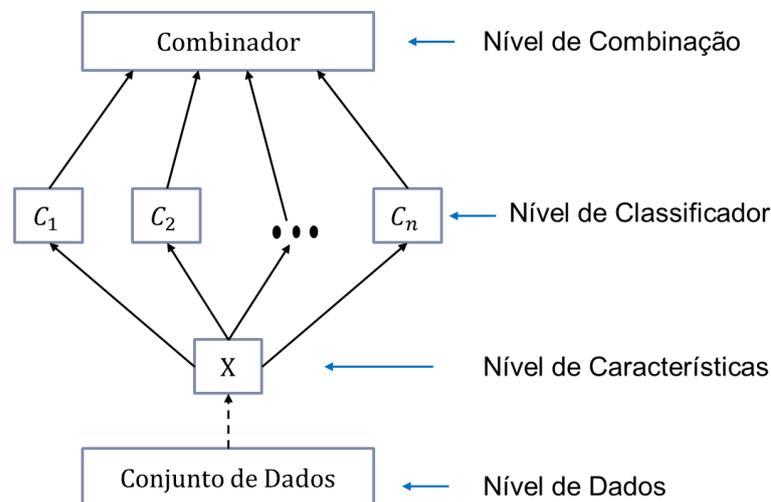
- Independência: cada membro do conjunto não pode ter sua resposta afetada pela opinião dos outros;
- Descentralização: cada membro deve ser capaz de se especializar e extrair informações de determinado conhecimento local;
- Agregação: uma técnica é responsável por tornar os julgamentos individuais em uma decisão coletiva;
- Diversidade de opiniões: cada membro deve ter pelo menos uma fração da informação do problema que está sendo tratado, mesmo que esta seja uma interpretação incomum dos fatos.

Um sistema de combinação de diversos classificadores é um novo classificador que tem como entrada as respostas de  $n$  classificadores, e gera como saída, uma decisão final, que pode ser em três níveis: abstração, *ranking* e medida de confiança.

No nível de abstração, cada classificador  $C_i$  gera um rótulo da classe sem dar informações sobre a certeza desta predição, nem dando alternativa a esta classificação. No *ranking*, a saída  $C_i$  é um subconjunto de escolhas preditas, cujas alternativas são ranqueadas em uma ordem possível de ser a rotulação correta, sendo bastante usadas para problemas com um grande número de classes. Por fim, no nível de medida de confiança, para cada  $C_i$  é produzido um vetor dimensional em que representa probabilidade de que o vetor  $X$ , submetido para a classificação, vem da classe em questão.

A combinação de classificadores tem sido usada em diversos trabalhos com o objetivo de melhorar os resultados obtidos no reconhecimento de padrões, sem que exista um aumento na complexidade do método, ou esta melhoria é obtida por meio da complementariedade entre eles. Mais especificamente, para obter essa complementariedade entre classificadores, deve ser estabelecido um conjunto de combinação, como é exibido na Figura 2.12, em que se pode variar combinadores (nível de combinação), classificadores individuais (nível de classificador), atributos (nível de características) e subconjuntos de dados (nível de dados).

Figura 2.12: Diferentes abordagens para a construção de conjuntos em combinação de classificadores



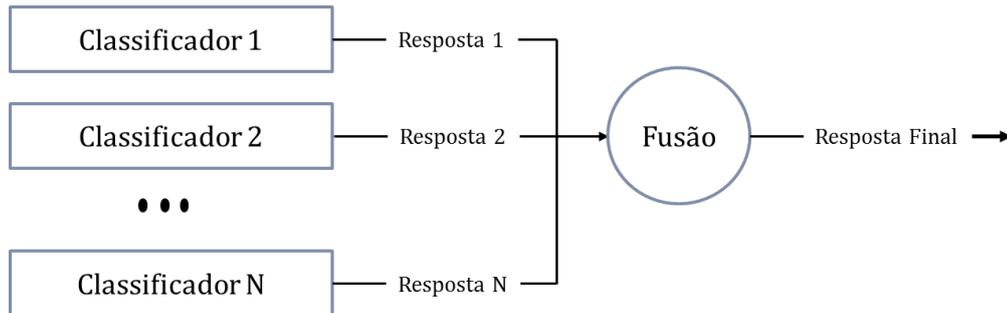
Fonte: O autor.

Neste trabalho, para fazer a combinação dos classificadores são utilizados o nível de características e nível de classificador, com o tipo de saída *ranking*. No nível de características são utilizados três descritores com características de textura e no nível de classificador, quatro classificadores distintos.

### 2.4.1 MÉTODOS DE FUSÃO

Os métodos de fusão classificam os objetos por meio da combinação das predições de diversos classificadores, aplicando uma função de fusão em suas predições, como ilustrado na Figura 2.13.

Figura 2.13: Arquitetura de fusão das respostas de diversos classificadores



Fonte: O autor.

Nesta Seção são apresentados algumas das regras de fusão mais comuns, presentes no trabalho de (KITTLER et al., 1998).

#### Voto Majoritário

Uma das estratégias mais simples de unir as respostas dos classificadores é por meio do voto majoritário das respostas. Essa regra estabelece que a classe de determinado objeto é a classe que mais recebeu votos (resposta final) dentre todos os classificadores. Caso haja empates no método de votação, a classe final pode ser predita por meio de uma seleção aleatória entre as classes empatada (KUNCHEVA, 2004)

Em um problema contendo  $z$  classes possíveis e  $w$  classificadores, a classe de um objeto  $x$ , dada pela votação majoritária, pode ser definida como:

$$classe(x) = \underset{k = 1}{\operatorname{arg\,max}} \sum_{i=1}^w g(y_i, k) \quad (2.10)$$

Em que  $y_i$  é o rótulo ao objeto pelo  $i$ -ésimo classificador, e  $g(y, k)$  é uma função indicativa definido como :

$$g(y, k) = \begin{cases} 1 & \text{se } y = k \\ 0 & \text{se } y \neq k \end{cases} \quad (2.11)$$

Às vezes, as respostas dos classificadores pode ser um número contínuo que representa probabilidade (*score*) de um objeto  $x$  pertencente à classe  $k$ . Considere  $dy, k$  o *score* dado pelo classificador  $y$  que  $x$  pertence a classe  $k$ , em um problema contendo  $n$  classificadores e  $c$  classes. Os métodos de fusão, apresentados a seguir, baseiam-se nesse *score* para definir a classe de um objeto.

### Regra do Produto

A regra do produto faz a multiplicação dos *scores* gerados pelos classificadores, tornando esta regra sensível aos baixos *scores*, pois um *score* próximo a 0 dado por um dos classificadores para determinada classe pode anular a chance de ela ser escolhida como a classe objeto predita. A regra produto é definida por:

$$classe(x) = \underset{k = 1}{arg \max} \prod_{i=1}^w dy_{i,k} \quad (2.12)$$

### Regra da Soma

A regra da soma faz a somatória dos *scores* gerados pelos classificadores e é definida por:

$$classe(x) = \underset{k = 1}{arg \max} \sum_{i=1}^w dy_{i,k} \quad (2.13)$$

### Regra da Média

A regra da média simplesmente realiza uma média dos *scores* atribuídos pelos classificadores, como definido na equação 2.14.

$$classe(x) = \underset{k = 1}{arg \max} \frac{1}{w} \sum_{i=1}^w dy_{i,k} \quad (2.14)$$

### Regra da Mediana

A regra da mediana consiste em fazer a mediana dos *scores* gerados pelos classificadores, como resultado da classificação combinada, e é definida pela equação 2.15.

$$classe(x) = \underset{k = 1}{arg \max} \underset{i = 1}{mediana} dy_{i,k} \quad (2.15)$$

### Regra do Máximo

A regra do máximo é realizada assumindo como resultado da combinação o maior *score* dado dentre os classificadores, como visto na equação 2.16.

$$classe(x) = \underset{k=1}{\overset{z}{arg\ max}} \underset{w}{\overset{i=1}{m\acute{a}ximo\ dy_{i,k}}} \quad (2.16)$$

### Regra do Mínimo

Na regra do mínimo, o resultado da combinação é o valor mínimo encontrado dentre os *scores* de cada classificador atribuídas a classes. Essa regra, como a regra do produto, é sensível a baixos *scores* e é definida por:

$$classe(x) = \underset{k=1}{\overset{z}{arg\ max}} \underset{w}{\overset{i=1}{m\acute{í}nimo\ dy_{i,k}}} \quad (2.17)$$

## 2.5. MÉTRICAS DE AVALIAÇÃO DE CLASSIFICADORES

Para avaliar os resultados obtidos é necessário definir algumas medidas de avaliação, onde o modelo de classificação treinado é posto à prova, a fim de investigar seu desempenho em uma base de dados de validação. No entanto, a taxa de acertos obtida sem nenhuma métrica pode não representar precisamente o quão apropriado é o classificador. Esse problema é maior quando é preciso comparar os resultados alcançados utilizando classificadores diferentes.

Nas próximas subseções são apresentadas algumas métricas utilizadas para a investigação do desempenho e na comparação de classificadores.

### 2.5.1 MATRIZES DE CONFUSÃO E SUAS DERIVAÇÕES

Muitos problemas de classificação normalmente são binários, ou seja, de duas classes, que são mapeadas como pertencentes à classe Positiva (Classe P - Correto) ou negativa (Classe N - Errado). Uma matriz de confusão, para esses problemas, mostra as informações sobre os verdadeiros positivos (VP) e falsos positivos (FP), bem como os verdadeiros negativos (VN) e falsos negativos (FN) originado pelo classificador. Um exemplo de Matriz de Confusão de duas classes (binário) pode ser visto na Tabela 2.1

Tabela 2.1: Matriz de Confusão para problemas de duas classes

	<b>P</b>	<b>N</b>
<b>P</b>	Verdadeiro Positivo (VP)	Falso Negativo (FN)
	Falso Positivo (FP)	Verdadeiro Negativo (VN)

Para o caso de identificação de espécies de plantas, que exige mais de duas classes, o problema é ainda maior, tendo em vista que as instâncias podem ser corretamente ou incorretamente classificadas em relação a qualquer classe.

Com  $x$  classes, a matriz de confusão de múltiplas classes se transforma em uma matriz contendo os  $y$  resultados do classificador. A tabela 2.2 apresenta a matriz de confusão para problema de múltiplas classes, aonde  $C$  representa a classe e  $W$  a amostra pertencente a classe.

Tabela 2.2: Matriz de Confusão para problemas de múltiplas classes

	<b>Predição <math>c_1</math></b>	<b>Predição <math>c_2</math></b>	<b>Predição <math>c_{...}</math></b>	<b>Predição <math>c_n</math></b>
<b>Classe <math>c_1</math></b>	$W_{11}$	$W_{12}$	$W_{...}$	$W_{1n}$
<b>Classe <math>c_2</math></b>	$W_{21}$	$W_{22}$	$W_{...}$	$W_{2n}$
<b>Classe <math>c_{...}</math></b>	$W_{...}$	$W_{...}$	$W_{...}$	$W_{...}$
<b>Classe <math>c_n</math></b>	$W_{1n}$	$W_{2n}$	$W_{...}$	$W_n$

Na matriz de confusão existe uma distribuição entre as classes e a ligação entre as linhas. Dessa forma, qualquer alteração na medida de desempenho obtida por uma classe que utilize valores de ambas as colunas, esta será necessariamente sensível à desproporção de valores entre as classes.

Com a matriz de confusão, além de ser possível ver a variação e a acurácia de cada classe pertencente ao conjunto de classes, também se pode derivar e calcular algumas métricas quantitativas (FAWCETT, 2006). Neste trabalho as métricas adotadas são:

- Taxa de acertos (TA)

$$TA = \frac{VP + VN}{P + N} \quad (2.18)$$

- Taxa de Verdadeiro Positivo - *True Positive Rate* (TPR)

$$TPR = \frac{VP}{P} \quad (2.19)$$

- Taxa Verdadeiro Negativo - *True Negative Rate* (TNR)

$$TNR = \frac{VN}{N} \quad (2.20)$$

- Taxa de Falso Positivo - *False Positive Rate* (FPR)

$$FPR = \frac{FP}{N} = -TNR \quad (2.21)$$

### 2.5.2 MÉTRICA IMAGECLEF

A fim de fazer uma comparação com os trabalhos relacionados expostos na seção 3.3, que utilizaram da mesma base de dados que este trabalho, uma métrica é avaliada para fins de comparação com uma aplicação ao mundo real no processo de identificação de plantas, onde o reconhecimento pode ser feito em qualquer área e, em diferentes estágios das plantas e observadas por qualquer usuário, é definida pela equação 2.22 (GOËAU et al., 2011).

$$S = \frac{1}{U} \sum_{u=1}^U \frac{1}{P_u} \sum_{p=1}^{P_u} \frac{1}{r_{u,p}} \quad (2.22)$$

onde  $U$  é o número de usuários que captura uma foto de uma folha de planta,  $P_u$  é o número de plantas individuais observadas pelos  $U$  usuários (de qual planta o usuário tirou a foto),  $r_{u,p}$  é a lista de espécies corretas retornadas pelo sistema proposto (levando em consideração cada  $P_u$  observada pelo  $U$ ).

Essa métrica é ponderada pelo fato de que o conjunto de dados ImageCLEF (GOËAU et al., 2011) foi construída de forma colaborativa, de modo que alguns contribuintes podem ter fornecido muito mais imagens do que outros colaboradores. Dessa forma, para avaliar a capacidade de um sistema para fornecer respostas corretas para todos os usuários, é feita uma acurácia média para as imagens de cada usuário, ao em vez de medir apenas a taxa média de classificação para todos os usuários.

Além disso, alguns usuários, por exemplo, fornecem muitas fotos da mesma planta individual (para enriquecer os dados de treinamento com menos esforço). Para resolver esse problema, é feita a média da taxa de classificação para cada planta individual, com a intenção de avaliar a capacidade de um sistema para fornecer a resposta correta com base em uma única observação da planta.

### 3. ESTADO DA ARTE

Neste capítulo são apresentadas as abordagens e ferramentas utilizadas por diversos autores para enfrentar o problema de classificação automática de plantas. Os trabalhos foram divididos em duas seções, sendo que a Seção 3.1 é composta por vários trabalhos que aplicam técnicas de extração de características da folha e classificação de imagens, a partir das mais variadas metodologias e diferentes protocolos de imagens. Na Seção 3.2 são apresentados os quatro melhores métodos avaliados apresentados no congresso ImageCLEF em 2011. O ImageCLEF é considerado evento que propõe o maior desafio para a contribuição e evolução de pesquisas na área de identificação automática de plantas, e conta com a participação de pesquisadores do mundo inteiro.

#### 3.1 SISTEMAS DE RECONHECIMENTO DE PLANTAS

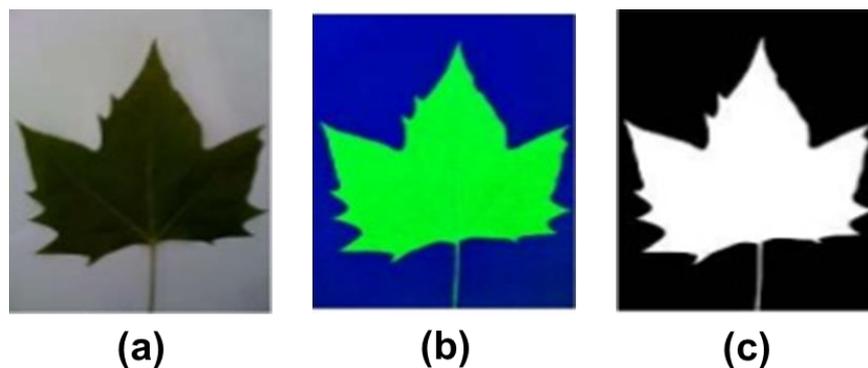
Em 2008, um esforço entre a Universidade de Columbia e do Instituto de Botânica Smithsonian (“Institution Smithsonian”, 2016) surge um dos primeiros serviços de identificação de plantas através da internet baseado em reconhecimento de padrões (KUMAR et al., 2012). Após a publicação do sistema, um desafio também foi enviado à comunidade de visão computacional, a fim de melhorar os métodos apresentados, e que abrangesse uma identificação ampla das espécies de plantas encontradas. O resultado dessa pesquisa é o aplicativo para smartphone LeafSnap, que oferece serviços de identificação para algumas espécies de plantas.

A solução apresentada consiste em o usuário tirar uma foto de uma folha a ser classificada, então a imagem capturada é enviada, por meio da internet, para um servidor que faz os cálculos de reconhecimento, que em seguida transmite os resultados de volta ao dispositivo, que apresenta informações da espécie. Por fim, o sistema retorna ao usuário uma lista de resultados, encomendada pelas probabilidades encontradas entre a imagem da amostra dada e várias classes preditas, para ser escolhida pelo usuário.

Assim, essa métrica não é a probabilidade de obter a classe correta no primeiro resultado, mas a probabilidade de que a classe correta possivelmente está contida nos 5 primeiros resultados, o que significa um *ranking* das 5 possíveis classificações corretas. Isso é importante do ponto de vista da usabilidade, onde uma pessoa pode refinar rapidamente os resultados visuais e escolher a classe correta.

Os autores desse trabalho criaram um método baseado em características de curvatura uni modal e pesquisa por similaridade usando KNN e que foi testado em um conjunto de dados construído por meio de árvores da América do Norte, utilizando 183 espécies no total. Essa metodologia foi baseada em 3 fases. Na primeira fase, eles segmentaram as imagens, que consistia em estimar as distribuições de cores entre o fundo uniforme e o objeto principal que é a folha; logo após, cada pixel é classificado de cada vez em fundo ou em objeto principal, de maneira que, ao final, o objeto principal possa ser identificado. Para fazer essa segmentação eles utilizaram um algoritmo de Expectativa de Maximização (*Expectation Maximization - EM*) (DEMPSTER; LAIRD; RUBIN, 1977), mas antes, eles fizeram uma conversão do domínio de cor de RGB para o espaço de cor Matiz, Saturação e Valor (*Saturation-Value Space - HSV*). Na Figura 3.1 é apresentado o processo desta segmentação.

Figura 3.1: (a) Imagem da Folha em RGB original, (b) imagem convertida para HSV, e (c) segmentação usando EM.



Fonte: (KUMAR et al., 2012)

A segunda fase consistiu na extração de características da folha, mais especificamente, o uso de um extrator de curvatura, onde a forma da folha é representada utilizando medidas de curvatura invariantes à escala. Foram utilizadas duas médias, a área de intersecção de um disco centrado no ponto do contorno, e o comprimento do arco, que é a fração do perímetro do disco interior do contorno. Vinte e cinco discos com raios diferentes foram gerados em cada ponto de apoio de escala. Finalmente um histograma é gerado em cada escala e, em seguida, concatenado para forma um histograma de curvatura sobre a escala.

Por fim, na terceira e última etapa é feita a classificação com o algoritmo KNN. Com essa metodologia os autores atingiram uma percentagem top-5 de classificação de 96,8% em

seus experimentos, sendo que se considerar apenas o primeiro resultado retornado, o top-1, tem-se apenas 69% de acerto.

Um sistema semelhante ao LeafSnap foi desenvolvido por (QUANG-KHUE; THILAN; NGOC-HAI, 2013), porém, este trabalho adicionou algumas particularidades no sistema, como o reconhecimento de folhas para aplicativo Android e a utilização de uma outra base de dados gratuita e bem avaliada pela comunidade de reconhecimento de plantas chamada Flavia (WU et al., 2007).

Os autores primeiramente utilizaram o descritor de característica Características de Imagem da Cena Global (*Global Image Scene Features - GIST*) (OLIVA et al., 2008), do qual foram calculados pixels das imagens das folhas, e um classificador de Máquina de Vetor de Suporte (*Support Vector Machine - SVM*) foi treinado para reconhecer se uma imagem recebida no aplicativo era folha ou não como um primeiro passo, para evitar um processamento desnecessário de uma imagem que não era folha.

Para a tarefa de classificação de espécies de folhas, as características SURF foram extraídas de uma imagem em escala de cinza. A fim de reduzir o tempo de processamento do aplicativo foi utilizado um método chamado *Bag-of-Words* (SIVIC, 2009), que reduz a dimensionalidade do vetor de características, uma vez que o vetor resultante de características SURF pode ser grande, com a extração de milhões de pontos-chaves encontradas em uma imagem. Para fazer a classificação das imagens foi utilizado SVM. A precisão relatada para este método foi de 95,94% no conjunto de dados Flavia, contendo 32 espécies de plantas.

Abrangendo mais de uma base de dados para validar e comprovar a robustez de um sistema de classificação de plantas, (MOUINE; YAHIAOUI; VERROUST-BLONDET, 2013a) propuseram um trabalho usando descritores locais associados com pontos de amostras da margem da folha para criar triângulos com base nestes pontos. Esses pontos de referência são  $N$  pontos distribuídos uniformemente em torno da forma, e cada ponto  $p_i$  é representado por  $N_s$  triângulos computados em diferentes escalas, com uma distância  $d(k)$  entre os pontos do triângulo.

Os autores utilizaram quatro tipos de descritores: Representação da Área do Triângulo (TAR), que é baseado em áreas do triângulo que são robustas ao ruído e fornece informações sobre concavidades locais; Triângulos Orientados a Ângulos (TOA), que descreve os ângulos internos dos triângulos; Tamanho das laterais dos triângulos (TSL), que descreve os comprimentos laterais; e Tamanho e ângulo do lado dos triângulos (TSLA), que consiste em

uma combinação de comprimentos laterais e ângulos, e é considerado o mais robusto dentre estes quatro descritores.

A busca de similaridade entre as características extraídas é feita usando Localizada Sensitiva de *Hashing* (Locality Sensitive Hashing - LSH) (PAULEVÉ; JÉGOU; AMSALEG, 2010), e KNN, projetando os vetores de características para um *hash*<sup>2</sup> para uma busca rápida.

Com a base de dados Flavia, o melhor resultado obtido foi de 69,83% de precisão usando KNN com a representação TSLA. Na base de dados de folhas suecas (SÖDERKVIST, 2001) o método TSLA alcançou 96,53% de precisão. Na base de dados ImageCLEF 2011 (GOËAU et al., 2011), o TSL foi o que teve os melhores resultados, com uma precisão de 80% na categoria Scan (folhas com fundo branco) e 75 % para a categoria Scan-like (folhas com fundo uniforme).

No mesmo ano esses autores realizaram outro trabalho, extraindo características das veias das folhas usando dois descritores de forma, um para descrever o limite das folhas e outro para representar a correlação espacial entre os pontos chaves e a margem da folha, com essa metodologia (MOUINE; YAHIAOUI; VERROUST-BLONDET, 2013b), alcançaram um percentual de 78% na base de dados ImageCLEF na categoria Scan e 70% na Scan-like; para ambos os métodos, a classificação é feita através de uma fusão dos métodos implementados pelos autores utilizando o algoritmo (*Leave Out Algoritm* - LO) (WANG et al., 2006).

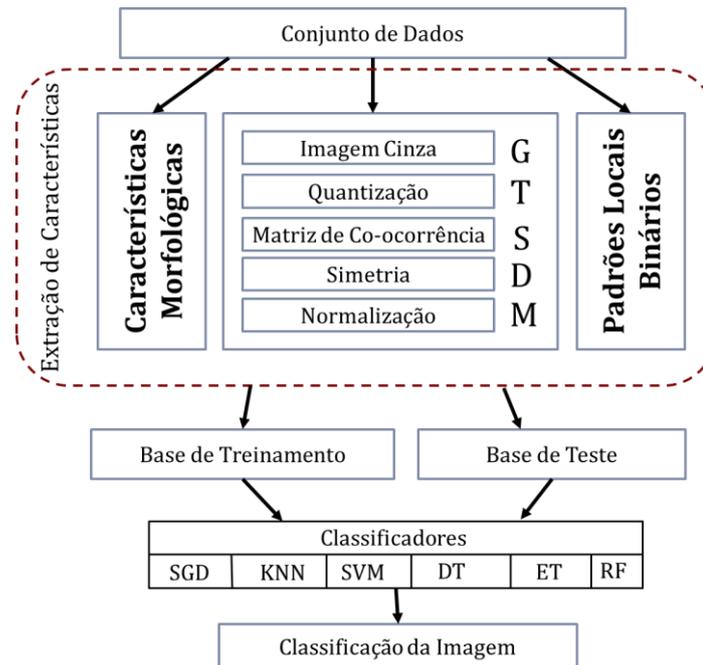
Uma forma de classificar espécies de plantas a partir da textura foi proposta por (ARUN; EMMANUEL; DURAIRAJ, 2013). Eles utilizaram três métodos de descrição das características: Padrões Locais Binários (*Local Binary Pattern* - SIFT), Matriz espacial dependente de tons de cinza (*Gray Tone Spatial Dependente Matriz* - GTSDM) e características morfológicas, como média, variância, assimetria e desvio padrão dos níveis de cinza da imagem.

Para obter a melhor combinação de características para a classificação automática de plantas, os autores utilizaram seis classificadores diferentes: KNN, SVM, RF, Árvores Extras (*Extra Trees* - ET), Árvores de Decisão (*Decision Trees* - DT) e uma Otimização de Gradiente (*Stochastic Gradient Descent* -SGD). Um funcionamento da metodologia proposta pelos autores, desde o recebimento de uma imagem a ser classificada até a classificação da mesma por ser vista na Figura 3.2.

---

<sup>2</sup> Uma função *hash* é um algoritmo que mapeia dados de comprimento variável para dados de comprimento fixo (SHAKHNAROVICH; DARRELL; INDYK, 2006).

Figura 3.2: Metodologia do trabalho proposto por Arun, Emmanuel e Durairaj.



Fonte: (ARUN; EMMANUEL; DURAIRAJ, 2013).

O conjunto de dados criado pelos autores é um conjunto pequeno, que contém 5 espécies diferentes de plantas em um total de 250 imagens, o que não permite confirmar se o método é robusto para propostas generalistas, mesmo quando um 94,7% de precisão é relatado pelos autores com esta metodologia.

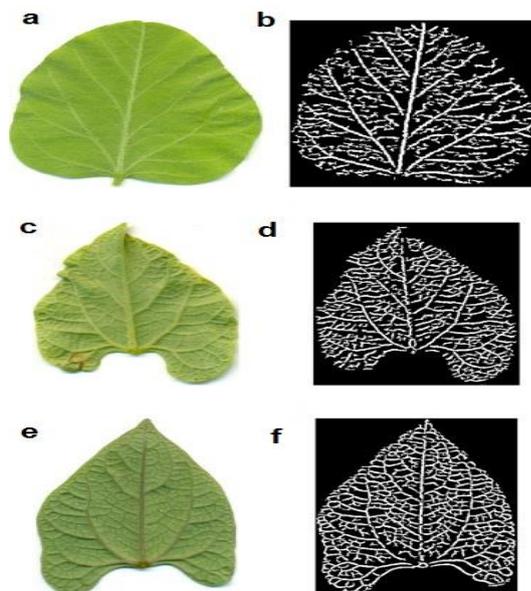
Um sistema para detecção de espécies de plantas medicinais baseado em combinação de um descritor de textura com um classificador Redes Neurais Probabilísticas (*Probabilistic Neural Network* – PNN) foi proposto por (HERDIYENI; KUSMANA, 2013). Utilizando SIFT de diferentes pontos e raios de amostragem, é calculado um histograma para cada conjunto de características contidas no comprimento do raio, e depois, os histogramas são concatenados.

A classificação é feita por meio de 4 camadas de PNN com a função *Radial Base Function* (RBF). A função da rede consiste em uma camada de  $X$  entradas com  $K$  valores a serem classificadores; em seguida, uma camada padrão é responsável por atribuir e aglomerar pesos semelhantes na rede, para fazer uma pré-divisão das classes contidas nas amostras  $X$ ; então, na camada de soma, cada padrão em cada classe é somado para produzir uma função de densidade de população para cada classe; por fim, na camada de saída, vai estar a decisão do classificador, sendo que a classe predita vai ser aquela em que o valor dos pesos aglomerados for maior do que qualquer outra classe. O conjunto de dados utilizado é composto por 30

espécies, com 1.440 imagens de plantas medicinais, e o melhor resultado obtido foi com uma combinação de classificadores utilizando a regra produto com 77%.

Outra possibilidade de extração de características de folhas por meio de suas veias é a de (LARESE et al., 2014). De fato, eles descartaram completamente cor, textura, tamanho e forma, e focaram apenas sobre as veias das folhas. Nesse caso, os autores primeiramente segmentaram a imagem usando uma máscara chamada de Transformação Livre de Acerto ou Erro (*Unconstrained Hit-or-Miss Transform - UHMT*) sobre os pixels de uma imagem em escala de cinza, que consiste em deixar em destaque apenas as veias das folhas (Figura 3.3).

Figura 3.3: (a) folha de feijão de soja. (b) segmentação da imagem para (a). (c) folha de feijão branco. (d) segmentação da imagem para (c). (e) folha de feijão vermelho. (f) segmentação da imagem para (e).



Fonte : (LARESE et al., 2014)

Três classes de legumes foram reconhecidas: soja (*Glycine max* (L) Merr), arroz vermelho e branco (*Phaseolus vulgaris*). Os arrozes pertenciam às mesmas espécies, apresentando formas similares, mas veias diferentes. No entanto, suas cores também foram descartadas, pela conversão da imagem para uma escala de cinzas, e então, a fim de segmentar as veias, a máscara UHMT foi computada com diferentes versões de tamanhos da imagem para obter diferentes níveis de detalhes das veias. Para a extração de características do modelo, Larese utilizou o software LeafGUI (PRICE et al., 2011), que consistiu em mensurar 35

morfologias diferentes das folhas. Usando 10 independentes iterações de 10-folds de validação cruzada. Com isso, os autores reportaram uma precisão do modelo com o classificador SVM de 87 %, sobre 866 imagens de folhas.

Pensando na variabilidade das folhas das plantas intra-espécies e inter-espécies, (MZOUGHFI et al., 2015) focaram nos conceitos e características que descrevem arquitetura das folhas de plantas, em outras palavras, na forma e na colocação de diferentes elementos que constituem a estrutura das folhas, como padrão das veias, arranjo folha, configurações marginais, formas de peças de folha etc. Os autores afirmam que dividir a base de dados nesses conceitos semânticos pode reduzir a complexidade do problema de identificação de folhas, causado principalmente pela variação dentro e entre as espécies, uma vez que permite a definição de características relevantes para cada categoria ou em parte, dependendo das suas propriedades discriminatórias.

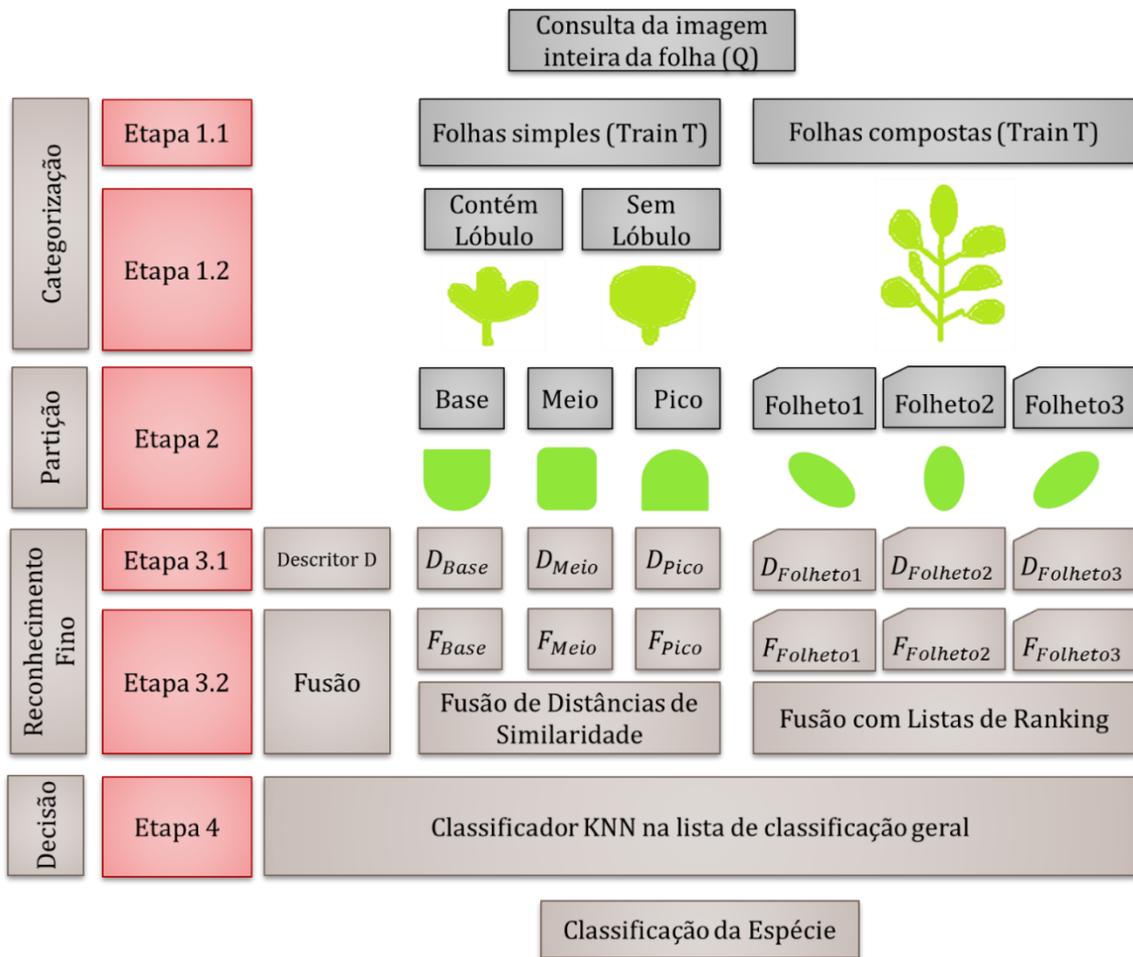
Para isso, eles avaliaram cada etapa da estruturação em um esquema de identificação clássica e de forma independente, onde a forma é representada por uma combinação de descritores única ou global (textura e contorno) e a decisão é feita usando um classificador KNN. Por conta disso, cinco cenários são detalhados na Figura 3.4 e foram construídos da seguinte forma:

- Cenário 1: Representa o esquema de identificação clássica (usando a imagem inteira e usando o conjunto de dados completo) Etapa 3.1 + Etapa 4
- Cenário 2: Divide o conjunto de dados em subconjuntos de folhas simples e compostas Etapa 1.1 + Etapa 3.1 + Etapa 4
- Cenário 3: Divisão da folha em regiões semânticas (parte do meio da folha, pico e base) Etapa 1.1 (primeira parte apenas) + Etapa 2 + Etapa 3 + Etapa 4
- Cenário 4: Divide folhas simples em subconjuntos de folhas com lóbulos<sup>3</sup> e folhas sem lóbulos Etapa 1 + Etapa 2 + Etapa 3 + Etapa 4
- Cenário 5: Faz uma seleção de combinação dos descritores.

---

<sup>3</sup> Partes de uma folha com várias divisões de recortes arredondados em sua volta

Figura 3.4: Arquitetura do trabalho proposto por Mzougui.



Fonte: (MZOUGHJI et al., 2015).

A Figura 3.4 mostra uma visão geral dos passos envolvidos nesta abordagem proposta por (MZOUGHJI et al., 2015). Em primeiro lugar é efetuada uma categorização de duas etapas. Essa categorização avalia a imagem de consulta  $Q$  para que seja separada, por meio de em um subconjunto do conjunto de dados de treinamento (Train  $T$ ), na categoria simples ou categoria compostas; quando uma folha é categorizada na categoria denominada simples, na etapa 1.2 é feito um processamento para verificar se é uma folha simples com ou sem lóbulos. Todas as folhas da categoria composta têm lóbulos. O segundo passo consiste em dividir a imagem da folha inteira em regiões que contenham as partes semânticas; dependendo da sua categoria essas partes podem ser do pico, base e meio para folhas simples, ou então a folha é dividida em três folíolos escolhidos aleatoriamente, para folhas compostas.

O terceiro passo é onde o reconhecimento fino é aplicado, a fim de determinar as espécies da consulta  $Q$ . Esse passo envolve duas etapas: a primeira é o cálculo do vetor de representação local para cada parte (etapa 3.1); essa representação é definida usando um descritor  $D$  padrão que, neste caso, foram utilizados 5 descritores  $D$  baseados em histogramas: Transformada de Fourier e Hough, Histograma de Orientação de borda (*Edge Orientation Histogram* – EOH), Histograma de Orientação de borda local (*Local Edge Orientation Histogram* – LEOH), e Histograma Fragmentado Direcional (*Directional Fragment Histogram* – DFH). Os quatro primeiros descritores descrevem a textura e os padrões de veias enquanto o último representa a margem da folha.

A etapa 3.2 trata da fusão de diferentes representações das partes. Isso é realizado de forma diferente, dependendo da categoria, podendo ser por *ranking* ou por fusão das partes divididas.

No quarto passo é feita a decisão, onde o classificador KNN retorna uma lista das espécies preditas para uma imagem de consulta  $Q$ . Finalmente, na última etapa da Figura 3.4, é feita uma combinação dos melhores descritores de características para cada categoria.

O melhor resultado relatado pelos autores foi encontrado para o cenário 5 (seleção de combinação dos descritores), com uma acurácia de 81,66%, para a categoria Scan da base de dados de 2011 do ImageCLEF, a única categoria avaliada no artigo.

## 3.2 IMAGECLEF

Nesta seção são apresentados os trabalhos expostos no congresso de identificação de plantas em 2011 (GOËAU et al., 2011), que conta com a participação de grupos do mundo inteiro, com o objetivo de automatizar o processo de reconhecimento de folhas de plantas. Mais especificamente, são apresentadas as quatro melhores metodologias dos grupos que participaram desse congresso.

### 3.2.1 GRUPO IFSC/ USP

O procedimento de reconhecimento de plantas descrito pelos autores em (CASANOVA; FLORINDO; GON, 2011), se baseia na ideia de representar o contorno da folha baseado em uma rede complexa para categorizar a folha.

Os autores utilizaram a técnica de segmentação de Otsu (OTSU, 1979) em todas as imagens. A partir das imagens segmentadas a detecção do contorno é extraída e, na sequência,

para a extração de características, é utilizada uma rede complexa. Essa rede complexa consiste em um conjunto de grafos<sup>4</sup> em que cada ponto no contorno da folha representa um nó, e a distância euclidiana entre os pontos representam o respectivo custo da borda.

Dessa forma, a representação resulta em uma matriz de distância entre os nós. A rede é então processada de forma iterativa com histogramas de vários valores de distância, e ao grau máximo e a média de nós do grafo para cada respectivo limiar são concatenados em um vetor de características.

Para fazer a discriminação dos dados foi utilizada a técnica de Análise de dados funcionais (*Functional Data Analysis* - FDA), onde a ideia é considerar o vetor de características como sendo uma representação discreta de uma função contínua. A intenção é transformar o vetor de características em uma função através de meios de interpolação.

O processo de classificação é realizado usando uma Rede Bayesiana sobre variáveis canônicas, que é utilizada para procurar por funções canônicas que maximizam a relação entre a variabilidade das características contidas nas classes similares e heterogêneas. Usando essa metodologia os autores alcançaram uma acurácia nas categorias de Scan e Scan-like de 56% e 40%, respectivamente, valores normalizados na métrica utilizada pelo congresso (GOËAU et al., 2011).

### 3.2.2 GRUPO INRIA

No trabalho realizado pelo grupo INRIA, os autores (BONNET et al., 2011) se baseiam nos seguintes passos: extração de características locais (características de textura SIFT e forma, misturadas com transformadas de Fourier, Hough e EOH calculadas em torno de pontos de Harris); Compactação dos vetores de características com *Hashing* de Margem Máxima Aleatória (*Random Maximum Margin Hashing* – RMMH); Utilização de um algoritmo de Filtragem sobre o modelo; e Decisão baseada no classificador top-K, onde, para cada espécie, um número de ocorrências nas top-K imagens é retornada e usada como uma pontuação, para predizer a classe correta.

Primeiramente as características locais de textura e forma são projetadas para complementar um método ao outro, onde a transformada de Hough oferece informações de forma e a transformada de Fourier oferece a informação de textura. Os pontos de Harris foram

---

<sup>4</sup> O termo redes complexas refere-se a um grafo que apresenta uma estrutura topográfica não trivial, composto por um conjunto de vértices (nós) que são interligados por meio de arestas (BARABASI, 2003)

limitados a 500 por imagem e concatenados em um vetor de 280 para cada um desses pontos principais encontrado em cada imagem.

RMMH é responsável pela concatenação desses vetores de características locais. Esse método foi utilizado para fazer a redução do espaço de características, selecionando cada projeção do vetor tendo como base o conjunto de dados de treinamento. Cada vetor é obtido randomicamente, particionando as características em duas classes, independentemente de seus rótulos e então uma SVM é treinada fazendo uma partição binária. O vetor resultante do qual divide as duas partições com margem máxima é um dos vetores RMMH.

Usando RMMH sobre características locais dá uma medida de semelhança entre os pontos chaves, mas não necessariamente de formas, por consequência, uma saída da filtragem do RMMH com um modelo rígido é aplicada. Um modelo geométrico representado por tamanho, translação e rotação da forma é utilizado para filtragem. Esses parâmetros são computados pelo algoritmo RANSAC, que seleciona aleatoriamente pontos-chaves combinados, verifica se suas distâncias correspondem ao modelo e atualiza o modelo em conformidade.

Finalmente um vetor de consulta é rotulado pela votação das top 10 imagens de treinamento retornadas pelo classificador TOP-K por pontuação de consistência geométrica. Nesse método se obteve uma acurácia de 68,5% na categoria Scan e 46% na categoria Scan-like.

No ano de 2014, com essa mesma metodologia, este grupo criou um sistema robusto de identificação de plantas online e um aplicativo móvel que permite qualquer usuário enriquecer sua base de dados (JOLY et al., 2014). Adicionalmente, eles adaptaram esse algoritmo para fazer uma identificação da combinação dos tipos de órgãos recebida pelo usuário do sistema.

Os órgãos poderiam ser definidos em diferentes visões de plantas de uma mesma espécie, tais como o fruto, caule, flor, folhas, bem como a planta inteira. Com essas informações adicionais o processo de reconhecimento da espécie se tornou mais fácil, pelo fato de se ter diferentes exemplos de órgãos de uma determinada imagem para uma mesma espécie a ser identificada.

### **3.2.3 GRUPO LIRIS**

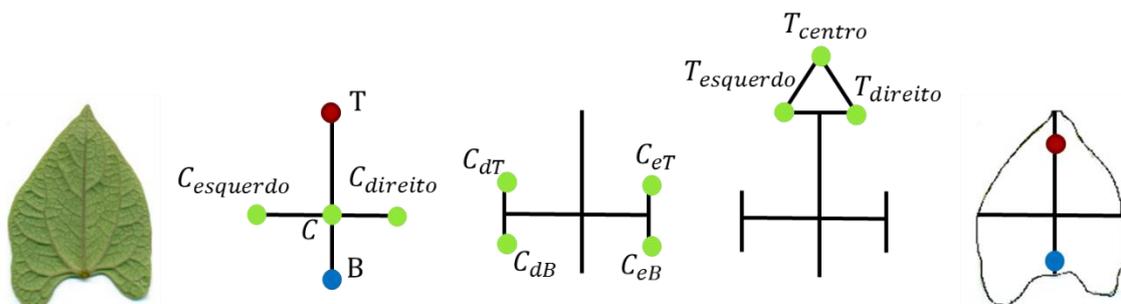
Quando são consideradas apenas folhas no processo de classificação, sabe-se que estas podem mostrar uma impressionante variedade de formas, no entanto, é necessário chegar a uma

representação que abranja qualquer tipo de planta. Pensando nisso, (CERUTTI et al., 2011) abordaram o conhecimento botânico sobre as variações de forma da folha, em um sistema que tenta combinar polígonos de forma padrão em imagens.

Para isso, a extração de características é feita através de um modelo de polígono de dois pontos (base B e ponta T) que definem o principal eixo de simetria da folha. A partir deste eixo é definido o polígono, figura 3.5, que é determinado pelos seguintes parâmetros:

- B, Ângulo de abertura na base da folha;
- T, Ângulo de abertura no topo;
- $w$ , A largura máxima;
- $p$ , Posição relativa, onde a largura é atingida.

Figura 3.5: Arquitetura do trabalho proposto por Cerutti.



Fonte: (CERUTTI et al., 2011).

A ideia de definir um modelo paramétrico simples para representar as variadas formas, segundo os autores, tem a vantagem de transformar uma classificação imprecisa e obscura em valores numéricos dos parâmetros, além de produzir, ao mesmo tempo, algumas descrições que representem a forma geral da folha.

As características finais obtidas são resultantes de um vetor misto e compactado, sendo composto de apenas 20 dimensões no total, com 12 descrevendo a forma geral da planta e 8 para contorno. A classificação é feita usando o método de vizinhos mais próximos (KNN) em uma métrica ponderada de distância, onde cada atributo do vetor de características é escalado separadamente para um desempenho ideal.

A principal vantagem dessa metodologia é ter uma dimensionalidade do vetor de características extremamente baixa comparada com os outros trabalhos. Os autores obtiveram um resultado de 53 pontos percentuais para Scan e 54 para Scan-like.

### 3.2.4 GRUPO SABANCI-OKAN

Sabendo que a planta consiste de diversas morfologias (YANIKOGLU; APTOULA; TIRKAZ, 2011) propuseram um sistema que aborda sete descritores em função da mesma. Uma lista dos descritores usados, bem como seu tipo, está na Tabela 3.1.

Tabela 3.1: Descritores utilizados no trabalho de (YANIKOGLU; APTOULA; TIRKAZ, 2011).

Identificador	Descritor	Tipo do Descritor
1	Histograma de covariância circular (CCH)	Textura
2	Pontos invariantes de rotação (RIT)	Textura
3	Fator de largura, comprimento e volume	Forma
4	Medida de convexidade	Forma
5	Estatística de forma básica	Forma
6	Covariância da borda	Forma
7	Momentos estáticos de cor RGB	Cor

Esses descritores foram então fundidos em dois vetores de características distintos para treinar classificadores SVM: o Vetor de características de forma, composto por todos os descritores de forma; e Vetor de características geral, composto por todos os sete descritores.

As saídas desses classificadores são distâncias da instância de teste para cada classe de plantas. Para isso, os autores treinaram um terceiro classificador para aprender como combinar aqueles dois classificadores. Assim, o vetor de características usadas na formação do combinador é de comprimento  $2 \times K$ , em que  $K$  é o número de classes no problema.

Na fase final de classificação, as cinco mais prováveis classes são selecionadas e uma SVM multi-classe é treinada especificamente para estas classes de desambiguação. Isso foi feito para treinar classificadores a aprender distinguir classes semelhantes (por exemplo, diferentes tipos de bordas que são muito semelhantes entre si, em comparação com as outras plantas). Com isso eles obtiveram a precisão com validação cruzada para cada classificador, Tabela 3.2.

Tabela 3.2: Resultado gerados através da validação cruzada no trabalho de (YANIKOGLU; APTOULA; TIRKAZ, 2011).

Descritores	Acurácia SVM
3,4,5,6	71.46
1,2,3,4,5,6,7	89.69
Combinação (1,2,3,4,5,6,7)	93.64

Normalizando o resultado para a tarefa de identificação de plantas do congresso, com o melhor método de combinação, eles obtiveram uma acurácia de 68,2% para Scan e 47% para Scan-like.

### 3.3 SUMÁRIO DO ESTADO DA ARTE

Nesta Seção é apresentada uma breve descrição sobre as metodologias aplicadas para a classificação de espécies de plantas utilizadas pelos autores citados nas seções 3.1 e 3.2, além de apresentar os resultados obtidos pelos mesmos ao testarem suas metodologias.

Na Tabela 3.3 é apresentado um resumo das metodologias empregadas, o tipo de extrator de características utilizado, bem como o método de classificação, a base de dados utilizada para os testes, e os resultados obtidos pelos autores. Os resultados podem ser exibidos como taxas de acerto ou taxas normalizadas da métrica ImageCLEF, dependendo da forma com que os autores disponibilizam seus resultados. Informações sobre essas métricas podem ser encontradas na Seção 2.5.

Tabela 3.3: Sumarização das Metodologias

Autores	Metodologia Empregada	Extrator de Características	Método de Classificação	Base de Testes	TA (%)	Métrica ImageCLEF (%)	
						Scan	Scan-like
(KUMAR et al., 2012)	Focou em analisar a forma das folhas usando o algoritmo Expectativa de Maximização (EM) para segmentar as imagens	Histogramas de curvatura sobre múltiplas escalas	KNN	Base criada, com 183 espécies de plantas	Top-5 (96,8%) Top-1 (69%)	-	-
(QUANG-KHUE; THI-LAN; NGOC-HAI, 2013)	Observou a textura das plantas	Algoritmo SURF com o método de redimensionamento do vetor de características Bag-of-Words	SVM	Base pública denominada Flavia com 32 espécies	95,94%	-	-
(MOUINE; YAHIAOUI; VERROUST-BLONDET, 2013a)	Utilizou descritores locais de forma e contorno na margem da folha para criar triângulos com base em pontos locais extraídos	Representações Triangulares: TAR, TOA, TSL, TSLA e LSH	KNN	ImageCLEF 2011	-	Scan	Scan-like
						80%	75%
(MOUINE; YAHIAOUI; VERROUST-BLONDET, 2013b)	Para identificação das plantas, verificou-se suas veias baseado em suas formas	Um modelo proposto pelos autores que extrai características das veias das plantas	Fusão das respostas	ImageCLEF 2011	-	Scan	Scan-like
						78%	70%
(ARUN; EMMANUEL; DURAIRAJ, 2013)	Combinação entre características e classificadores.	SIFT, GTSDM e características morfológicas dos níveis de cinza da imagem	KNN, SVM, RF, Extra Trees (ET), Decision Trees (DT) e Stochastic Gradient Descent (SGD)	Base criada, contendo 5 espécies, com 250 imagens	94,7%	-	-

(Continua)

Tabela 3.3: Sumarização das Metodologias

(Continuação)

<b>Autores</b>	<b>Metodologia Empregada</b>	<b>Extrator de Características</b>	<b>Método de Classificação</b>	<b>Base de Testes</b>	<b>TA (%)</b>	<b>Métrica ImageCLEF (%)</b>	
(HERDIYENI; KUSMANA, 2013)	Usou plantas medicinais para avaliar seu método que descreve a textura da mesma	SIFT	Probabilidades Neurais (PNN) com função RBF	Base criada, contendo 30 espécies, com 1.440 imagens	77%	-	
(LARESE et al., 2014)	Para de uma classificação automática de plantas, segmentou as imagens com o algoritmo UHMT, focando em suas veias	Software LeafGUI que consiste em mensurar 35 morfologias diferentes das folhas	SVM, com validação cruzada de 10-folds	Base criada com 3 espécies, contendo 866 imagens de folhas	87%	-	
(MZOUGHFI et al., 2015)	Dividiu folhas de plantas compostas e simples em partes, e verificou as variações entre elas	Combinação de descritores de diferentes partes das plantas, transformada de Fourier, EOH, LEOH, Hough e DFH	KNN	ImageCLEF 2011, apenas categoria Scan	-	Scan	
						81,66%	
(CASANOVA; FLORINDO; GON, 2011)	Analisaram o contorno da folha, bem como a utilização de um algoritmo de segmentação denominado OTSU	Rede Complexa + FDA	Rede Bayesiana	ImageCLEF 2011	-	Scan	Scan-like
						56,20%	40,20%

(Continua)

Tabela 3.3: Sumarização das Metodologias

(Continuação)

Autores	Metodologia Empregada	Extrator de Características	Método de Classificação	Base de Testes	TA (%)	Métrica ImageCLEF (%)	
						Scan	Scan-like
(BONNET et al., 2011)	Uma fusão das características locais de textura e forma é abordada	Hough, EOH e Fourier extraídos de cada ponto de Harris, redutor de vetores de dimensão (RMMH), estimador de modelos matemáticos (RANSAC)	TOP-K	ImageCLEF 2011	-	Scan	Scan-like
						68,50%	46,40%
(CERUTTI et al., 2011)	Esboçaram um polígono nas folhas, para extrair características do lados, base e topo a partir de sua forma.	Modelo de polígono de dois pontos (base e topo) que definem o eixo principal da folha.	KNN	ImageCLEF 2011	-	Scan	Scan-like
						53,90%	54,30%
(YANIKOGLU; APTOULA; TIRKAZ, 2011)	Abordaram a morfologia das plantas	Combinação dos descritores de textura: CHH, RIT; descritores de forma: Morfologias da planta e descritores da cor: Momentos estáticos de cor RGB	SVM	ImageCLEF 2011	-	Scan	Scan-like
						68,2%	47,60%

### 3.4 DISCUSSÃO CRÍTICA SOBRE OS MÉTODOS APRESENTADOS

Os principais desafios para a classificação em problemas de reconhecimento de espécies de plantas, por meio de identificação em imagens, são: a grande quantidade de espécies de plantas (classes), sua grande variabilidade inter e intra-espécies, a similaridade entre as classes, e as condições inerentes a maneira da aquisição da imagem. Por isso, para fazer um levantamento de quais características são relevantes, é necessário um conhecimento de domínio específico.

De maneira geral, os resultados recentes da literatura estão longe de atingir as exigências do mundo real em termos de ferramentas automáticas ou semi-automatizadas aplicada à identificação de espécies de plantas, devido a maioria dos trabalhos ou ferramentas disponíveis conseguirem identificar apenas algumas dezenas ou centenas de espécies, ou seja, é muito pouco se comparado com a totalidade das espécies catalogadas ao redor do mundo. Também se observa que os recursos extraídos da imagem ainda não são bem explorados para a classificação de plantas, dada a complexidade do problema e a grande similaridade entre as espécies.

Já os sistemas que não utilizaram de um recurso de pré-processamento em suas imagens, como o processo de segmentação, quando levados a campo, será difícil obter uma classificação exata, pelo fato de não prever ruídos como sombras, reflexos de luzes e até mesmo indivíduos que não fazem parte do objeto principal a ser identificado.

A grande deficiência dos sistemas apresentados está em não conseguir gerar um modelo para aquelas plantas que tem poucas amostras, sendo assim, estas amostras ficam vulneráveis a se confundir no meio de tantas amostras de diversas espécies de plantas em um sistema de classificação. Alguns trabalhos se limitam a usar poucas características para descrever uma folha, bem como em seu processo de classificação, onde a decisão é tomada por meio de um único classificador. Um exemplo de trabalho baseado nesse tipo de metodologia é o de (HERDIYENI; KUSMANA, 2013), que mesmo criando sua própria base de dados obteve uma acurácia baixa, pelo fato de usar apenas um descritor de características modelado em um único classificador. Esse método se assemelha a um processo de classificação que conta com apenas a opinião de um especialista para uma determinada visão da planta. No entanto, se esse mesmo trabalho obtivesse a opinião de vários especialistas, bem como diversas visões que correspondessem partes da planta, essa classificação seria incrementada, pelo fato de se ter a opinião de diferentes especialistas e, para se dar o rótulo de classificação da planta teria um senso comum entre eles.

Com relação às bases de dados empregadas na literatura, foi verificado que somente os trabalhos que utilizam a base de dados ImageCLEF e o trabalho de Kumar, são postos à prova na abordagem de classificação de plantas do mundo real. As imagens contidas nestas bases de dados são provenientes de diversos colaboradores ao redor do mundo e também de plantas diferentes nas mais variadas posições, condições de iluminação e estágio, a fim de imitar o mundo real, onde o usuário tenta reconhecer uma planta em qualquer área; caso contrário, se as imagens fossem sempre da mesma planta e na mesma localização, talvez o reconhecimento não seria generalizado, e então o sistema reconheceria uma planta somente naquela determinada região ou situação. Por causa disso, os trabalhos que tem suas próprias bases criadas tem seu devido valor, porém, não é possível avaliar o quão generalizado é sua metodologia proposta.

Portanto, é importante que um sistema de reconhecimento de plantas seja posto à prova em uma base de dados que contenha um elevado número de espécies disponíveis, coletadas em diferentes localizações, condições climáticas, e ao longo de várias estações do ano em um período razoável de tempo. Além disso, havendo uma base de dados ampla e padrão, podem-se padronizar os testes neste tipo de sistema, possibilitando uma melhor comparação entre as metodologias propostas e uma possível generalização do método quando testadas em outros ambientes.

## 4. MÉTODO PROPOSTO

A seguir é descrito o método proposto neste trabalho para a classificação de espécies de plantas baseada em técnicas de reconhecimento de padrões com aprendizagem supervisionada e combinação de classificadores. O reconhecimento é feito através de imagens capturadas das folhas das plantas, das quais são extraídas características baseadas em textura para o treinamento dos classificadores.

Mais especificamente, busca-se fornecer uma metodologia preparada para classificar espécies de plantas com diferentes problemas de ruídos utilizando imagens predominantemente focadas nas folhas, coletadas por diferentes usuários e em distintas áreas.

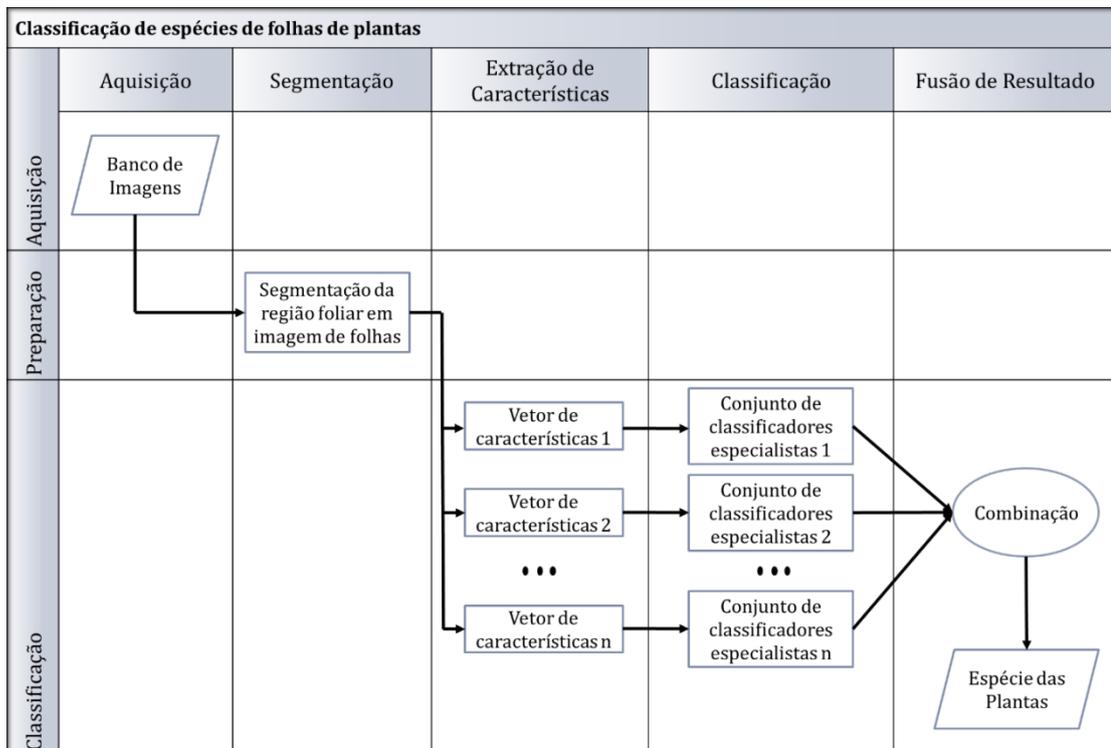
### 4.1 MODELAGEM DO PROBLEMA

Em qualquer problema básico de reconhecimento de padrões é desejado atribuir classes pré-definidas a objetos. No reconhecimento de padrões com treinamento supervisionado, em primeiro lugar são extraídas características de objetos com classes conhecidas, os quais são utilizados para treinamento de um modelo. Em seguida, depois de treinado, esse modelo pode ser utilizado para classificar novos objetos dos quais não se sabe até então a qual classe pertencem.

A classificação de espécies de plantas utilizando suas imagens é modelada, neste trabalho, como um problema de reconhecimento de padrão com treinamento supervisionado. Para esse fim, são extraídos vetores de características baseadas em textura das folhas das plantas individuais das quais os estados são conhecidos com antecedência. Essas características são então utilizadas para o treinamento do modelo, logo, novas folhas de plantas podem ser classificadas, satisfazendo para isto a extração de suas características e que estas sejam submetidas a teste no modelo treinado.

A Figura 4.1 exhibe a sequência seguida para a classificação de uma espécie de folha de planta.

Figura 4.1: Fluxograma para classificação de espécies de plantas



Fonte: O autor.

Primeiramente, na fase de Aquisição é selecionado um banco de imagens consistente ao problema para que, na fase de preparação, as imagens das folhas individuais sejam segmentadas e convertidas para escala de cinza, com o intuito de fazer um pré-processamento. Na fase de classificação é realizada a extração das características das amostras, onde cada vetor gerado é emitido para categorização em um classificador especialista para que, enfim, os resultados sejam combinados e a folha seja classificada.

O banco de imagens de folhas de plantas utilizadas na fase de aquisição é detalhado na seção 4.2, a fase de preparação das imagens é descrita na seção 4.3, enquanto a seção 4.4 detalha a fase de classificação das amostras.

## 4.2 BASE DE DADOS

Uma base de imagens de espécies de folhas de plantas foi selecionada a fim de que esta pudesse prover os dados necessários para os treinamentos e testes neste trabalho. A propósito,

como visto na Seção 3.3, muitos autores têm relatado resultados utilizando o conjunto de dados ImageCLEF 2011.

ImageCLEF é um conjunto de imagens proposto por (GOËAU et al., 2011) com 41 espécies de plantas, e detém estados variáveis dos conceitos morfológicos da planta, tal como, localização, formas e período da planta, bem como distorções derivadas das condições de aquisição de imagem (principalmente rotação, sobreposição e sombras). A base de dados permite o estudo caracterizado de dados, o que não é possível com outros conjuntos mais populares que se concentram no nível básico de informações.

As informações detalhadas sobre a aquisição do banco de imagens são descritas no tópico 4.2.1, as características atuais da base são detalhadas na Seção 4.2.2, enquanto no tópico 4.2.3 são descritos os descritores das imagens.

#### **4.2.1 BANCO DE IMAGENS**

A principal originalidade desta base de dados é que ela foi construída através de uma iniciativa de cientistas conduzida pela Tela botânica (“Tela Botanica”, 2016), que é considerada a principal rede francesa de botânicos amadores e especialistas, e que tem mais de 20 mil membros registrados em volta do mundo em mais de 70 países. Essa rede organizada conta com diversos projetos, e um desses projetos foi lançado em 2010 especificamente para coletar imagens, para fins de identificação automática de plantas.

As imagens recebidas são definidas por um protocolo de aquisição, que é espalhado para os membros do projeto, para que cada imagem seja capturada e, logo após, associada a várias determinações feitas por membros distintos; dessa forma, uma imagem é rotulada, pelo consenso calculado através do voto de cada membro. Se existir pessoas que fazem determinações contraditórias, é realizada uma discussão através de um fórum associado a cada imagem ambígua e as determinações podem ser revistas a qualquer momento. Finalmente, cada imagem disponível na base de dados é associada com um descritor de imagem (tópico 4.2.3), que disponibiliza informações sobre o autor, data, localidade, identificador e o tipo da planta, bem como sua espécie, família e gênero.

A principal razão de se usar o conjunto de dados ImageCLEF neste trabalho é que a criação da base por meio do compartilhamento social torna a avaliação do método mais próxima das condições de um cenário de identificação do mundo real. Partindo do princípio que as imagens da mesma espécie são provenientes de plantas que vivem em áreas distintas, as

imagens são adquiridas por diferentes usuários, que podem não ter utilizado o mesmo protocolo de aquisição; adicionalmente, as imagens são adquiridas em diferentes períodos do ano.

#### 4.2.2 CARACTERÍSTICAS DA BASE DE DADOS

A base de dados selecionada tem três categorias de plantas distintas, utilizando-se o protocolo de aquisição descrito na Seção 4.2.1. No entanto, neste trabalho foram utilizadas duas categorias de plantas chamadas Scan e Scan-like, sendo que na primeira categoria constam imagens de folhas com um fundo branco, sem muitos problemas de ruídos nas amostras, e na segunda categoria estão imagens com empecilhos como objetos indesejados, sombras, variações de posições e iluminação.

Depois de coletadas, as imagens foram separadas em uma estrutura de diretórios de acordo com suas respectivas categorias e suas espécies seguindo a hierarquia: "Categoria/ espécie/ nome\_da\_espécie\_contador.jpg", onde Categoria refere-se à condição da imagem capturada, podendo ser Scan ou Scan-like. Já o caminho espécie é em função das espécies existentes na base de dados, e, por fim, é dado o nome da imagem final seguido de um contador sequencial, que consiste em contar quantas imagens existe para aquela determinada espécie, diferenciando-as para não haver confusão.

Sendo assim, o caminho completo de uma imagem na base de dados poderia, por exemplo, ser da seguinte forma "Scan-like/ Ginkgo-biloba/ Ginkgo-biloba\_1.jpg".

Essa estrutura foi feita com o intuito de estabelecer uma forma simples de separar as imagens, onde fosse possível encontrar todas as informações sobre determinada imagem no seu próprio caminho na estrutura de diretórios, além de facilitar a busca por imagens de determinada espécie.

Todas as imagens presentes na base de dados foram rotuladas de forma automática, com suas respectivas espécies, devido à leitura dos descritores *Extensible Markup Language* (XML) (DEITEL, 2011) de suas respectivas imagens (seção 4.2.3). Um detalhamento do nome científico de todas as espécies usadas neste trabalho pode ser visto na Tabela 4.4.

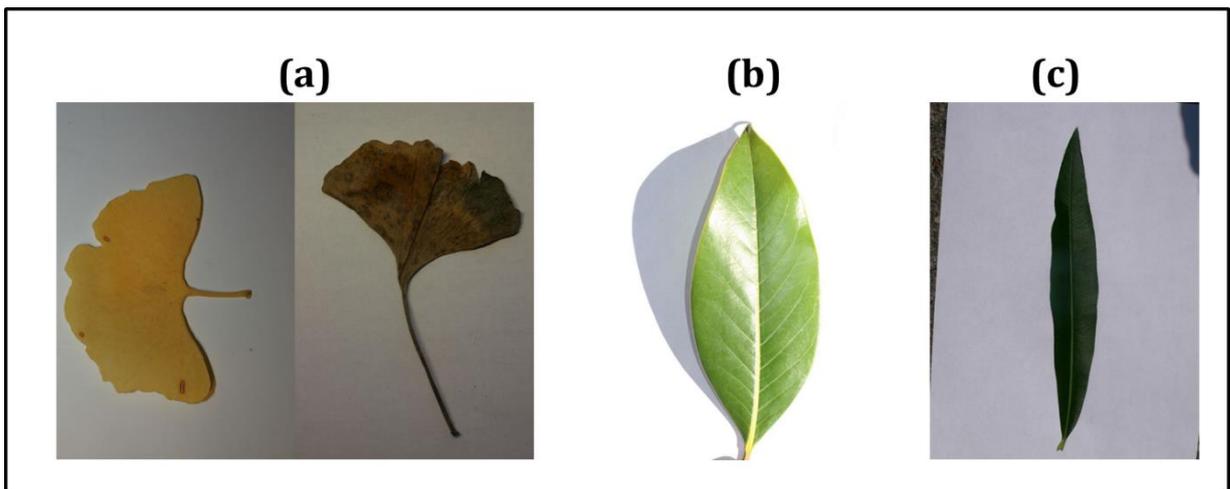
Atualmente a base conta com um total de 2.586 imagens para as duas categorias. A Tabela 4.1 contém uma síntese das informações sobre o estado atual da base de imagens e suas categorias.

Tabela 4.1: Detalhamento da base de dados

Categoria	Número de Espécies	Imagens
Scan-like	13	683
Scan	28	1.903
<b>Total</b>	<b>41</b>	<b>2.586</b>

Na Figura 4.2 são expostas algumas amostras complicadas existentes na categoria Scan-like. Na primeira categoria, Scan-like, Figura 4.2 (a), um dos desafios para a classificação das amostras de plantas nesta categoria, é a variação de rotação das plantas, que pode variar de acordo com o indivíduo que fez a aquisição da imagem.

Figura 4.2: Ruídos na categoria Scan-like. (a) variação de rotação das plantas de mesma espécie *Ginkgo biloba*, (b) presença de sombras e luminosidade da espécie *Quercus ilex*, (c) ruídos na espécie *Nerium oleander*.

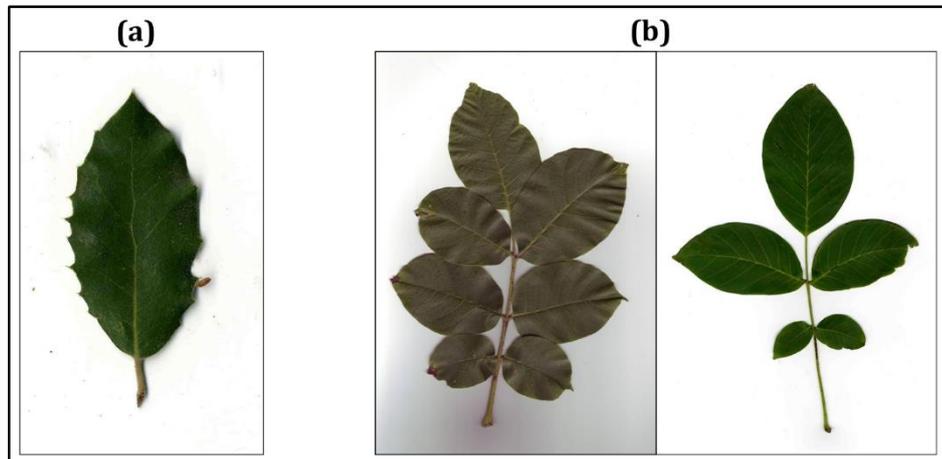


Fonte: O autor.

Outros desafios da categoria Scan-like podem ser observados nas imagens 4.2(b) e 4.2(c) respectivamente, que é a presença de sombras projetadas pela própria folha, e a grande variação de luminosidade, que ocorre em má condições de ambiente para obter a foto, bem como, ruídos no fundo da imagem e objetos que podem confundir a classificação na cena.

Assim como nas imagens da categoria Scan-like, as imagens da categoria Scan também apresentam manchas ao redor da folha da planta e até mesmo alguns restos de terra, sobrepostos a imagem (Figura 4.3 (a)).

Figura 4.3: Ruídos na categoria Scan. (a) *Quercus ilex*: manchas ao redor da folha e pedaço de terra, (b) *Juglans regia*: variações da morfologia na mesma espécie de planta.



Fonte: O autor.

Além disso, devido ao fato de diferentes usuários do mundo inteiro capturar imagens de uma mesma espécie de plantas, há algumas espécies que variam seu crescimento, pela questão de que em uma região a planta se desenvolve melhor que na outra, dificultando a classificação das amostras principalmente quando a folha é composta por vários lóbulos, como pode ser ver na Figura 4.3(b).

#### 4.2.3 DESCRITORES DAS IMAGENS

Para cada uma das imagens presentes no conjunto de dados ImageCLEF 2011 existe um descritor no formato *Extensible Markup Language* (XML) (DEITEL, 2011). Esses descritores contêm informações sobre o estado de cada uma das imagens, como exibido na Figura 4.4, que exemplifica o descritor XML de uma das imagens de folha de plantas.

Os arquivos XML são estruturados da seguinte forma: o elemento raiz, denominado Imagem, define o identificador da imagem, categoria, autor e de qual planta individual foi capturada a imagem.

Dentro do nodo raiz encontram-se a taxonomia, que definem todos os graus de classificação da planta, contendo o reino, ordem, família, espécie e gênero da planta. Neste trabalho é focado apenas na identificação das espécies das plantas, que é considerada uma das

partes mais difíceis, considerando que em uma hierarquia das plantas, as diferenças morfológicas entre as espécies de uma família podem ser pequenas.

Figura 4.4: Exemplo de Descritor XML de uma planta.

```
<?xml version="1.0"?>
<Imagem>
  <Id>4.jpg</Id>
  <Categoria>Scan</Categoria>
  <Autor>Fulano</Autor>
  <PlantaIndividual>374</PlantaIndividual>
  <Taxonomia>
    <Reino>Plantae</Reino>
    <Ordem>Laurales</Ordem>
    <Familia>Lauraceae</Familia>
    <Especie>nobilis</Especie>
    <Genero>Laurus</Genero>
  </Taxonomia>
</Imagem>
```

Fonte: O autor.

Estes arquivos foram importantes para fazer a separação das imagens de treinamento e teste e dar seus respectivos rótulos. Também pontos importantes foram extraídos como: quantos autores capturaram as imagens, quantas amostras cada autor capturou e quantas plantas individuais existem em uma determinada espécie, para que este seja utilizado na métrica ImageCLEF disponibilizada pelos autores (Seção 2.5.2).

### 4.3 PRÉ-PROCESSAMENTO E SEGMENTAÇÃO DA IMAGEM

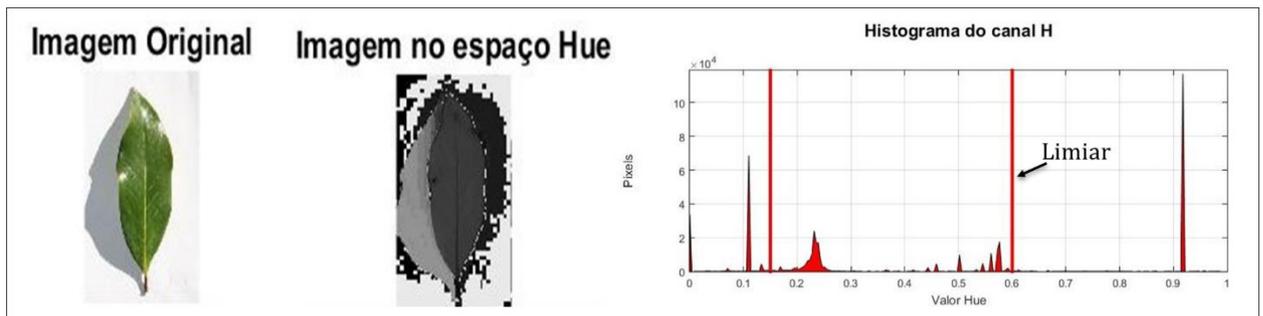
Primeiramente é realizada a eliminação da sombra da folha. Para isso, a imagem é convertida do espaço de cores RGB, para o sistema de cor Matiz, Saturação e Valor (*Saturation-Value Space* - HSV) (FILHO; NETO, 1999).

Segundo (SHUGEN; ZEJIN; DEREN, 2003), no espaço HSV os pixels da região de sombra apresentam características de alto tom (H), alta saturação (S) e baixa intensidade (V). Neste trabalho as características de alto tom (H) foram suficientes para detectar a sombra presente nas imagens.

A Figura 4.5 apresenta um histograma médio do canal H das imagens que contêm sombras, em que é possível perceber que o histograma tem uma divisão, destacando-se a classe

de pixels com valores H entre 0,2 e 0,6, que representam a folha. Já os pixels com valores de H maiores que 0,6, são pouco frequentes, podendo induzir que tais valores pertencem a pequena região de sombra destacada na imagem no espaço H, pois não armazenam informações de cores. Portanto, esses valores de H são desconsiderados, descartando assim a sombra da folha.

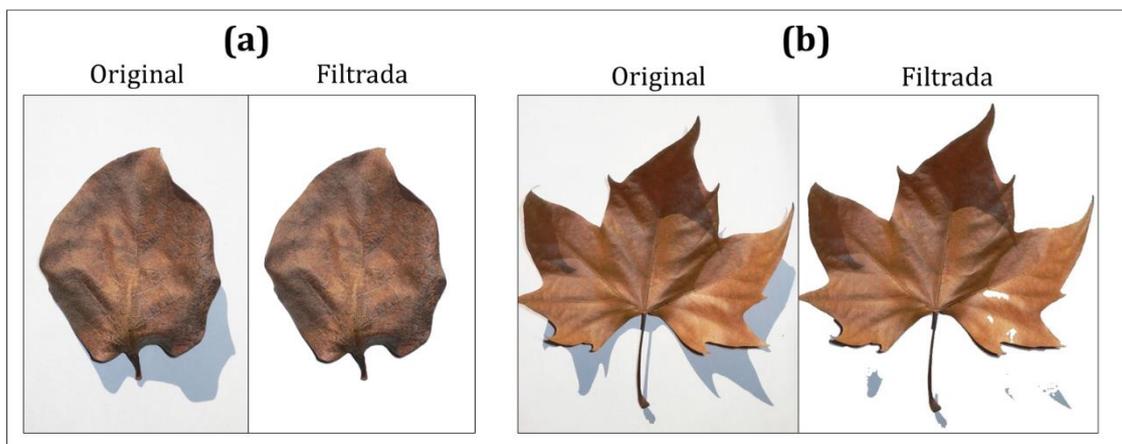
Figura 4.5: Delimitação do histograma da imagem para a separação da sombra contida na imagem.



Fonte: O autor.

A Figura 4.6 (a) apresenta um exemplo de exclusão da sombra, onde os pixels classificados como sombra são subtraídos e a imagem é convertida para tons de cinza, com a função de diminuir o volume de informação a ser processada. Entretanto, para algumas amostras, o pré-processamento não atinge o esperado, como pode ser observado na Figura 4.6 (b), onde a exclusão da sombra está incompleta.

Figura 4.6: (a) Diospyros kaki: separação da sombra completa, (b) Platanus x hispanica: separação incompleta da sombra.



Fonte: O autor.

Segmentar uma imagem neste trabalho tem o conceito de separar o objeto (folha) do fundo da imagem. Para isso, a segmentação foi realizada pelo algoritmo de limiarização automática de OTSU (OTSU, 1979). Esse método considera que o histograma de uma imagem é formado por duas classes, por exemplo, os pixels do objeto (folha) e os pixels do fundo. O que diferencia essas duas classes é um valor global chamado de limiar, que é calculado automaticamente pelo algoritmo de Otsu, potencializando a variância entre as classes objeto e fundo.

Dessa maneira, todos os pixels com valores inferiores deste limiar são classificados como objeto e os demais são classificados como fundo. A Figura 4.7 apresenta apenas o resultado do processo de segmentação pelo uso do algoritmo de Otsu.

Figura 4.7: Acer campestre: remoção apenas do fundo pelo algoritmo de Otsu.



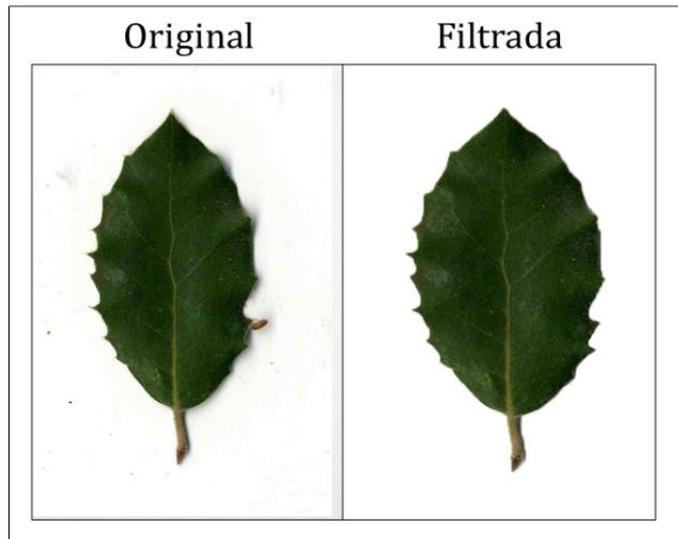
Fonte: O autor.

Após a segmentação, a imagem pode conter objetos que não sejam elementos de uma amostra de folha, como ciscos e pequenos pedaços que se desintegram da amostra principal. Para resolver esse problema foi utilizado um algoritmo de rotulação (FILHO; NETO, 1999), que tem como base identificar todos os objetos ligados a imagem. Esses objetos são considerados regiões que apresentem pixels com valor 0 (preto). Após a identificação são removidos todos os objetos, deixando apenas o de maior tamanho que é a folha a ser analisada.

Um exemplo do aparecimento de objetos indesejáveis é apresentado pela Figura 4.8. Objetos indesejáveis na folha podem influenciar na caracterização precisa da mesma, o que

pode levar um algoritmo de extração de característica usado, definir um objeto indesejado como um ponto característico da imagem.

Figura 4.8: Quercus ilex: remoção de objetos indesejados.



Fonte: O autor.

#### 4.4 CRIAÇÃO DOS CLASSIFICADORES

Nas seguintes subseções são estabelecidos os vetores de características extraídos das amostras, o processo de treinamento dos indutores e o esquema de fusão dos resultados, que mostra o funcionamento da fase de classificação da Figura 4.1.

##### 4.4.1 EXTRAÇÃO DE CARACTERÍSTICAS

Após o pré-processamento das imagens, são extraídas características baseadas em textura das mesmas, para isso são utilizados extratores do tipo SIFT (seção 2.2.1), SURF (seção 2.2.2) e HOG (seção 2.2.3).

As características baseadas em SURF e SIFT são extraídas com invariância a rotação e escala. Já as características baseadas em HOG são extraídas utilizando grades de células uniformemente espaçadas sobrepostas à imagem, que calcula os histogramas de gradientes. Os valores dos vetores dos três descritores de características são normalizados entre 0 e 1.

Esses descritores foram avaliados quanto a capacidade de representar características texturais nas imagens das plantas, que implicou na variação dos parâmetros utilizados, de acordo com a possibilidades de configuração de cada descritor; e na identificação dos valores que maximizaram as taxas individuais de reconhecimento.

Como as imagens possuem número variável de pontos característicos, que podem gerar vetores de alta dimensionalidade pelos os algoritmos SIFT e SURF, uma estratégia usada foi reduzir o conteúdo para um tamanho fixo, sendo criado um vocabulário de descrição da imagem. Assim, apenas um conjunto definido de pontos característicos extraídos é utilizado pelo vocabulário.

Para agrupar algumas centenas de pontos característicos visuais de uma imagem em um vocabulário de tamanho fixo, o método *Bag-of-Words* é utilizado. O Histograma de palavras visuais (*Bag-of-Words*) proposto por (SIVIC; ZISSERMAN, 2003) realiza a quantização vetorial por meio do algoritmo K-médias, que permite uma aceleração ao processo de classificação de uma imagem com uma grande quantidade de pontos característicos, onde cada ponto chave possui uma referência da imagem de origem. Segundo (SIVIC; ZISSERMAN, 2003), o tamanho do vocabulário utilizado deve ser grande o suficiente para distinguir mudanças relevante em partes da imagem, mas não tão grande para distinguir variações irrelevantes.

Portanto, por meio de um primeiro experimento (Apêndice A) é que foi definido que o tamanho ideal do vetor de características (vocabulário) dos algoritmos SIFT e SURF é de 512 elementos. Com relação à quantidade de pontos chaves extraídas de cada imagem, foi utilizado a definição padrão para cada um dos descritores, sendo 800 para SIFT e 700 para SURF. Vale ressaltar que, conforme descrito nas Seções 2.2.1 e 2.2.2, os vetores de características que foram gerados para cada ponto de interesse identificado são compostos de 128 elementos para o algoritmo SIFT e 64 elementos para SURF respectivamente.

Já a definição do tamanho do vetor para HOG foi feita através dos passos comentados na Seção 2.2.3, que consiste em um descritor final que possui uma dimensão igual a 3.780 elementos. A Tabela 4.2 sumariza os descritores e o tamanho dos vetores de características gerado.

Tabela 4.2: Extração de características

Descritor	Características
SIFT	512
SURF	512
HOG	3.780

#### 4.4.2 CLASSIFICAÇÃO E FUSÃO DOS RESULTADOS

A metodologia de classificação apresentada neste trabalho consiste em, empregando um mesmo indutor, treinar um classificador para cada um dos tipos de descritores de características da Seção 4.4.1, totalizando 12 classificadores diversos. O resultado da classificação de uma amostra estabelecida é concebido por meio da combinação das respostas de todos os classificadores.

As estratégias de combinação de classificadores apresentadas neste trabalho são a combinação pela multiplicação (produto) das respostas dos classificadores, já que esses métodos apresentaram os melhores resultados dentre todas as estratégias de combinação.

Os indutores que foram selecionados para a criação dos classificadores propostos são: SMO, RNA, KNN e RF. Na Tabela 4.3 são expostos os parâmetros de cada classificador, que foram selecionados por meio de buscas em *grid* em validação cruzada de 5-folds (CHEN, 2012), totalizando 12 classificadores distintos.

Há alguns parâmetros que não são possíveis definir utilizando a busca em *grid*, por conta de seu alto processamento. Então, para estabelecer o parâmetro relacionado aos neurônios da camada oculta utilizada pelo classificador RNA é utilizada a definição padrão do algoritmo de múltiplas camadas (*MultiLayerPerceptron*) disponibilizada pelo programa WEKA (HALL et al., 2009). Contudo a definição dos neurônios da camada oculta da rede neural artificial é dada pela soma entre atributos expostos na Tabela 4.2 de cada descritor e classes de cada categoria, e então é feita a divisão por dois. Portanto, na camada oculta da categoria Scan obteve-se 1.904 neurônios para o algoritmo HOG e 270 para SIFT e SURF. Já para a categoria Scan-like na camada oculta obteve-se 1.896 neurônios para HOG e 262 para SIFT e SURF.

Tabela 4.3: Parâmetros Utilizados na Classificação.

Classificadores		Categoria	
		SCAN	SCAN-LIKE
SIFT	KNN	K = 5	K = 3
	RF	Características (200) Árvores (200)	Características (200) Árvores (175)
	SMO	Filtro (Standardize) Coeficiente C (1.0) Kernel (PolyKernel)	Sem filtro Coeficiente C (1.0) Kernel (PolyKernel)
	RNA	Épocas (200) Momento (0.2) Taxa Aprendizado (0.3)	Épocas (200) Momento (0.2) Taxa Aprendizado (0.3)
SURF	KNN	K = 5	K = 37
	RF	Características (200) Árvores (175)	Características (175) Árvores (150)
	SMO	Filtro (Standardize) Coeficiente C (1.0) Kernel (PolyKernel)	Sem Filtro Coeficiente C (1.0) Kernel (PolyKernel)
	RNA	Épocas (200) Momento (0.2) Taxa Aprendizado (0.3)	Épocas (200) Momento (0.2) Taxa Aprendizado (0.3)
HOG	KNN	K = 1	K = 1
	RF	Características (200) Árvores (200)	Características (175) Árvores (150)
	SMO	Sem filtro Coeficiente C (1.0) Kernel (PolyKernel)	Sem Filtro Coeficiente C (1.0) Kernel (PolyKernel)
	RNA	Épocas (200) Momento (0.2) Taxa Aprendizado (0.3)	Épocas (200) Momento (0.2) Taxa Aprendizado (0.3)

## 4.5 METODOLOGIA DE AVALIAÇÃO DO MÉTODO PROPOSTO

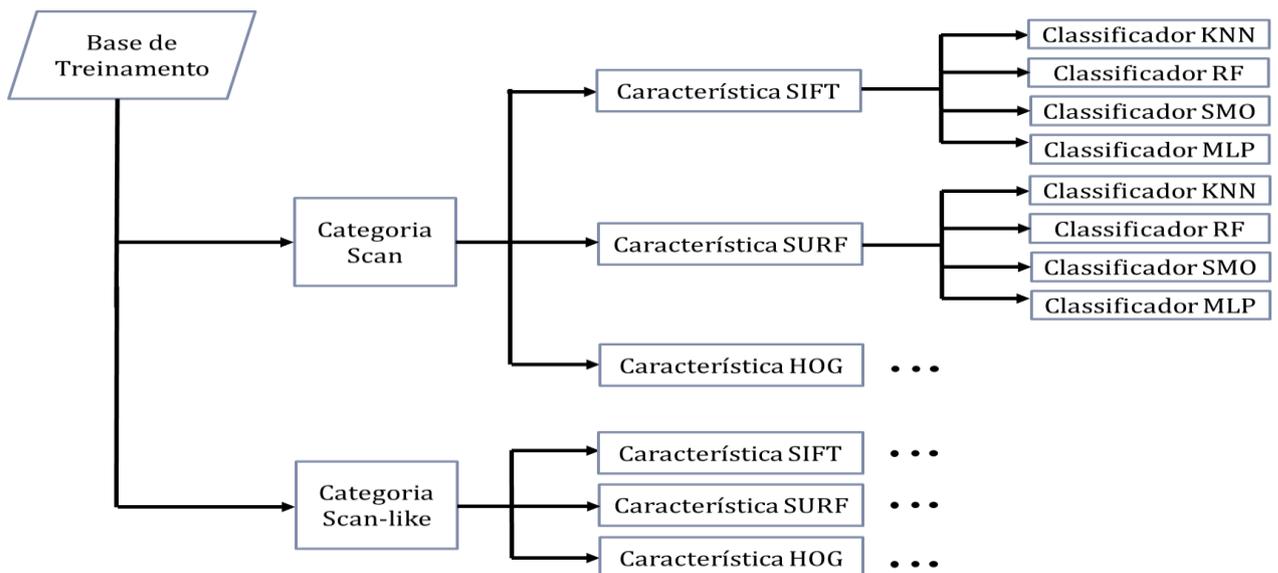
Nesta seção é apresentada a metodologia de avaliação do método proposto. Os principais testes realizados e seus objetivos em ambas as categorias são os seguintes:

- Testes com classificadores individuais (Seção 5.2): tem por objetivo verificar a validade das características texturais na classificação de folhas de plantas.
- Testes com múltiplos classificadores (Seção 5.3): visa verificar quais são as melhores estratégias de combinações de características texturais que podem trazer um resultado melhor ou pior que um classificador individual.
- Testar o desempenho do método proposto (Seção 5.4): visa avaliar o método proposto com outros trabalhos do estado da arte, promovendo um comparativo entre os resultados obtidos por outros autores que utilizaram a mesma base de dados, e a metodologia proposta.

### 4.5.1 CLASSIFICADORES INDIVIDUAIS

Nesta seção é detalhada abordagem realizada com classificadores que possuem apenas um indutor; os testes consistiram em, para cada uma das categorias presentes na base de dados de treinamento, extrair as características de SIFT, SURF e HOG das amostras e treinar quatro classificadores distintos (KNN, RF, SMO, RNA) para cada um dos conjuntos de características. O processo de criação dos classificadores individuais é detalhado na Figura 4.9.

Figura 4.9: Esquema de treinamento de classificador por categoria.



Fonte: O autor.

#### 4.5.2 SEPARAÇÃO DAS BASES DE TREINAMENTO E TESTES

Para realizar os experimentos, primeiramente foi necessário entender a separação da base de dados, para fazer o treinamento e testes dos modelos.

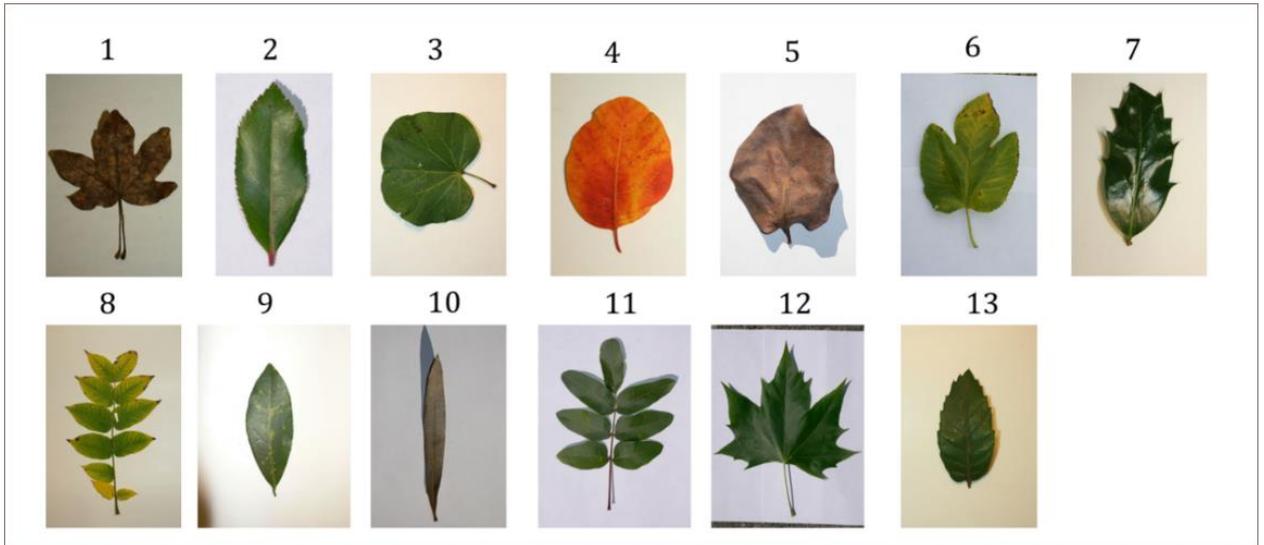
A base utilizada é disponibilizada pelo congresso ImageCLEF, portanto, a separação de amostras de treinamento e teste já está pré-definida pela organização. Contudo não é permitido incrementar a base de dados com amostras de outras bases de dados, porque seria uma forma de aumentar a robustez dos modelos gerados com mais exemplos de amostras de uma determinada espécie, facilitando assim o reconhecimento daquelas espécies (classes) que tem poucas imagens de treinamento.

A Tabela 4.4 exhibe como estão distribuídas as bases de dados de treinamento e teste para cada categoria. Essas duas categorias foram as utilizadas para os treinamentos e testes dos classificadores deste trabalho. Ainda na Tabela 4.4 consta um identificador chamado “Classe” que é seguida de números sequenciais para cada categoria. Esse identificador é para simplificar a visualização de cada amostra da base de dados, cujos alguns exemplos estão mostrados nas Figuras 4.10 e 4.11, para as categorias Scan-like e Scan, respectivamente.

Tabela 4.4: Separação entre a base de treinamento e testes.

Categoria	Classe	Espécie	Treinamento	Teste	Total
Scan	1	Acer campestre	24	22	46
	2	Acer monspessulanum	45	22	67
	3	Acer negundo	17	21	38
	4	Acer platanoides	12	10	22
	5	Aesculus hippocastanum	25	4	29
	6	Arbutus unedo	34	41	75
	7	Betula pendula	76	3	79
	8	Broussonetia papyrifera	12	58	70
	9	Castanea sativa	24	45	69
	10	Celtis australis	39	24	63
	11	Cercis siliquastrum	49	20	69
	12	Corylus avellana	55	20	75
	13	Crataegus monogyna	23	54	77
	14	Eriobotrya japonica	14	21	35
	15	Ginkgo biloba	34	15	49
	16	Juglans regia	30	4	34
	17	Laurus nobilis	36	6	42
	18	Olea europaea	125	25	150
	19	Paliurus spina christi	65	7	72
	20	Pistacia lentiscus	41	36	77
	21	Quercus coccifera	16	33	49
	22	Quercus ilex	123	18	141
	23	Quercus pubescens	24	18	42
	24	Rhamnus alaternus	54	52	106
	25	Robinia pseudoacacia	32	26	58
	26	Syringa vulgaris	51	20	71
	27	Viburnum tinus	47	94	141
	28	Vitex agnus castus	55	2	57
<b>Total</b>	<b>28</b>		<b>1182</b>	<b>721</b>	<b>1903</b>
Categoria	Classe	Espécie	Treinamento	Teste	Total
Scan-like	1	Acer campestre	12	9	21
	2	Arbutus unedo	8	1	9
	3	Cercis siliquastrum	49	11	60
	4	Cotinus coggygria	64	29	93
	5	Diospyros kaki	30	7	37
	6	Ficus carica	43	19	62
	7	Ilex aquifolium	40	26	66
	8	Juglans nigra	16	22	38
	9	Laurus nobilis	36	16	52
	10	Nerium oleander	12	8	20
	11	Pistacia terebinthus	47	1	48
	12	Platanus x hispanica	23	2	25
	13	Quercus ilex	123	29	152
<b>Total</b>	<b>13</b>		<b>503</b>	<b>180</b>	<b>683</b>
<b>Total</b>	<b>41</b>		<b>1685</b>	<b>901</b>	<b>2586</b>

Figura 4.10: Exemplos de imagem de cada espécie para a categoria Scan-like.



Fonte: O autor.

Figura 4.11: Exemplos de imagem de cada espécie para a categoria Scan.



Fonte: O autor.

## 5. RESULTADOS EXPERIMENTAIS

Todos os resultados alcançados, para facilidade de comparação, serão normalizados pela métrica ImageCLEF (Equação 2.22) e expostos na forma de matrizes de confusão. As classes dos objetos computados nas matrizes de confusão são definidas como a classe com maior *score* gerados pelos classificadores. Empates são resolvidos de forma aleatória. Na visualização das matrizes de confusão apresentadas, as linhas da matriz representam a classe real dos objetos, ao passo que as colunas apresentam as classes resultantes do processo de classificação.

Os resultados e discussões sobre os testes de classificadores individuais com as amostras das categorias Scan-like e Scan são detalhados nas Seções 5.1 e 5.2, respectivamente, enquanto os resultados das combinações de classificadores para ambas as categorias são apresentados na Seção 5.3.

### 5.1 AVALIAÇÕES DOS DESCRITORES: MELHORES RESULTADOS INDIVIDUAIS PARA A CATEGORIA SCAN-LIKE

Nesta seção são apresentados os resultados obtidos com classificadores individuais treinados com amostras da categoria Scan-like. Na Tabela 5.1 contém as taxas de acertos (TA) dos extratores e seus respectivos classificadores dos testes realizados.

Tabela 5.1: Categoria Scan-like: Melhores classificadores individuais ordenados pela taxa de acerto.

Extrator	Classificador	TA (%)
HOG	RNA	61,66
SURF	RF	57,22
HOG	SMO	57,22
SURF	SMO	52,22
HOG	KNN	50,00
SURF	RNA	49,44
SIFT	RNA	45,55
HOG	RF	45,00
SIFT	SMO	42,77
SURF	KNN	42,22
SIFT	RF	41,11
SIFT	KNN	37,77

A partir das taxas apresentadas na Tabela 5.1, as medidas de desempenho (taxas de acerto) definem claramente a formação de dois grupos, no primeiro grupo tem-se um conjunto formado pelos quatro primeiros classificadores que, acertam mais da metade das amostras disponíveis.

O segundo grupo contém todos os classificadores com taxas iguais ou inferiores a 50 pontos percentuais. Destaque que nesse grupo estão todos os classificadores que contém o extrator SIFT. Embora as diferenças entre os grupos não sejam significativas, pode se observar que os extratores do tipo HOG e SURF, conseguem alcançar uma percentagem de acerto mais acurada que SIFT.

Já em relação aos classificadores pode-se notar que não há um classificador individual que consiga se destacar, satisfazendo uma boa percentagem de acerto para todos os extratores de características. Isso se deve ao fato de que os parâmetros utilizados nos classificadores de cada extrator têm diferenças na configuração dos parâmetros, ou seja, eles são distintos.

Afim de entender melhor cada extrator, nas Tabelas 5.2, 5.3 e 5.4 são expostas as melhores matrizes de confusões para os extratores de características HOG, SURF e SIFT, respectivamente.

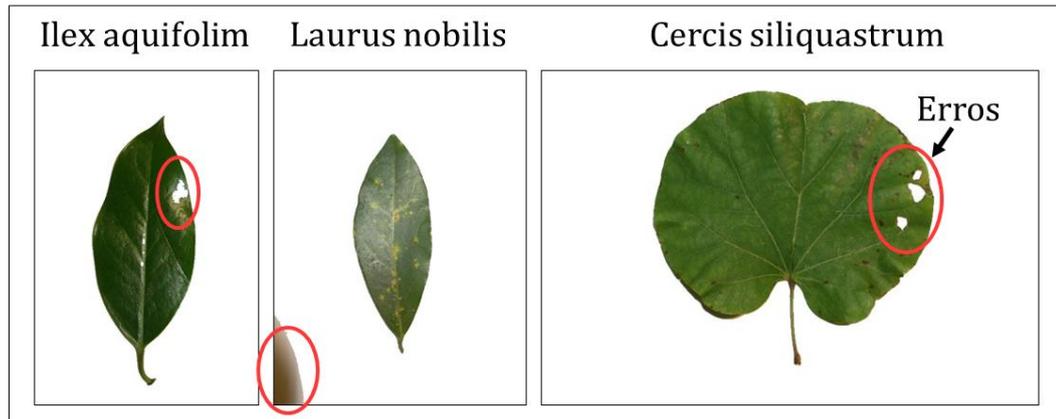
Tabela 5.2: Categoria Scan-like: Matriz de confusão do melhor classificador individual HOG+RNA.

a	b	c	d	e	f	g	h	i	j	k	l	m	id	Classes
2	0	0	0	0	0	0	0	1	0	1	2	3	a	<i>Acer campestre</i>
0	0	0	0	0	0	1	0	0	0	0	0	0	b	<i>Arbutus unedo</i>
0	0	0	0	0	0	0	0	0	0	0	0	11	c	<i>Cercis siliquastrum</i>
0	0	0	24	1	0	0	0	0	0	0	0	4	d	<i>Cotinus coggygria</i>
0	0	0	0	6	0	0	0	0	0	0	0	1	e	<i>Diospyros kaki</i>
0	0	0	5	0	14	0	0	0	0	0	0	0	f	<i>Ficus carica</i>
0	3	0	0	0	0	1	0	5	0	0	0	17	g	<i>Ilex aquifolium</i>
0	0	0	0	0	0	0	22	0	0	0	0	0	h	<i>Juglans nigra</i>
0	4	0	0	0	0	0	0	3	3	0	0	6	i	<i>Laurus nobilis</i>
0	1	0	0	0	0	0	0	0	7	0	0	0	j	<i>Nerium oleander</i>
0	0	0	0	0	0	0	0	0	0	1	0	0	k	<i>Pistacia terebinthus</i>
0	0	0	0	0	0	0	0	0	0	0	2	0	l	<i>Platanus x hispanica</i>
0	0	0	0	0	0	0	0	0	0	0	0	29	m	<i>Quercus ilex</i>

Dentre os descritores, HOG teve a melhor acurácia com o classificador RNA para a categoria Scan-like. Analisando a matriz de confusão (Tabela 5.2), verifica-se que as amostras classificadas incorretamente pelo classificador de textura HOG são aquelas em que o pré-processamento não conseguiu ocultar todo o conteúdo indesejado, ou seja, quando há a presença de ruídos (Figura 5.1) como: luminosidade (classe com id “G” *Ilex aquifolium*), objetos

projetando padrões de textura nas imagens (classe com id “I” *Laurus nobilis*), e até mesmo o não preenchimento de espaços vazios (cortes) contidas nas folhas das plantas (classe com id “C” *Cercis siliquastrum*).

Figura 5.1: Luminosidade na classe *Ilex aquifolium*, objetos projetando padrões de textura na classe *Laurus nobilis* e o não preenchimento de espaços vazios na classe *Cercis siliquastrum*.



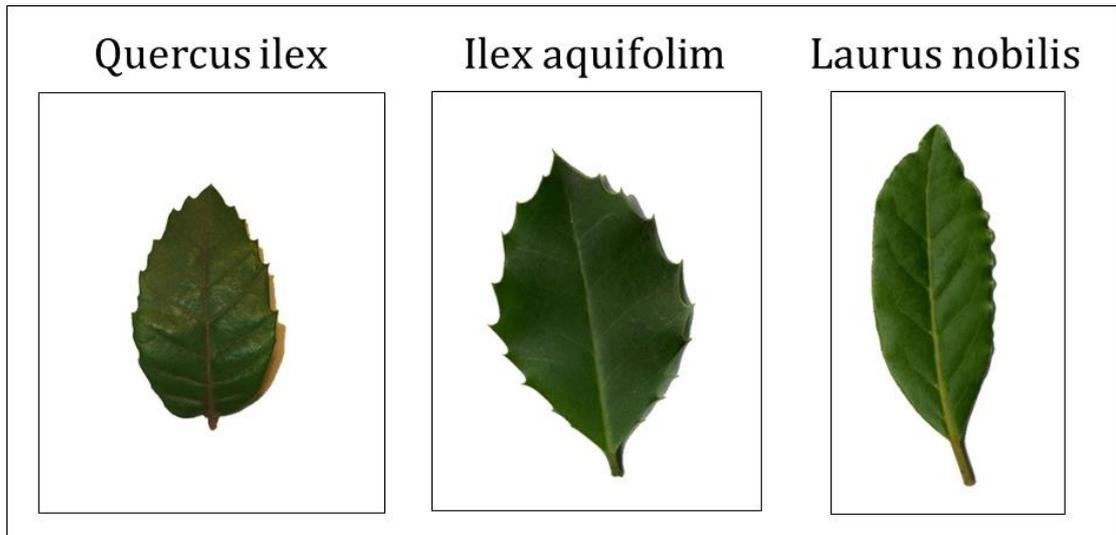
Fonte: O autor.

Já na matriz de confusão de SURF (Tabela 5.3), verifica-se que a maioria dos erros está na coluna “M” que corresponde a classe *Quercus ilex*. Esse fato acontece porque essa espécie tem morfologia similar a outras espécies. Quando isso acontece, tem um problema de confusão entre classes distintas, a Inter-classe. Por exemplo, a classe *Ilex aquifolium* é confundida com outras duas classes *Laurus nobilis* e *Quercus ilex*. De fato, observando a Figura 5.2 as folhas são parecidas, mesmo sendo de espécies diferentes; portanto, o classificador acaba errando nesses casos.

Tabela 5.3: Categorias Scan-like: Matriz de confusão do classificador individual SURF+RF.

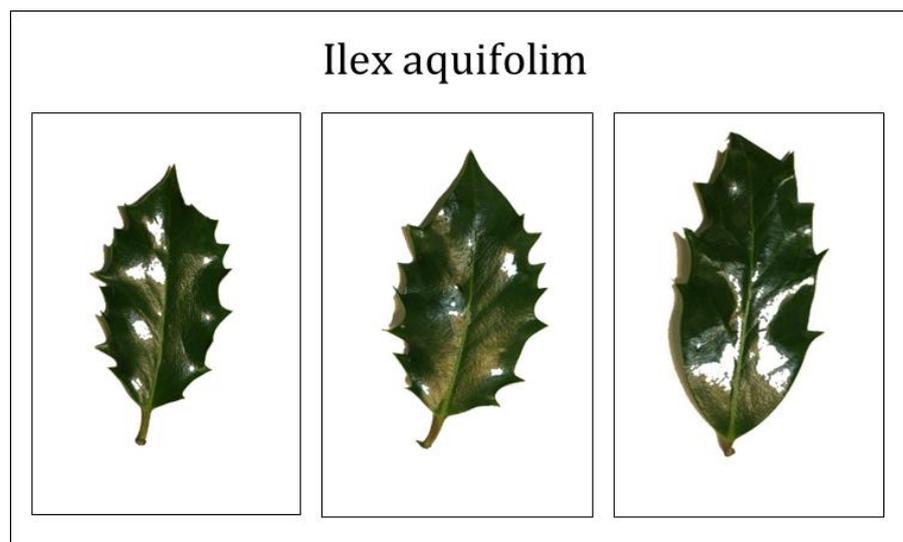
a	b	c	d	e	f	g	h	i	j	k	l	m	id	Classes
7	1	0	0	0	0	0	0	0	0	0	1	0	a	<b>Acer campestre</b>
0	0	0	0	0	0	0	0	0	0	0	0	1	b	<b>Arbutus unedo</b>
0	0	7	1	0	0	0	0	0	0	0	0	3	c	<b>Cercis siliquastrum</b>
1	0	4	7	14	0	0	0	0	0	0	1	2	d	<b>Cotinus coggygria</b>
1	0	0	0	6	0	0	0	0	0	0	0	0	e	<b>Diospyros kaki</b>
0	0	0	1	0	10	0	0	0	0	0	0	8	f	<b>Ficus carica</b>
0	0	0	0	0	0	14	0	3	0	0	0	9	g	<b>Ilex aquifolium</b>
0	0	0	0	0	0	0	21	0	0	0	0	1	h	<b>Juglans nigra</b>
0	0	0	1	0	0	0	0	0	0	0	0	15	i	<b>Laurus nobilis</b>
1	2	0	0	0	0	0	0	1	0	0	4	0	j	<b>Nerium oleander</b>
0	0	0	0	0	0	0	0	0	0	1	0	0	k	<b>Pistacia terebinthus</b>
0	0	0	0	0	0	0	0	0	0	0	1	1	l	<b>Platanus x hispanica</b>
0	0	0	0	0	0	0	0	0	0	0	0	29	m	<b>Quercus ilex</b>

Figura 5.2: Semelhança entre espécies distintas.



Fonte: O autor.

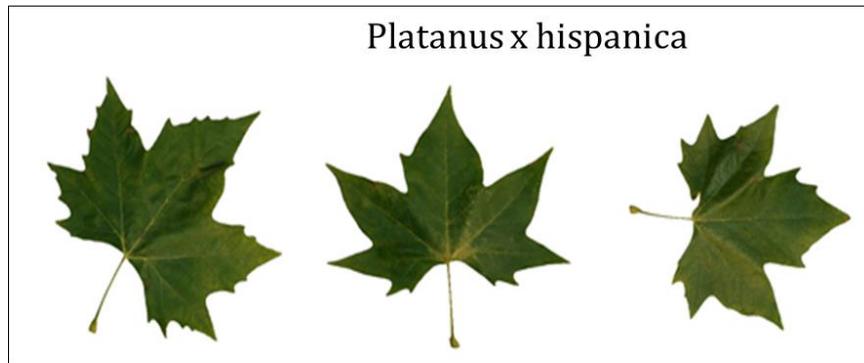
Ainda na tabela de SURF (Tabela 5.3), pode-se observar que este descritor se comporta bem, quando se diz que ele é robusto em relação a luminosidade dos objetos, dentre os três classificadores em análise (BAY; TUYTELAARS; VAN GOOL, 2006). SURF conseguiu obter um número maior de amostras corretas na classe *Ilex aquifolium*, mesmo sabendo que algumas amostras de teste, Figura 5.3, tinham uma carga excessiva de luz.

Figura 5.3: Luminosidade excessiva na espécie *Ilex aquifolium*.

Fonte: O autor.

Por fim, com relação a rotação dos objetos, Figura 5.4, pode-se observar que a matriz de confusão do descritor SIFT (Tabela 5.4) corresponde a este quesito, acertando todas as amostras de testes para a classe *Platanus x hispanica*.

Figura 5.4: Rotação nas amostras da espécie *Platanus x hispanica*.



Fonte: O autor.

Tabela 5.4: Categoria Scan-like: Matriz de confusão do classificador individual SIFT+RNA.

a	b	c	d	e	f	g	h	i	j	k	l	m	id	Classes
5	0	0	0	1	0	0	0	0	0	0	3	0	a	<i>Acer campestre</i>
0	0	0	0	0	0	1	0	0	0	0	0	0	b	<i>Arbutus unedo</i>
0	0	2	7	0	0	0	0	0	0	1	1	0	c	<i>Cercis siliquastrum</i>
0	0	9	14	1	0	0	0	0	1	0	4	0	d	<i>Cotinus coggygia</i>
0	0	1	0	2	0	0	0	0	0	0	0	4	e	<i>Diospyros kaki</i>
0	0	0	6	4	6	0	0	0	1	0	2	0	f	<i>Ficus carica</i>
1	0	0	0	5	0	11	0	2	0	0	0	7	g	<i>Ilex aquifolium</i>
0	0	0	0	0	0	0	12	0	0	2	8	0	h	<i>Juglans nigra</i>
0	1	0	2	2	0	0	0	0	0	0	1	10	i	<i>Laurus nobilis</i>
0	0	0	0	1	0	0	0	0	7	0	0	0	j	<i>Nerium oleander</i>
0	0	0	0	0	0	0	0	0	0	1	0	0	k	<i>Pistacia terebinthus</i>
0	0	0	0	0	0	0	0	0	0	0	2	0	l	<i>Platanus x hispanica</i>
0	1	0	0	2	0	6	0	0	0	0	0	20	m	<i>Quercus ilex</i>

## 5.2 AVALIAÇÕES DOS DESCRITORES: MELHORES RESULTADOS INDIVIDUAIS PARA A CATEGORIA SCAN

Os resultados obtidos com os classificadores individuais treinados com as amostras da categoria Scan são descritos nesta Seção. A Tabela 5.5 apresenta um resumo com os melhores resultados alcançados pelos classificadores com suas médias individuais.

As matrizes de confusões dos melhores classificadores individuais de cada extrator podem ser visualizadas nas Tabelas 5.6, 5.7 e 5.8 para SIFT, SURF e HOG, respectivamente.

Tabela 5.5: Categoria Scan: Melhores classificadores individuais ordenados pela taxa de acerto.

Extrator	Classificador	TA (%)
SIFT	RNA	65,46
SURF	RNA	62,96
SIFT	SMO	58,66
SURF	SMO	58,52
HOG	RNA	57,55
HOG	SMO	56,03
SURF	RF	55,20
SIFT	RF	51,59
HOG	RF	49,65
HOG	KNN	48,12
SIFT	KNN	47,98
SURF	KNN	42,57

Os resultados mostram que as medidas de desempenho (taxas de acerto) na categoria Scan tem uma nova divisão diferente da categoria Scan-like, desta vez com relação aos classificadores, onde os 6 melhores resultados são obtidos com indutores RNA e SMO, além dos extratores SIFT e SURF superaram HOG nesta categoria. Isso ocorre pois na categoria Scan as imagens não sofrem de qualquer tipo de ruído, o que faz com que os descritores SIFT e SURF obtenham pontos de características apenas no espaço das folhas, não identificando pontos característicos em ruídos como luminosidade, sombras e pontos indesejados nas amostras em que o pré-processamento não foi útil, potencializando-os.

As melhores matrizes de confusão geradas para cada extrator da categoria Scan pertence ao indutor RNA (Tabelas 5.6 e 5.7), e pode-se notar, nas matrizes de confusão, que mesmo utilizando o mesmo protocolo de indução, cada extrator de textura tem sua maneira individual para classificar amostras. Verifica-se também uma pequena semelhança entre os descritores SIFT e SURF em seus acertos na maioria das classes, por conta de utilizarem o mesmo princípio de extração em pontos locais.

Tabela 5.6: Categoria Scan: Matriz de confusão do melhor classificador individual SIFT+RNA.

a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	aa	ab	id	Classes	
7	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	11	a	Acer campestre
0	22	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	b	Acer monspessulanum
1	0	12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8	0	0	c	Acer negundo	
0	0	0	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	d	Acer platanoides	
0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	e	Aesculus hippocastanum	
0	0	0	0	0	32	2	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	5	0	0	0	0	1	f	Arbutus unedo	
0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	g	Betula pendula	
0	1	0	3	1	0	0	2	11	1	1	35	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	h	Broussonetia papyrifera	
0	0	0	0	2	0	0	0	41	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	i	Castanea sativa	
0	0	0	0	0	0	5	0	3	15	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	j	Celtis australis	
0	0	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	k	Cercis siliquastrum	
0	0	0	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	l	Corylus avellana	
0	7	2	0	0	0	0	0	0	0	0	0	27	0	0	0	0	0	3	9	0	0	5	0	0	0	0	1	m	Crataegus monogyna	
0	0	0	0	0	0	0	0	0	1	0	0	0	13	0	0	0	0	0	0	0	0	0	0	0	0	0	7	n	Eriobotrya japonica	
2	1	0	0	0	0	0	0	0	0	0	0	0	11	0	0	0	0	0	0	0	0	0	0	0	0	0	1	o	Ginkgo biloba	
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	p	Juglans regia	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	0	0	0	0	0	0	0	0	0	0	0	q	Laurus nobilis	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	25	0	0	0	0	0	0	0	0	0	0	0	r	Olea europaea	
0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	3	0	s	Paliurus spina christi	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	36	0	0	0	0	0	0	0	0	0	t	Pistacia lentiscus	
0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	31	0	0	0	0	0	u	Quercus coccifera	
0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	3	4	0	1	9	0	0	0	0	0	0	0	0	v	Quercus ilex	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	16	0	0	0	0	0	w	Quercus pubescens		
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	17	0	32	0	0	0	0	x	Rhamnus alaternus		
0	0	2	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	21	0	0	0	0	y	Robinia pseudoacacia		
0	0	0	0	0	0	0	0	0	0	2	0	0	1	0	1	0	0	0	0	0	1	0	0	0	13	2	0	z	Syringa vulgaris	
0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	2	0	0	19	0	0	0	0	0	70	aa	Viburnum tinus	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	ab	Vitex agnus castus	

Tabela 5.7: Categoria Scan: Matriz de confusão do classificador individual SURF+RNA.

a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	aa	ab	id	Classes
21	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	a	Acer campestre
0	20	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	b	Acer monspessulanum
6	0	11	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	c	Acer negundo	
0	0	0	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	d	Acer platanoides	
0	0	0	0	3	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	e	Aesculus hippocastanum	
0	0	0	0	0	24	0	0	0	0	1	0	0	0	0	0	1	1	1	0	0	10	1	2	0	0	0	f	Arbutus unedo	
0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	g	Betula pendula	
0	10	3	3	6	0	0	0	0	6	4	6	0	10	0	0	0	2	0	0	0	7	1	0	0	0	0	h	Broussonetia papyrifera	
0	0	0	1	2	1	5	0	16	3	1	3	0	6	1	1	2	0	0	0	0	1	1	0	0	0	1	i	Castanea sativa	
0	0	0	0	0	0	2	0	1	7	0	7	1	2	0	0	0	0	0	0	0	1	3	0	0	0	0	0	j	Celtis australis
0	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	13	k	Cercis siliquastrum	
0	0	0	0	0	0	0	0	0	0	0	19	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	l	Corylus avellana	
1	2	0	1	0	0	0	0	0	0	0	47	0	1	0	0	0	0	0	0	0	2	0	0	0	0	0	m	Crataegus monogyna	
0	0	0	0	0	0	0	0	2	0	0	1	0	0	1	1	0	0	0	0	0	0	0	2	0	1	13	n	Eriobotrya japonica	
0	2	0	1	0	0	0	0	0	0	0	0	12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	o	Ginkgo biloba	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	p	Juglans regia	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	3	0	0	0	0	1	0	0	0	0	0	q	Laurus nobilis	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	25	0	0	0	0	0	0	0	0	0	r	Olea europaea	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0	0	0	2	0	s	Paliurus spina christi
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	36	0	0	0	0	0	0	0	0	t	Pistacia lentiscus	
0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	31	0	0	0	0	0	0	0	u	Quercus coccifera	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	16	0	0	0	0	0	v	Quercus ilex	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	15	0	0	0	0	w	Quercus pubescens	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	3	0	0	3	0	45	0	0	0	x	Rhamnus alaternus	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	26	0	0	0	y	Robinia pseudoacacia	
0	0	0	0	0	0	0	0	0	0	5	0	0	0	0	1	1	0	0	0	0	0	0	0	12	1	0	z	Syringa vulgaris	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	7	0	0	20	0	0	0	0	66	aa	Viburnum tinus	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	ab	Vitex agnus castus

Na matriz de confusão do classificador HOG+RNA (Tabela 5.8) pode-se considerar a maioria dos erros nas classes, onde o número de amostras de teste é superior ao número de amostras de treinamento, como pode ser visto na tabela de distribuições da categoria Scan

(Tabela 4.4). Por exemplo, a classe *Broussonetia papyrifera* da matriz de confusão de HOG, tem apenas um acerto, sendo que essa classe contém 58 amostras para ser testada. Porém, quando um classificador se depara com poucas amostras, nesse caso 12 amostras de treinamento, um classificador individual não consegue gerar um modelo apropriado para classificar as 58 amostras. Esse caso também acontece nos outros descritores como SIFT e SURF.

Essa desproporção de amostras de treinamento e teste é feita pela organização do ImageCLEF justamente para avaliar o quão robusto é a metodologia proposta. Por conta disso, o congresso não permite incrementar amostras na base de dados de treinamento.

Tabela 5.8: Categoria Scan: Matriz de confusão do classificador individual HOG+RNA.

a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	aa	ab	id	Classes	
18	1	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	a	Acer campestre
2	18	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	b	Acer monspessulanum
3	4	10	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	1	c	Acer negundo
0	0	0	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	d	Acer platanoides	
0	0	0	0	3	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	e	Aesculus hippocastanum	
0	0	0	0	0	30	0	0	0	0	0	0	0	0	0	0	2	1	0	1	6	0	1	0	0	0	0	0	f	Arbutus unedo	
0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	g	Betula pendula	
6	1	2	0	3	1	17	1	0	2	0	2	1	0	0	0	0	0	9	0	0	0	0	2	0	7	0	4	h	Broussonetia papyrifera	
0	0	0	0	0	3	0	0	24	0	0	9	0	1	0	0	2	0	0	2	0	0	2	0	2	0	2	0	i	Castanea sativa	
0	0	0	0	0	0	1	0	1	18	0	2	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	j	Celtis australis	
0	0	0	0	0	0	0	0	0	0	16	0	0	0	1	0	0	0	0	0	0	0	0	3	0	0	0	0	k	Cercis siliquastrum	
0	0	0	0	0	0	0	0	0	0	0	19	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	l	Corylus avellana	
2	7	0	0	0	1	0	0	0	0	0	36	0	5	0	0	0	0	1	0	0	0	0	0	0	0	0	2	m	Crataegus monogyna	
0	0	0	0	0	2	0	0	7	0	0	1	0	0	0	0	3	0	0	0	0	0	0	1	0	0	7	0	n	Eriobotrya japonica	
2	0	0	0	0	0	0	0	0	0	0	0	0	0	11	0	0	0	0	0	0	0	1	0	0	0	1	0	o	Ginkgo biloba	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	2	0	0	0	0	0	0	0	0	p	Juglans regia	
0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	2	0	1	0	0	0	0	q	Laurus nobilis	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	17	0	0	0	5	0	1	0	0	0	2	r	Olea europaea		
0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	6	0	0	0	0	0	0	0	0	0	0	s	Paliurus spina christi	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	35	0	0	0	0	1	0	0	0	0	t	Pistacia lentiscus	
0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	32	0	0	0	0	0	0	0	u	Quercus coccifera	
0	0	0	0	0	2	0	0	0	0	1	0	0	0	0	0	0	4	0	0	9	0	0	0	0	2	0	v	Quercus ilex		
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7	11	0	0	0	0	0	0	w	Quercus pubescens		
0	0	0	0	0	14	0	0	0	0	0	0	0	0	0	0	2	0	0	3	0	31	0	0	2	0	0	x	Rhamnus alaternus		
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	1	0	0	0	23	0	0	0	0	y	Robinia pseudoacacia		
0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	1	0	0	1	0	0	0	0	0	0	14	2	0	z	Syringa vulgaris	
0	0	0	0	0	7	0	0	0	1	0	0	0	0	0	7	3	7	0	0	9	0	12	0	0	48	0	aa	Viburnum tinus		
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	ab	Vitex agnus castus	

### 5.3 COMBINAÇÃO DE CLASSIFICADORES

Os experimentos nesta seção consideram diferentes estratégias de combinação entre os 12 classificadores gerados por meio das estratégias: voto majoritário, soma, média, produto, máximo, mínimo e mediana. É importante salientar que, embora tenham sido avaliados diferentes tipos de combinação, somente os resultados obtidos com a regra produto são apresentados, por apresentarem as melhores acurácias.

### 5.3.1 RESULTADOS DA CATEGORIA SCAN

Diante da amplitude da avaliação realizada, verificou-se uma maior influência dos classificadores que obtiveram as melhores taxas de reconhecimento, onde as inclusões nos conjuntos de combinação de classificadores da Tabela 5.9 seguiram a sequência apresentada na Tabela 5.5, e aqui compuseram o conjunto de classificadores para o qual se alcançou as melhores taxas de acerto.

Tabela 5.9: Categoria Scan: Melhores resultados de combinação com a regra produto.

Combinações	TA(%)
SIFT+RNA	65,46
SIFT+RNA E SURF+RNA	67,11
SIFT+RNA, SURF+RNA E SIFT+SMO	71,56
SIFT+RNA, SURF+RNA, SIFT+SMO E SURF+SMO	72,10
<b>SIFT+RNA, SURF+RNA, SIFT+SMO, SURF+SMO E HOG+RNA</b>	<b>73,09</b>
<b>SIFT+RNA, SURF+RNA, SIFT+SMO, SURF+SMO, HOG+RNA E HOG+SMO</b>	<b>73,09</b>
SIFT+RNA, SURF+RNA, SIFT+SMO, SURF+SMO, HOG+RNA, HOG+SMO E SURF+RF	72,55
SIFT+RNA, SURF+RNA, SIFT+SMO, SURF+SMO, HOG+RNA, HOG+SMO, SURF+RF E SIFT+RF	71,14
SIFT+RNA, SURF+RNA, SIFT+SMO, SURF+SMO, HOG+RNA, HOG+SMO, SURF+RF, SIFT+RF E HOG+RF	70,99
SIFT+RNA, SURF+RNA, SIFT+SMO, SURF+SMO, HOG+RNA, HOG+SMO, SURF+RF, SIFT+RF, HOG+RF E HOG+KNN	70,54
SIFT+RNA, SURF+RNA, SIFT+SMO, SURF+SMO, HOG+RNA, HOG+SMO, SURF+RF, SIFT+RF, HOG+RF, HOG+KNN E SIFT+KNN	71,51
SIFT+RNA, SURF+RNA, SIFT+SMO, SURF+SMO, HOG+RNA, HOG+SMO, SURF+RF, SIFT+RF, HOG+RF, HOG+KNN, SIFT+KNN E SURF+KNN	72,10

Combinados, os classificadores SIFT+RNA, SURF+RNA, SIFT+SMO, SURF+SMO, HOG+RNA E HOG\_SMO alcançaram 73,09 % por meio da regra de combinação produto, obtendo um aumento de 7,63 pontos percentuais comparado com o melhor classificador individual.

As principais vantagens obtidas com a utilização de estratégias de combinação de classificadores é possibilitar que as deficiências de um determinado classificador consigam ser omitidas pelo bom desempenho de outros classificadores. Essa questão é considerável quando se faz a combinação dos classificadores SIFT+RNA, SURF+RNA, SIFT+SMO, SURF+SMO, HOG+RNA E HOG\_SMO e verifica-se que os acertos são maximizados na matriz de confusão gerada para a melhor combinação de classificadores na Tabela 5.10.

Tabela 5.10: Categoria Scan: Matriz de confusão da combinação de classificadores com a regra produto.

a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	aa	ab	id	Classes	
21	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	a	Acer campestre
0	22	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	b	Acer monspessulanum	
0	0	13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8	0	c	Acer negundo		
0	0	0	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	d	Acer platanoides		
0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	e	Aesculus hippocastanum		
0	0	0	0	0	32	1	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	5	0	0	0	1	f	Arbutus unedo		
0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	g	Betula pendula		
0	1	0	3	1	0	0	0	13	1	1	35	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	h	Broussonetia papyrifera		
0	0	0	2	2	0	0	0	33	0	0	2	0	0	0	5	0	0	0	0	0	0	1	0	0	0	0	i	Castanea sativa		
0	0	0	0	0	0	5	0	1	17	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	j	Celtis australis		
0	0	0	0	0	0	0	0	0	0	18	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	k	Cercis siliquastrum		
0	0	0	0	0	0	0	0	0	0	19	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	l	Corylus avellana		
0	0	2	0	0	0	0	0	0	0	0	49	0	0	0	0	0	0	0	0	0	0	2	0	0	0	1	m	Crataegus monogyna		
0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	13	n	Eriobotrya japonica		
2	1	0	0	0	0	0	0	0	0	0	0	0	0	12	0	0	0	0	0	0	0	0	0	0	0	0	o	Ginkgo biloba		
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	p	Juglans regia		
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	1	q	Laurus nobilis		
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	25	0	0	0	0	0	0	0	0	0	r	Olea europaea		
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7	0	0	0	0	0	0	0	0	s	Paliurus spina christi		
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	36	0	0	0	0	0	0	0	t	Pistacia lentiscus		
0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	31	0	0	0	u	Quercus coccifera		
0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	3	0	0	1	13	0	0	0	0	0	0	v	Quercus ilex		
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	14	0	0	0	0	w	Quercus pubescens		
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	50	0	0	0	x	Rhamnus alaternus		
0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	24	0	0	y	Robinia pseudoacacia		
0	0	0	0	0	0	0	0	0	0	2	0	0	1	0	1	0	0	0	0	0	0	1	0	0	0	15	z	Syringa vulgaris		
0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	2	0	0	9	0	0	0	0	80	aa	Viburnum tinus		
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	ab	Vitex agnus castus	

Fazendo uma comparação da matriz de confusão do melhor classificador individual (Tabela 5.6) com esta matriz de combinação da categoria Scan (Tabela 5.10), verifica-se que existe uma perda mínima de acertos em algumas classes. Por exemplo, na classe *Castanea sativa* teve uma perda de 8 amostras, no entanto se fizer uma análise em todas as classes, nota-se que a quantidade de incremento de acerto é maior que o número de erros nas classes, fato este que possibilita o aumento da acurácia do classificador, até mesmo naquelas imagens que contêm poucas amostras de treinamento.

### 5.3.2 RESULTADOS DA CATEGORIA SCAN-LIKE

Assim como na seção anterior, nesta categoria o método de combinação produto supera todas as outras estratégias de combinação de classificadores. A Tabela 5.11 apresenta os resultados obtidos.

Tabela 5.11: Categoria Scan-like: Melhores resultados de combinação com a regra produto.

Combinações	TA(%)
HOG+RNA	61,66
HOG+RNA E SURF+RF	62,22
HOG+RNA, SURF+RF E HOG+SMO	62,22
HOG+RNA, SURF+RF, HOG+SMO E SURF+ SMO	61,66
HOG+RNA, SURF+RF, HOG+SMO, SURF+ SMO E HOG+KNN	56,66
HOG+RNA, SURF+RF, HOG+SMO, SURF+ SMO, HOG+KNN E SURF+RNA	60,55
HOG+RNA, SURF+RF, HOG+SMO, SURF+ SMO, HOG+KNN, SURF+RNA E SIFT+RNA	63,33
HOG+RNA, SURF+RF, HOG+SMO, SURF+ SMO, HOG+KNN, SURF+RNA, SIFT+RNA E HOG+RF	61,66
HOG+RNA, SURF+RF, HOG+SMO, SURF+ SMO, HOG+KNN, SURF+RNA, SIFT+RNA, HOG+RF E SIFT+SMO	62,22
HOG+RNA, SURF+RF, HOG+SMO, SURF+ SMO, HOG+KNN, SURF+RNA, SIFT+RNA, HOG+RF, SIFT+SMO E SURF+KNN	61,11
HOG+RNA, SURF+RF, HOG+SMO, SURF+ SMO, HOG+KNN, SURF+RNA, SIFT+RNA, HOG+RF, SIFT+SMO, SURF+KNN E SIFT+RF	62,77
HOG+RNA, SURF+RF, HOG+SMO, SURF+ SMO, HOG+KNN, SURF+RNA, SIFT+RNA, HOG+RF, SIFT+SMO, SURF+KNN, SIFT+RF E SIFT+KNN	62,22
<b>HOG+RNA, SIFT+RNA, SURF+ RF, SURF+RNA</b>	<b>71,66</b>

Analisando a Tabela 5.11, percebe-se que a acurácia não aumenta à medida em que se incrementa um classificador na combinação, como observado na Tabela 5.9, da categoria Scan. Por conta disso, para melhorar os resultados encontrados, seguiu-se a mesma ideia da categoria Scan, que dentre os melhores classificadores, continha extratores e classificadores diferentes independente de suas acurácias.

Para isso, utilizou-se uma combinação de classificadores adicional contendo os três melhores extratores de características que são HOG+RNA, SIFT+RNA e SURF+RF, mesmo eles não estando dentre os melhores classificadores individuais no geral (Tabela 5.1). Para reforçar essa ideia, foi incrementado na combinação o classificador SURF+RNA e, com isto, foi alcançado uma acurácia de 71,66 pontos percentuais para a categoria Scan-like.

Notando esse resultado, cabe ressaltar que foram aproveitadas informações fornecidas pelos três extratores de características individuais distintos, mesmo com a grande variação de desempenho entre eles.

Nesta categoria observa-se claramente que as explorações de diferentes características geram resultados diferentes uns dos outros, de maneira que proporcione informações complementares umas às outras. Fazendo uma comparação entre a acurácia do melhor classificador individual e da combinação, teve-se um acréscimo de 10 pontos percentuais na categoria Scan-like.

A fim de discutir os resultados expostos através da combinação de resultados, na Tabela 5.12 é apresentada a matriz de confusão da combinação de HOG+RNA, SIFT+RNA, SURF+RF e SURF+RNA.

Tabela 5.12: Categoria Scan-like: Matriz de confusão da combinação de classificadores com a regra produto

a	b	c	d	e	f	g	h	i	j	k	l	m	id	Classes
7	0	0	0	0	0	0	0	0	0	0	2	0	a	<b>Acer campestre</b>
0	0	0	0	0	0	1	0	0	0	0	0	0	b	<b>Arbutus unedo</b>
0	0	4	0	0	0	0	0	0	0	0	0	7	c	<b>Cercis siliquastrum</b>
0	0	0	25	0	0	0	0	0	0	0	0	4	d	<b>Cotinus coggygia</b>
0	0	0	0	6	0	0	0	0	0	0	0	1	e	<b>Diospyros kaki</b>
0	0	0	4	0	15	0	0	0	0	0	0	0	f	<b>Ficus carica</b>
0	0	0	0	0	0	11	0	0	0	0	0	15	g	<b>Ilex aquifolium</b>
0	0	0	0	0	0	0	22	0	0	0	0	0	h	<b>Juglans nigra</b>
0	5	0	0	0	0	0	0	0	4	0	0	7	i	<b>Laurus nobilis</b>
0	1	0	0	0	0	0	0	0	7	0	0	0	j	<b>Nerium oleander</b>
0	0	0	0	0	0	0	0	0	0	1	0	0	k	<b>Pistacia terebinthus</b>
0	0	0	0	0	0	0	0	0	0	0	2	0	l	<b>Platanus x hispanica</b>
0	0	0	0	0	0	0	0	0	0	0	0	29	m	<b>Quercus ilex</b>

Quando se utiliza de combinação de classificadores pode-se notar na matriz de confusão que todos os resultados são incrementados, com exceção da classe *Lauros nobilis*, que em comparação com a matriz de confusão do melhor classificador individual HOG (Tabela 5.2), decai de 3 para 0. De fato, ao se analisar as melhores matrizes de confusões dos classificadores individuais compostos por esta combinação, apenas o classificador individual HOG+RNA acerta em 3 amostras e, quando se faz uma combinação pelo método produto verifica-se que este método é sensível aos baixos *scores*, portanto, se a maioria dos classificadores compostos pela combinação discordam, este está propenso a errar pela a maioria dos *scores* baixos.

Para finalizar esta seção e a discussão sobre combinação de classificadores, pode-se ressaltar que os testes com combinação de classificadores para ambas as categorias geram modelos com taxas de acerto superiores ou próximas ao melhor classificador individual. Esses resultados mostram que as técnicas de combinação podem ampliar os resultados ou até mesmo minimizar a possibilidade de se escolher um modelo ineficiente para classificar as amostras das plantas (DIETTERICH, 2000).

Analisando as amostras classificadas incorretamente pelos classificadores combinados, verifica-se que os erros mais comuns dos classificadores é quando há a presença de ruídos como sombras, luminosidade ou objetos projetando padrões de textura complexos nas imagens, ou seja, casos em que o pré-processamento não foi eficaz. Também foram identificadas limitações nos casos em que as amostras que contém um número de exemplos de treinamento inferior ao teste. Nesse caso, há uma maior probabilidade de classificação incorreta. Porém, mesmo nesses

exemplos, a combinação de classificadores gera um aumento de acertos, comparados com classificadores individuais.

Assim, os resultados obtidos indicam que a combinação de classificadores possibilita a melhor capacidade de generalização para o problema de identificação de plantas, além de proporcionar resultados melhores quando comparados a classificadores individuais.

#### 5.4 COMPARAÇÃO DOS RESULTADOS OBTIDOS

A fim de prover uma análise comparativa do método proposto com trabalhos de outros autores, o método foi posto à prova com diversas metodologias que utilizaram a mesma distribuição de base de dados. A Tabela 5.13 é uma versão reduzida da tabela da Seção 3.3, pois apresenta apenas os trabalhos que utilizaram da base de dados disponibilizada pela organização ImageCLEF, a mesma utilizada neste trabalho.

Tabela 5.13: Comparação dos resultados.

Autores	Metodologia Empregada	Extrator de Características	Método de Classificação	TA (%). Normalizada	
				Scan	Scan-like
(CASANOVA; FLORINDO; GON, 2011)	Analisaram o contorno da folha, bem como a utilização de um algoritmo de segmentação denominado OTSU	Rede Complexa + FDA	Rede Bayesiana	56,20	40,20
(BONNET et al., 2011)	Uma fusão das características locais de textura e forma é abordada	Hough, EOH e Fourier extraídos de cada ponto de Harris, redutor de vetores de dimensão (RMMH), estimador de modelos matemáticos (RANSAC)	TOP-K	68,50	46,40
(CERUTTI et al., 2011)	Esboçaram um polígono nas folhas, para extrair características do lados, base e topo a partir de sua forma.	Modelo de polígono de dois pontos (base e topo) que definem o eixo principal da folha.	KNN	53,90	54,30
(YANIKOGLU; APTOULA; TIRKAZ, 2011)	Abordaram a morfologia das plantas	Combinação dos descritores de textura: CHH, RIT; descritores de forma: Morfologias da planta e descritores da cor: Momentos estáticos de cor RGB	SVM	68,20	47,60
(MOUINE; YAHIAOUI; VERROUST-BLONDET, 2013b)	Para identificação das plantas, verificou-se suas veias baseado em suas formas	Um modelo proposto pelos autores que extrai características das veias das plantas	Fusão das respostas	78,00	70,00
(MOUINE; YAHIAOUI; VERROUST-BLONDET, 2013a)	Utilizou descritores locais de forma e contorno na margem da folha para criar triângulos com base em pontos locais extraídos	Representações Triangulares: TAR, TOA, TSL, TSLA e LSH	KNN	80,00	75,00
(MZOUGHFI et al., 2015)	Dividiu folhas de plantas compostas e simples em partes, e verificou as variações entre elas	Combinação de descritores de diferentes partes das plantas, transformada de Fourier, EOH, LEOH, Hough e DFH	KNN	81,66	-
Método Proposto	Pré-processamento com o método Otsu e Combinação de Classificadores	SIFT, SURF e HOG	RNA, KNN, RF e SMO	73,09	71,66

Os quatro primeiros métodos da Tabela 5.13 são de autores que participaram do desafio de identificação de plantas em 2011. Os trabalhos que vem em sequência são recentes e utilizaram do mesmo protocolo de imagens, e que tem como objetivo aprimorar e desenvolver métodos mais acurados que os desenvolvidos em 2011.

Ao comparar os grupos, para identificar as abordagens com melhor desempenho sobre a taxa correta de identificação, verifica-se que o melhor resultado para a categoria Scan é referente ao trabalho de (MZOUGHFI et al., 2015), onde sua metodologia é composta pela separação de partes das folhas, para fazer a identificação através de descritores distintos e, em seguida, é feita a combinação de seus resultados. De fato, trabalhos que utilizam métodos de identificação de partes de folhas são os que conseguem obter melhores resultados.

O próximo destaque de desempenho para a categoria Scan é atribuído aos dois trabalhos de Mouine e sua equipe, um que foca em característica de veias das folhas (MOUINE; YAHIAOUI; VERROUST-BLONDET, 2013b) e, uma segunda abordagem, utiliza identificação de representações de área triangulares através do contorno das folhas (MOUINE; YAHIAOUI; VERROUST-BLONDET, 2013a). Ou seja, esses trabalhos abordam uma forma diferente de extração de características, por meio das veias e contorno das folhas, conseguindo ter uma discriminação maior sobre o método proposto neste trabalho.

Já para a categoria Scan-like, os resultados alcançados com esta pesquisa conseguem ser próximos aos trabalhos de Mouine. Destaca-se, no entanto, que este trabalho apenas considera extração de características através da textura das folhas. Não foi abordado nenhum outro tipo de característica da folha como cor, contorno ou veias. Entretanto, para alcançar resultados mais acurados, uma investigação de características diferentes daquelas exploradas neste trabalho, combinadas com as já expostas, apresenta-se interessante.

De forma geral, a comparação dos resultados traz a certificação de que o método desenvolvido se apresenta competitivo. Considerando a categoria Scan, o método proposto distancia-se 4,59 pontos percentuais em relação aos quatro melhores métodos apresentado no desafio de identificação de plantas em 2011. Por outro lado, na categoria Scan-like, por apenas 3,34 pontos percentuais, o método não atinge a melhor acurácia em todos os casos. Vale lembrar que os resultados obtidos já estão normalizados na métrica disponibilizada pela organização da ImageCLEF, possibilitando assim uma comparação justa de acordo com o maior desafio de identificação de folhas de plantas por meio da imagem atualmente.

## 6. CONCLUSÃO

Neste trabalho foi apresentada uma metodologia para classificação de folhas de plantas utilizando características de textura, extraídas de imagens coletadas de diversos colaboradores. As características de textura utilizadas no trabalho são do tipo SIFT, SURF e HOG, as quais foram utilizadas para treinar distintos indutores do tipo SMO, RNA, KNN e RF.

Para execução dos experimentos foi utilizada uma base de dados de plantas disponibilizada pela organização ImageCLEF, a qual atualmente contém cerca de 2.586 imagens capturadas e divididas em duas categorias com diferentes complexidades na imagem, totalizando 41 espécies de folhas de plantas.

Os experimentos mostraram que os descritores SIFT e SURF que utilizam da extração de pontos característicos invariantes a escala e rotação, conseguiram um desempenho superior ao do descritor HOG, as amostras que atrapalharam no desempenho do HOG estavam contidas nesta categoria Scan. Já na categoria Scan-like os descritores HOG e SURF abstraíram o problema de ruído e conseguiram as melhores acurácias.

As estratégias de combinação criaram classificadores com desempenhos superiores ao melhor classificador individual nas duas categorias de imagens. A combinação de classificadores se mostrou de essencial importância, principalmente por suavizar as quedas de desempenho dos classificadores individuais, quando os testes são realizados com imagens que tinham um certo nível de alteração, pela presença de ruídos nas imagens, e em espécies de plantas que continham mais amostras de imagens de teste do que treinamento.

Como trabalhos futuros pretende-se combinar classificadores treinados com características de outras naturezas com classificadores de textura, bem como utilizar um pré-processamento robusto, que consiga ser eficaz para todas as imagens, para tentar atenuar ainda mais o efeito de similaridade e erros de classificação entre as espécies no sistema de classificação. Futuramente pretende-se também construir uma base de dados de plantas invasoras (ervas daninhas), para que se faça o reconhecimento automático deste tipo de planta, para que então seja possível prever um produto para combatê-la. Nessa abordagem pretende-se não somente utilizar combinação de classificadores, mas sim explorar as técnicas de seleção dinâmica de classificadores, que é considerada uma das principais abordagens de Sistemas de Múltiplos Classificadores (*Multi-Classifier Systems - MCS*).

## 7. REFERÊNCIAS BIBLIOGRÁFICAS

- ADAM, Z. et al. Ficus deltoidea: A Potential Alternative Medicine for Diabetes Mellitus. Evidence-Based Complementary and Alternative Medicine, v. 2012, 2012.
- ARUN, C. H.; EMMANUEL, W. R. S.; DURAIRAJ, D. C. Texture Feature Extraction for Identification of Medicinal Plants and Comparison of Different Classifiers. v. 62, n. 12, 2013.
- ASLINA, N. et al. Plant Species Identification by using Scale Invariant Feature Transform ( SIFT ) and Grid Based Colour Moment ( GBCM ). IEEE Conference on Open Systems (ICOS), p. 226–230, 2013.
- BARABASI, A.-L. Linked: How Everything Is Connected to Everything Else and What It Means for Business, Science, and Everyday Life. 2003.
- BAY, H. et al. Speeded-Up Robust Features (SURF). Computer Vision and Image Understanding, v. 110, n. 3, p. 346–359, 2008.
- BAY, H.; TUYTELAARS, T.; VAN GOOL, L. Surf: Speeded up robust features. Lecture notes in computer science, v. 3951, p. 14, 2006.
- BONNET, P. et al. The ImageCLEF 2011 plant images classification task. 2011.
- BRAGA, A. P.; CARVALHO, A. P. L. F.; LUDEMIR, T. B. Redes Neurais Artificiais Teoria e Aplicacoes. 2000.
- BREIMAN, L. Random Forests. Machine Learning, v. 45, n. 1, p. 5–32, 2001.
- CASANOVA, D.; FLORINDO, B.; GON, W. N. IFSC / USP at ImageCLEF 2012 : Plant identification task. 2011.
- CERUTTI, G. et al. Guiding Active Contours for Tree Leaf Segmentation and Identification. CLEF Notebook Papers/Labs/Workshop, v. 005, p. 1, 2011.
- CHAKI, J.; PAREKH, R.; BHATTACHARYA, S. Plant leaf recognition using texture and shape features with neural classifiers. Pattern Recognition Letters, v. 58, p. 61–68, 2015.
- CHEN, C. . Signal and Image Processing for Remote Sensing, Second Edition. Taylor & Francis, 2012. ISBN 9781439855966. 2012.
- DALAL, N. Finding People in Images and Videos. n. Phd Thesis, p. 149, 2006.
- DALAL, N.; TRIGGS, B. Histograms of Oriented Gradients for Human Detection. In CVPR, 2005.
- DEITEL, H. Xml Como Programar. 2011.
- DEMPSTER, A. P. A.; LAIRD, N. M. N.; RUBIN, D. D. B. Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society Series B Methodological, v. 39, n. 1, p. 1–38, 1977.
- DIETTERICH, T. G. Ensemble Methods in Machine Learning. Proceedings of the First International Workshop on Multiple Classifier Systems, p. 1–15, 2000.
- DUDA, R. O.; HART, P. E.; STORK, D. G. Pattern Classification 2nd Edition. 2000.
- FAWCETT, T. An introduction to ROC analysis. Pattern Recognition Letters, v. 27, n. 8, p.

861–874, 2006.

FILHO, O. M.; NETO, H. V. Processamento Digital de Imagens. *Revista Brasileira de Geofísica*, v. 21/03, n. 1, p. 331, 1999.

GOËAU, H. et al. The ImageCLEF 2011 plant images classification task. v. 2011, p. 1–23, 2011.

GONZALEZ, R. C.; WOODS, R. E. *Digital image processing*. Leonardo, v. 14, n. 3, p. 256, 2002.

GUYON, I. et al. Feature Extraction Foundations and Applications. *Soft Computing*, v. 18, n. 3-4, p. 318, 2006.

HALL, M. et al. The WEKA data mining software. *ACM SIGKDD Explorations Newsletter*, v. 11, n. 1, p. 10, 2009.

HAYKIN, S. *Neural networks: a comprehensive foundation* *The Knowledge Engineering Review*, 1999.

HERDIYENI, Y.; KADEK, N.; WAHYUNI, S. Identification using Fuzzy Local Binary Pattern and. p. 978–979, 2012.

HERDIYENI, Y.; KUSMANA, I. Fusion of Local Binary Patterns Features for Tropical Medicinal Plants Identification. p. 978–979, 2013.

HUSIN, Z. et al. Embedded portable device for herb leaves recognition using image processing techniques and neural network algorithm. *Computers and Electronics in Agriculture*, v. 89, p. 18–29, 2012.

Institution Smithsonian. Disponível em: <<http://botany.si.edu/>>. Acesso em: 10 ago. 2015.

JAIN, A. K.; DUIN, R. P. W.; MAO, J. Statistical Pattern Recognition: A Review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 22, n. 1, p. 4–37, 2000.

JAMIL, N. et al. Automatic Plant Identification: Is Shape the Key Feature. *Procedia Computer Science*, v. 76, n. Iris, p. 436–442, 2015.

JOLY, A. et al. Interactive plant identification based on social image data. *Ecological Informatics*, v. 23, p. 22–34, 2014.

KITTLER, J. et al. On combining classifiers. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1998.

KUMAR, N. et al. Leafsnap: A Computer Vision System for Automatic Plant Species Identification. p. 502–516, 2012.

KUNCHEVA, L. I. *Combining Pattern Classifiers: Methods and Algorithms*. 1 jul. 2004.

LARESE, M. G. et al. Automatic classification of legumes using leaf vein image features. *Pattern Recognition*, v. 47, n. 1, p. 158–168, 2014.

LE, T.-L.; TRAN, D.-T.; PHAM, N.-H. Kernel descriptor based plant leaf identification. 2014 4th International Conference on Image Processing Theory, Tools and Applications (IPTA), p. 1–5, out. 2014.

LOWE, D. G. Object recognition from local scale-invariant features. *International Conference on Computer Vision*, p. 7, 1999.

- MITCHELL, T. M. Machine Learning. *Annual Review Of Computer Science*, n. 1, p. 417–433, 1997.
- MOUINE, S.; YAHIAOUI, I.; VERROUST-BLONDET, A. A shape-based approach for leaf classification using multiscale triangular representation. *ICMR-3rd ACM International Conference on Multimedia Retrieval*, 2013a.
- MOUINE, S.; YAHIAOUI, I.; VERROUST-BLONDET, A. Plant Species Recognition Using Spatial Correlation Between the Leaf Margin and the Leaf Salient Points. p. 1–5, 2013b.
- MZOUGH, O. et al. Semantic-based automatic structuring of leaf images for advanced plant species identification. *Multimedia Tools and Applications*, p. 1–32, 2015.
- OJALA, T.; PIETIKÄINEN, M. Unsupervised texture segmentation using feature distributions. *Pattern Recognition*, v. 32, n. 3, p. 477–486, 1999.
- OJANSIVU, V.; HEIKKILÄ, J. Blur insensitive texture classification using local phase quantization. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, v. 5099 LNCS, p. 236–243, 2008.
- OLIVA, A. et al. Building the Gist of a Scene: The Role of Global Image Features in Recognition. *Progress in Brain Research*, 2008.
- OTSU, N. A Threshold Selection Method from Gray-Level Histograms. v. 20, n. 1, p. 62–66, 1979.
- PAULEVÉ, L.; JÉGOU, H.; AMSALEG, L. Locality sensitive hashing: A comparison of hash function types and querying mechanisms. *Pattern Recognition Letters*, v. 31, n. 11, p. 1348–1358, 2010.
- PEDRINI, H.; SCHWARTZ, W. R. *ANÁLISE DE IMAGENS DIGITAIS: Princípios, Algoritmos e Aplicações*. p. 528, 2007.
- PHAM, N.; GRARD, P. Computer aided plant identification system. p. 134–139, 2013.
- PLATT, J. C. Fast Training of Support Vector Machines Using Sequential Minimal Optimization. *Advances in kernel methods*, p. 185 – 208, 1999.
- PRICE, C. A. et al. Leaf Extraction and Analysis Framework Graphical User Interface: Segmenting and Analyzing the Structure of Leaf Veins and Areoles. *Plant Physiology*, v. 155, n. 1, p. 236–245, 2011.
- PROCHNOW, A. et al. Bioenergy from permanent grassland – A review: 1. Biogas. v. 100, n. 21, p. 4931–4944, 2009.
- QUANG-KHUE, N.; THI-LAN, L.; NGOC-HAI, P. Leaf based plant identification system for Android using SURF features in combination with Bag of Words model and supervised learning. *Advanced Technologies for Communications (ATC), 2013 International Conference on*, p. 404–407, 2013.
- ROMUALDO, L. M. et al. Use of artificial vision techniques for diagnostic of nitrogen nutritional status in maize plants. *Computers and Electronics in Agriculture*, v. 104, p. 63–70, 2014.
- SANKARAN, S. et al. A review of advanced techniques for detecting plant diseases. *Computers and Electronics in Agriculture*, v. 72, n. 1, p. 1–13, 2010.

- SHAKHNAROVICH, G.; DARRELL, T.; INDYK, P. Nearest-Neighbor Methods in Learning and Vision: Theory and Practice (Neural Information Processing). 1 mar. 2006.
- SHUGEN, W.; ZEJIN, G.; DEREN, L. Shadow detection and compensation for color aerial images. *Geo-spatial Information Science*, v. 6, n. 3, p. 20–24, jan. 2003.
- SILVA, L. O. L. A et al. Comparative assessment of feature selection and classification techniques for visual inspection of pot plant seedlings. *Computers and Electronics in Agriculture*, v. 97, p. 47–55, 2013.
- SIVIC, J. Efficient visual search of videos cast as text retrieval. *Ieee Transactions on Pattern Analysis and Machine Intelligence*, Vol. 31, No. 4, p. 591 – 605, 2009.
- SIVIC, J.; ZISSERMAN, A. Video Google: a text retrieval approach to object matching in videos. *Proceedings Ninth IEEE International Conference on Computer Vision*, n. Iccv, p. 2–9, 2003.
- SÖDERKVIST, O. J. O. Computer Vision Classification of Leaves from Swedish Trees. p. 74, 2001.
- SONKA, M.; HLAVAC, V.; BOYLE, R. *Image Processing, Analysis, and Machine Vision*. p. 180–182, 1999.
- SUROWIECKI, J. *The Wisdom of Crowds. How Collective Wisdom Shapes Business Economies Societies and Nations* New York Doubleday, p. 296, 2005.
- Tela Botanica. Disponível em: <<http://www.tela-botanica.org>>. Acesso em: 20 ago. 2015.
- THEODORIDIS, S.; KOUTROUMBAS, K. *Pattern Recognition*. p. 967, 2009.
- TSOLAKIDIS, D. G.; KOSMOPOULOS, D. I.; PAPADOURAKIS, G. Plant Leaf Recognition Using Zernike Moments and Histogram of Oriented Gradients. *Artificial Intelligence: Methods and Applications*, p. 406–417, 2014.
- WANG, L. et al. (EDS.). *Image Retrieval Based on Similarity Score Fusion from Feature Similarity Ranking Lists*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006. v. 4223
- WOODS, K.; KEGELMEYER, W. P.; BOWYER, K. Combination of multiple classifiers using local accuracy estimates. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 19, n. 4, p. 405–410, 1997.
- WU, S. G. et al. A Leaf Recognition Algorithm for Plant Classification Using Probabilistic Neural Network. p. 1–6, 2007.
- YANIKOGLU, B.; APTOULA, E.; TIRKAZ, C. Sabanci-Okan System at ImageClef 2011 : Plant identification task. p. 1–12, 2011.
- ZHAO, Z.-Q. et al. ApLeaf: An efficient android-based plant leaf identification system. *Neurocomputing*, v. 151, p. 1112–1119, 2015.

## 8. APÊNDICE A – Geração do *codebook* para os descritores SIFT e SURF

O experimento consistiu em determinar o tamanho do vocabulário (*codebook*) dos descritores SIFT e SURF que melhor represente as características de cada espécie, segundo os extratores relatados, foram gerados dicionários com tamanhos: 32, 64, 128, 256, 512, 1024, 2048 para ambos os extratores. Os experimentos a seguir foram executados com validação cruzada de 5 partições nas amostras de treinamento com todos os parâmetros citados na Tabela 8.1.

Tabela 8.1: Avaliação do *codebook* dos descritores SIFT e SURF

Tamanho do <i>Codebook</i>	SMO		KNN		RF		RNA	
	Scan	Scan-Like	Scan	Scan-Like	Scan	Scan-Like	Scan	Scan-Like
32	90.99 %	90.19 %	73.34 %	93.23 %	84.49 %	90.59 %	90.33 %	95.23 %
64	92.70 %	96.71 %	74.56 %	94.20 %	86.11 %	93.10 %	91.23 %	97.21 %
128	93.90 %	97.10 %	76.44 %	94.93 %	86.60 %	95.32 %	93.10 %	96.99 %
512	<b>95.29 %</b>	<b>98.18%</b>	<b>77.26 %</b>	<b>96.76 %</b>	<b>87.94%</b>	<b>96.56%</b>	<b>93.76%</b>	<b>97.57 %</b>
1024	94.84 %	98.00 %	76.87 %	95.11 %	85.54 %	96.14 %	91.93 %	97.10 %
2048	90.84 %	91.24 %	73.89 %	90.86 %	84.33 %	91.11 %	90.11 %	93.10 %

Os resultados em ambas as categorias Scan e Scan-like demonstraram que o melhor tamanho *codebook* a ser usado foi de 512. Apenas os resultados da melhor configuração de parâmetros entre os descritores SIFT e SURF são mostrados na tabela acima.

Analisando os resultados, é visto que quando o tamanho do *codebook* é muito pequeno ou muito grande, os classificadores tem uma dificuldade para conseguir encontrar as características essenciais para definir uma classe, pois um vocabulário muito pequeno é difícil de achar um ponto característico entre as 28 classes para a categoria Scan e 13 classes para Scan-like, ou seja, as precisões de classificação reduz rapidamente se o tamanho do *codebook* é inferior a 512, fato esse que confirma que um vocabulário compacto, com muito poucas entradas tem uma capacidade discriminativa limitado entre as classes, bem como um *codebook* de tamanho extenso por exemplo 2.048, pode confundir o classificador, pois utiliza-se de pontos característicos semelhantes em todas as imagens, pegando pontos desnecessários das amostras, fazendo uma confusão entre as classes.

Neste contexto o melhor tamanho de *codebook* encontrado para esta base de dados foi de 512 para ambas as categorias, pois é um número que satisfaz a melhor classificação para todos os classificadores.