

---

*Deep learning* aplicado à inspeção visual de  
madeira, pólen e vírus.

---

André Reis de Geus



UNIVERSIDADE FEDERAL DE UBERLÂNDIA  
FACULDADE DE COMPUTAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

Uberlândia  
2021



André Reis de Geus

*Deep learning* aplicado à inspeção visual de  
madeira, pólen e vírus.

Tese de doutorado apresentada ao Programa de Pós-graduação da Faculdade de Computação da Universidade Federal de Uberlândia como parte dos requisitos para a obtenção do título de Doutor em Ciência da Computação.

Área de concentração: Ciência da Computação

Orientador: Prof. Dr. Jefferson Rodrigo de Souza

Coorientador: Prof. Dr. André Ricardo Backes

Uberlândia

2021

Ficha Catalográfica Online do Sistema de Bibliotecas da UFU  
com dados informados pelo(a) próprio(a) autor(a).

G395 2021	<p>Geus, André Reis de, 1991- Deep learning aplicado à inspeção visual de madeira, pólen e vírus [recurso eletrônico] / André Reis de Geus. - 2021.</p> <p>Orientador: Jefferson Rodrigo de Souza. Coorientador: André Ricardo Backes. Tese (Doutorado) - Universidade Federal de Uberlândia, Pós-graduação em Ciência da Computação. Modo de acesso: Internet. Disponível em: <a href="http://doi.org/10.14393/ufu.te.2021.280">http://doi.org/10.14393/ufu.te.2021.280</a> Inclui bibliografia.</p> <p>1. Computação. I. Souza, Jefferson Rodrigo de, 1985-, (Orient.). II. Backes, André Ricardo, 1981-, (Coorient.). III. Universidade Federal de Uberlândia. Pós-graduação em Ciência da Computação. IV. Título.</p> <p style="text-align: right;">CDU: 681.3</p>
--------------	--

Bibliotecários responsáveis pela estrutura de acordo com o AACR2:

Gizele Cristine Nunes do Couto - CRB6/2091



## ATA DE DEFESA - PÓS-GRADUAÇÃO

Programa de Pós-Graduação em:	Ciência da Computação				
Defesa de:	Tese de doutorado, 10/2021, PPGCO				
Data:	19 de maio de 2021	Hora de início:	13:00	Hora de encerramento:	15:48
Matrícula do Discente:	11713CCP001				
Nome do Discente:	André Reis de Geus				
Título do Trabalho:	Deep learning aplicado à inspeção visual de madeira, pólen e vírus				
Área de concentração:	Ciência da Computação				
Linha de pesquisa:	Inteligência Artificial				
Projeto de Pesquisa de vinculação:	-				

Reuniu-se, por videoconferência, a Banca Examinadora, designada pelo Colegiado do Programa de Pós-graduação em Ciência da Computação, assim composta: Professores Doutores: Bruno Augusto Nassif Travençolo - FACOM/UFU, Marcelo Zanchetta do Nascimento - FACOM/UFU, Andrea Gomes Campos Bianch - UFOP, Cláudio Rosito Jung - INF/UFRGS, André Ricardo Backes - FACOM/UFU (coorientador) e Jefferson Rodrigo de Souza - FACOM/UFU orientador do candidato.

Os examinadores participaram desde as seguintes localidades: Andrea Gomes Campos Bianch - Ouro Preto/MG; Cláudio Rosito Jung - Porto Alegre/RS; Bruno Augusto Nassif Travençolo, Marcelo Zanchetta do Nascimento, André Ricardo Backes e Jefferson Rodrigo de Souza - Uberlândia/MG. O discente participou da cidade de Uberlândia/MG.

Iniciando os trabalhos o presidente da mesa, Prof. Dr. Jefferson Rodrigo de Souza, apresentou a Comissão Examinadora e o candidato, agradeceu a presença do público, e concedeu ao Discente a palavra para a exposição do seu trabalho. A duração da apresentação do Discente e o tempo de arguição e resposta foram conforme as normas do Programa.

A seguir o senhor presidente concedeu a palavra, pela ordem sucessivamente, aos examinadores, que passaram a arguir o candidato. Ultimada a arguição, que se desenvolveu dentro dos termos regimentais, a Banca, em sessão secreta, atribuiu o resultado final, considerando o candidato:

**Aprovado.**

Esta defesa faz parte dos requisitos necessários à obtenção do título de Doutor.

O competente diploma será expedido após cumprimento dos demais requisitos, conforme as normas do Programa, a legislação pertinente e a regulamentação interna da UFU.

Nada mais havendo a tratar foram encerrados os trabalhos. Foi lavrada a presente ata que após lida e achada conforme foi assinada pela Banca Examinadora.



Documento assinado eletronicamente por **Jefferson Rodrigo de Souza, Professor(a) do Magistério Superior**, em 19/05/2021, às 16:28, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Marcelo Zanchetta do Nascimento, Professor(a) do Magistério Superior**, em 19/05/2021, às 16:28, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **André Ricardo Backes, Professor(a) do Magistério Superior**, em 19/05/2021, às 16:28, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Andrea Gomes Campos Bianchi, Usuário Externo**, em 19/05/2021, às 16:38, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Claudio Rosito Jung, Usuário Externo**, em 19/05/2021, às 17:09, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Bruno Augusto Nassif Travençolo, Professor(a) do Magistério Superior**, em 19/05/2021, às 18:11, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site [https://www.sei.ufu.br/sei/controlador\\_externo.php?acao=documento\\_conferir&id\\_orgao\\_acesso\\_externo=0](https://www.sei.ufu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0), informando o código verificador **2778743** e o código CRC **03FDEE39**.

---

# Agradecimentos

Agradeço ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de Uberlândia pela oportunidade e à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo auxílio financeiro.

Agradeço ao meu orientador Dr. Jefferson Rodrigo de Souza e co-orientador Dr. André Ricardo Backes pelo auxílio, direcionamento deste trabalho e amizade. Também à Dra. Celia Zorzo Barcelos, ao Dr. Marcos Aurélio Batista e ao Dr. Sérgio Francisco da Silva pelas contribuições neste trabalho.

Agradeço ao Me. Alexandre Bahia Gontijo pelo auxílio na construção das bases de imagens de madeira e diversas discussões sobre as suas características anatômicas, ao pessoal da equipe RCPol pelo compartilhamento das imagens de pólen, e ao meu amigo de longa data Teodoro Terra pelas discussões e contribuições.

Agradeço aos meus pais, que sempre ofereceram suporte em todos os momentos.



*“The greatest challenge to any thinker is stating the problem in a way that will allow a solution.”*

*(Bertrand Russell)*



---

# Resumo

O reconhecimento de diferentes espécies de pólen, vírus e madeira tem se mostrado importantes para diferentes áreas, contribuindo para as investigações criminais, estudos climáticos, desenvolvimento de medicamentos, dentre outros. Entretanto, esses estudos dependem de profissionais altamente qualificados para a análise de imagens microscópicas, que se tornaram escassos e dispendiosos. Contudo, a automação destas tarefas utilizando métodos computacionais se mostra promissora. Recentemente, *Deep Learning* provou ser o melhor conjunto de técnicas para diversas tarefas de visão computacional, porém construir um conjunto de imagens com um tamanho suficiente para treinar essas técnicas a partir de sua inicialização com pesos aleatórios se mostra uma tarefa árdua. Este estudo investigou o uso de transferência de aprendizado em redes neurais profundas pré-treinadas para classificação de imagens microscópicas de amostras de pólen, madeira e vírus, comparando seus resultados com treinamento com pesos aleatórios e os métodos de extração de características pré-projetadas. Também foram introduzidos os dois maiores conjuntos em relação ao número de classes de imagens microscópicas de pólen e madeira até o momento, com 134 e 281 classes, respectivamente. Os resultados apontam que, mesmo com um elevado número de classes das bases de imagens, as metodologias propostas são capazes de atingir uma acurácia satisfatória para sua classificação. A saber, 96,24% na classificação de pólen e 98,75% na classificação de madeira. Adicionalmente, foi desenvolvido um novo dispositivo portátil combinado a um protocolo de aquisição de imagens simplificado para realização da classificação de madeiras fora de um laboratório. O resultado inicial avaliando 11 espécies amazônicas alcançou 98,13% de acurácia. Já para a classificação de vírus, as redes neurais profundas se apresentam mais eficazes atingindo a acurácia de 89%, sendo 2,8% superior quando comparadas com o melhor estudo de características pré-projetadas da literatura.

**Palavras-chave:** Classificação. *Deep Learning*. Imagens microscópicas. Pólen. Madeira. Vírus.



---

# Abstract

Pollen, virus, and wood species recognition has been shown to be an important task for a number of areas, contributing to criminal investigations, climate studies, drug developments, among others. However, these studies rely on highly qualified professionals to analyze microscopic images, which have become scarce and costly. Therefore, the automation of these tasks using computational methods is promising. Recently, Deep Learning has proven to be the ultimate set of techniques for many computer vision tasks, however, it is a very difficult task to build a data set with enough samples to train these techniques from scratch. In this study, the use of transfer learning was investigated pre-training deep neural networks for pollen, wood, and virus image classification and compared their results with training from scratch and with promising pre-designed features. It also introduced the largest data sets of pollen and wood images to the present date, with 134 and 281 classes, respectively. Results indicate that even with a high number of classes, the proposed methodologies are capable of achieving acceptable classification accuracy: 96.24% in the pollen and 98.75% in the wood classification. Additionally, a new portable device combined with a new image acquisition protocol was developed to perform the wood identification outside a laboratory. Initial results evaluating 11 amazon wood species achieved 98.13% accuracy. As for the virus classification, the deep neural networks have shown to be more effective achieving 89% accuracy, being 2.8% superior when compared to the best pre-designed features study in the literature.

**Keywords:** Classification. Deep Learning. Microscopic images. Pollen. Wood. Virus.



---

## Lista de ilustrações

Figura 1 – Exemplo de uma amostra de cada espécie de pólen investigadas em (MARCOS et al., 2015). . . . .	28
Figura 2 – Etapas de remoção do plano de fundo das imagens de pólen (MARCOS et al., 2015). . . . .	28
Figura 3 – Exemplo de uma amostra de cada espécie de pólen investigadas em (TELLO-MIJARES; FLORES, 2016) . . . . .	29
Figura 4 – Exemplo de uma amostra de cada espécie de pólen investigadas em (POZO-BAÑOS et al., 2015) . . . . .	30
Figura 5 – Exemplo de uma amostra de cada espécie de pólen investigadas em (DAOOD; RIBEIRO; BUSH, 2016). . . . .	31
Figura 6 – Exemplo de uma amostra de cada espécie de pólen investigadas em (GONCALVES et al., 2016) . . . . .	32
Figura 7 – (a) Exemplo de imagem microscópica analisada e (b) segmentação manual da imagem realizada por especialista ( <i>ground truth</i> ) (ITO et al., 2018). . . . .	34
Figura 8 – Tipos de vírus agrupados nos formatos icosaedral, regular e irregular. (KYLBERG et al., 2011) . . . . .	34
Figura 9 – Modelo de arquitetura ResNet. . . . .	42
Figura 10 – Módulo <i>inception</i> de arquitetura InceptionV3. . . . .	43
Figura 11 – Módulo <i>Fire</i> da arquitetura SqueezeNet. . . . .	43
Figura 12 – Módulo <i>Dense Block</i> de arquitetura DenseNet. . . . .	44
Figura 13 – Exemplo de um conjunto de 40 filtros de Gabor. . . . .	46
Figura 14 – Exemplos de amostras de cada tipo de pólen da base de imagens. . . .	48
Figura 15 – Amostra de pólen: (a) Lâmina Original; (b) Região de interesse. . . .	48
Figura 16 – Exemplos de amostras do gênero <i>Chamaecrista</i> ; da esquerda para direita IDs 40 ( <i>diphylla</i> ), 41 ( <i>ensiformis</i> ) e 42 ( <i>flexuosa</i> ). . . . .	50
Figura 17 – Os três principais planos de um tronco: transversal, radial, and tangencial. Adaptada de (FRANKE; QUENNEVILLE, 2011) . . . . .	50

Figura 18 – Exemplos de amostras nos três cortes: à esquerda ID 254 ( <i>Tachigali poeppigiana</i> ) e à direita ID 269 ( <i>Virola michelli</i> ). . . . .	51
Figura 19 – Exemplo de uma amostra de cada espécie de madeira no corte transversal.	52
Figura 20 – Exemplo de uma amostra de cada espécie de madeira no corte radial. . . . .	53
Figura 21 – Exemplo de uma amostra de cada espécie de madeira no corte tangencial.	54
Figura 22 – Captura das imagens com o microscópio portátil de baixo custo. . . . .	57
Figura 23 – Exemplo de uma amostra de cada espécie de madeira. . . . .	58
Figura 24 – Na primeira linha amostras da espécie do ID 1 ( <i>Alantoma decandra</i> ), enquanto na segunda linha amostras da espécie do ID 10 ( <i>Osteophloeum platyspermum</i> ). . . . .	59
Figura 25 – Vista explodida do microscópio portátil . . . . .	60
Figura 26 – Microscópio portátil focado na área em vermelho indicada no cartão de calibração, onde cada divisão equivale a 0,1mm. . . . .	60
Figura 27 – Base de imagens de vírus. Do canto superior esquerdo ao canto inferior direito: Adenovirus, Astrovirus, CCHF, Cowpox, Dengue, Ebola, Influenza, Lassa, Marburg, Norovirus, Orf, Papilloma, Rift Valley, Rotavirus and WestNile . . . . .	61
Figura 28 – Acurácia comparando a inicialização aleatória dos pesos, ajuste-fino e CNN como extrator de características. . . . .	65
Figura 29 – Exemplos de amostras dos cortes radial, tangencial e transversal de 16 espécies dentre as 281 espécies da base de imagens de madeira. . . . .	68
Figura 30 – Gráficos de barra da acurácia sobre épocas para cada arquitetura. . . . .	70
Figura 31 – Exemplo de variações em uma classe de vírus. Fonte: (KYLBERG; UPPSTRÖM; SINTORN, 2011) . . . . .	71
Figura 32 – Gráficos da perda média e desvio padrão dos <i> folds</i> em cada época para cada arquitetura. . . . .	72
Figura 33 – Gráficos da perda média e desvio padrão dos <i> folds</i> em cada época da arquitetura SqueezeNet usando 100 épocas. . . . .	73
Figura 34 – Gráficos de barra da acurácia sobre épocas para cada arquitetura. . . . .	73
Figura 35 – Acurácia média por espécie de madeira. . . . .	76
Figura 36 – Exemplos de amostras de madeira e seus respectivos mapas de calor gerados pelo método Grad-CAM++ utilizando as arquiteturas SqueezeNet (segunda linha) e DenseNet (terceira linha). . . . .	77
Figura 37 – Amostras das espécies de ID 1 ( <i>Alantoma decandra</i> ), ID 2 ( <i>Caraipa densifolia</i> ), ID 3 ( <i>Cariniana micrantha</i> ), ID 4 ( <i>Caryocar villosum</i> , ID 6 ( <i>Dipteryx odorata</i> ) e ID 9 ( <i>Lueheopsis duckeana</i> ); e abaixo de cada uma a sobreposição do mapa de calor gerado pelo método Grad-CAM++. . . . .	79

Figura 38 – Amostras das espécies de ID 5 (*Clarisia racemosa*), ID 7 (*Goupia glabra*), ID 8 (*Handroanthus incanus*), ID 10 (*Osteophloeum platyspermum*) e ID 11 (*Pouteria caimito*); e abaixo de cada uma a sobreposição do mapa de calor gerado pelo método Grad-CAM++. . . . . 80



---

## Lista de tabelas

Tabela 1 – Sumarização de estudos para classificação de madeira. . . . .	37
Tabela 2 – Sumarização de estudos para classificação de madeira. . . . .	38
Tabela 3 – Espécie, ID e número de amostras (#) da base de imagens de pólen. . . . .	49
Tabela 4 – Espécies e ID da base de imagens de madeira. . . . .	55
Tabela 5 – Espécies e ID da base de imagens de madeira. . . . .	56
Tabela 6 – Espécies e ID da base de imagens de madeira. . . . .	58
Tabela 7 – Descrição dos métodos de característica pré-projetados. Não foram experimentados Filtros Gabor individualmente devido ao número elevado de características, que resultam em um alto custo computacional para classificação. . . . .	65
Tabela 8 – Resultados obtidos para métodos de extração de características pré-projetadas comparados a aprendizado de características com ajuste-fino em diferentes arquiteturas de CNN. . . . .	66
Tabela 9 – Métodos de características pré-projetados comparados a aprendizado de características. . . . .	68
Tabela 10 – Número de parâmetros em cada arquitetura de CNN e resultados de desempenho em tempo de inferência. . . . .	69
Tabela 11 – Acurácia média das arquiteturas de CNN. . . . .	69
Tabela 12 – Comparação da acurácia média de métodos aplicados à classificação da base de imagens de vírus. . . . .	71
Tabela 13 – Métodos de características pré-projetados comparados a aprendizado de características. . . . .	74



---

## Lista de siglas

**CNN** Convolutional Neural Network

**CLBP** Completed Local Binary Patterns

**CRFK** Crandell-Rees Feline Kidney

**DTM** Discrete Tchebichef Moments

**FDA** Fisher Discriminant Analysis

**GLCM** Gray-Level Co-occurrence Matrices

**HOG** Histogram of Oriented Gradients

**ICA** Independent Component Analysis

**kNN** K Nearest Neighbor

**LGF** Log-Gabor Filters

**LBP** Local Binary Patterns

**LDA** Linear Discriminant Analysis

**LDP** Local Directional Patterns

**LPQ** Local Phase Quantization

**LQP** Local Quinary Patterns

**MLP** Multilayer Perceptron

**AM** Aura Matrix

**PSO** Particle Swarm Optimization

**PCA** Principal Component Analysis

**RF** Random Forest

**RCPol** Rede de Catálogos Polínicos online

**ReLU** Rectified Linear Unit

**RNN** Randomized Neural Network

**SVM** Support Vector Machine

**SGD** Stochastic Gradient Descent

---

# Sumário

<b>1</b>	<b>INTRODUÇÃO</b> . . . . .	<b>23</b>
<b>1.1</b>	<b>Motivação</b> . . . . .	<b>24</b>
<b>1.2</b>	<b>Objetivos</b> . . . . .	<b>26</b>
<b>1.3</b>	<b>Organização</b> . . . . .	<b>26</b>
<b>2</b>	<b>TRABALHOS RELACIONADOS</b> . . . . .	<b>27</b>
<b>2.1</b>	<b>Classificação de pólen</b> . . . . .	<b>27</b>
<b>2.2</b>	<b>Classificação de vírus</b> . . . . .	<b>33</b>
<b>2.3</b>	<b>Classificação de madeira</b> . . . . .	<b>35</b>
<b>2.4</b>	<b>Considerações finais</b> . . . . .	<b>36</b>
<b>3</b>	<b>MATERIAIS E MÉTODOS</b> . . . . .	<b>39</b>
<b>3.1</b>	<b>Redes Neurais Convolucionais</b> . . . . .	<b>39</b>
3.1.1	Transferência de Aprendizado . . . . .	40
3.1.2	Arquiteturas de CNN . . . . .	41
<b>3.2</b>	<b>Extratores de características pré-projetados</b> . . . . .	<b>44</b>
3.2.1	LBP . . . . .	44
3.2.2	LPQ . . . . .	45
3.2.3	RI-LPQ . . . . .	45
3.2.4	Filtro de Gabor 2D . . . . .	46
3.2.5	HOG . . . . .	46
<b>3.3</b>	<b>Bases de Imagens</b> . . . . .	<b>47</b>
3.3.1	Base de Pólen . . . . .	47
3.3.2	Base de Madeira . . . . .	50
3.3.3	Base de Madeira - Dispositivo portátil . . . . .	57
3.3.4	Base de Vírus . . . . .	60
<b>3.4</b>	<b>Considerações finais</b> . . . . .	<b>61</b>

4	EXPERIMENTOS E ANÁLISE DOS RESULTADOS . . . . .	63
4.1	Treinamento das CNNs . . . . .	63
4.2	Classificação de pólen . . . . .	64
4.3	Classificação de madeira . . . . .	66
4.4	Classificação de vírus . . . . .	69
4.5	Classificação de madeira - Dispositivo portátil . . . . .	71
4.6	Considerações finais . . . . .	78
5	CONCLUSÃO . . . . .	81
5.1	Contribuições . . . . .	82
5.2	Trabalhos futuros . . . . .	82
5.3	Contribuições em Produção Bibliográfica . . . . .	83
	REFERÊNCIAS . . . . .	85

---

# Introdução

Por meio de desenvolvimentos em *software* e *hardware*, os sistemas apoiados por computador tornaram-se presentes e essenciais nas mais diversas áreas do conhecimento humano. Esses sistemas vêm sendo utilizados em áreas de aplicação cada vez mais diversificadas, tais como agronomia, na contagem de plantas em imagens aéreas (KITANO et al., 2019); negócios, na previsão do preço de ações (BAO; YUE; RAO, 2017); medicina, na identificação de câncer em tomografia computadorizada (LAKSHMANAPRABU et al., 2019); entre outras. Tais aplicações normalmente lidam com conjuntos de dados complexos como sons, imagens, vídeos, cadeias de DNA, dados geo-referenciados, de telemetria, de engenharia, exames médicos e séries temporais. Consequentemente, o volume dessas informações manipulada por tais sistemas têm crescido tanto no número de objetos que compõem os conjuntos de dados, quanto na quantidade e complexidade das características. Esse panorama tem motivado o desenvolvimento de técnicas e ferramentas computacionais que possam, de maneira inteligente e automática, auxiliar os seres humanos nas tarefas de análise, manipulação, compreensão e extração de conhecimentos dos dados para diversos fins.

O processo tradicional de classificação de imagens envolve a extração de características, como forma, cor e textura, de modo a encontrar um conjunto de características que maximize a diferenciação de seus grupos, também chamados de classes. Neste contexto, o termo características pré-projetadas utilizado nesse trabalho se refere a este grupo de algoritmos que são vastos e necessitam de avaliação em cada domínio aplicado. Vale ressaltar que mesmo em domínios bem definidos, a literatura investiga diferentes extratores, expondo a dificuldade em encontrar uma boa representação para cada problema através de características relevantes das imagens. Isso motiva a criação de técnicas que possam de maneira automática extrair características, da mesma forma como as Redes Neurais Artificiais aprendem a realizar a classificação.

CNN (do inglês, *Convolutional Neural Network*) é um dos mais populares algoritmos de *Deep Learning* (LECUN; BENGIO; HINTON, 2015). Sua popularidade se deve a ampla gama de tarefas em que pode ser utilizado, como a classificação de imagens, vídeo ou som.

Isso se deve ao fato desse tipo de algoritmo ser capaz de aprender os dados de entrada sem a necessidade de extração de características pré-projetadas. O ressurgimento desse tipo de arquitetura, inicialmente proposta com o nome de *Neocognitron* em (FUKUSHIMA, 1980), foi impulsionada na última década através do aumento do poder computacional possibilitado através de paralelismo em GPU e também aumento das bases disponíveis.

Dentre as diversas aplicações, destacam-se a de interesse deste projeto de pesquisa, a classificação de imagens microscópicas de pólen, madeira e vírus. A grande maioria dos conjuntos de imagens microscópicas apresentam um grande desafio para sua classificação, já que apresentam grande variabilidade dentre suas classes. Esse fenômeno é descrito em (VU et al., 2019), onde estruturas celulares cancerosas tem núcleo diferente das demais e influenciam o tecido em várias formas. O mesmo ocorre nas características anatômicas dos cortes em tronco de árvore, devido às inúmeras variações climáticas que ocorrem ao longo dos anos, sofrem modificações principalmente nos anéis de crescimento (JAYAWICKRAMA et al., 1997).

Neste trabalho, além da introdução de duas bases de imagens microscópicas de pólen e madeira e respectivas metodologias para a classificação destas imagens, são explorados estudos comparativos entre aprendizado automático de características e características pré-projetadas, para levantar vantagens e limitações destes métodos computacionais.

## 1.1 Motivação

O combate à extração ilegal de madeira, certificação e correta tributação é a principal motivação para a classificação das diferentes espécies de madeira. Em seu ecossistema, as espécies arbóreas possuem muitas características anatômicas, como aspectos de sementes, flores, folhas, frutos e formato do tronco, o que permite a correta classificação das espécies. Usando apenas uma seção da madeira, a análise principal é realizada com base nos padrões de textura presentes. Esses padrões observados na madeira por um especialista incluem a disposição de vasos ou poros, parênquima da madeira, fibras, floema, traços de látex e canais intercelulares. Tais características de textura permanecem na madeira mesmo após tratamentos químicos. Para uma classificação correta, as texturas de madeira são comparadas com materiais científicos armazenados em uma xiloteca ou banco de dados como o InsideWood (WHEELER, 2011). No entanto, existe um número limitado de especialistas na classificação da madeira que, além de gastar tempo considerável no procedimento, podem falhar devido à fadiga de trabalho.

Métodos alternativos para facilitar a classificação de espécies de madeira através de amostras do tronco tem sido gradualmente desenvolvidos. Os métodos de DNA, baseados em marcadores moleculares, fornecem uma alternativa, porém, além de ser dispendioso coletar amostras, é necessário um especialista em genética para fazer a interpretação dos dados. A classificação de espécies de madeira usando visão computacional é um campo

de estudo emergente, já que o arranjo dos grãos da madeira pode ser caracterizado por métodos de análise de textura. De acordo com (ZAMRI et al., 2016), vários sistemas automáticos de classificação de espécies de madeira são mais precisos do que os sistemas tradicionais de classificação de madeira baseados exclusivamente na experiência humana e na análise de padrões de texturas. Com o surgimento de microscópios eletrônicos e câmeras digitais de alta resolução, é possível capturar imagens da madeira com uma visualização detalhada de seus grãos, arranjos e até de suas estruturas celulares. Tais características anatômicas da madeira são dominantes no desenvolvimento de sistemas biométricos para classificação de espécies de madeira através de imagens de cortes no tronco.

A virologia é uma outra área que recebeu destaque recentemente, principalmente pelo estado de pandemia que estamos vivenciando. Devido ao tamanho reduzido de um vírus, muito inferior ao de uma célula, eles não podem ser visualizados por microscópios ópticos. Nesse contexto, através do uso de imagens de microscopia eletrônica de transmissão foi possível identificar e distinguir tipos de vírus através de sua morfologia (VALE et al., 2010). A característica mais importante em diversos vírus é o envelope externo feito de glicoproteínas ou lipídios (BANERJEE; MUKHOPADHYAY, 2016). É através dele que os vírus recebem nomes populares como o caso do SARS-CoV-2, identificado como corona pela sua glicoproteína *spike* se parecer com uma coroa (BENVENUTO et al., 2020).

A palinologia, lida com o estudo de pólen de plantas, esporos e organismos planctônicos microscópicos, decorre da formação morfológica de tais partículas ou grãos. As principais características dos grãos de pólen estão relacionadas ao tamanho, forma (polaridade, simetria), aberturas e ornamentação. Os grãos de pólen não estão sujeitos a deterioração física devido a variações nas condições ambientais, tornando suas características geneticamente estabelecidas, estáveis e de grande valor diagnóstico. A identificação do pólen ajudou a delinear a distribuição geográfica de muitos grupos de plantas de milhões de anos atrás até o presente (AUGUSTYN et al., 2017). Estudos palinológicos têm sido úteis para emitir alertas para pessoas com alergias ao pólen (D'AMATO et al., 2007) e estabelecer o local ou período sazonal para cenas de crime (SZIBOR et al., 1998), para certificados de produção de mel (ALJOHAR et al., 2018), determinar práticas agrícolas que ocorrem em sítios arqueológicos (AUGUSTYN et al., 2017) e traçar relações entre diferentes grupos de plantas e linhas evolutivas (SARWAR; HOSHINO; ARAKI, 2015).

De maneira semelhante à classificação de madeiras e vírus, a classificação de pólen é usualmente realizada por meio da inspeção visual da amostra por um profissional altamente qualificado. Catálogos polínicos, como o disponibilizado pela Rede de Catálogos Polínicos online (RCPol), permitem facilitar o procedimento limitando grupos de espécies por características relevantes. No entanto, assim como na classificação de espécies de madeira, aqui também existe um número limitado de especialistas disponíveis, sendo o mesmo um trabalho moroso e suscetível a falhas por fadiga.

## 1.2 Objetivos

Esta tese tem como objetivo principal estudar e comparar diferentes métodos computacionais para a classificação de imagens microscópicas de pólen, madeira e vírus. Para tanto, serão estudadas técnicas que envolvam o aprendizado automático de características através de CNNs e o uso de características pré-projetadas.

Os objetivos específicos deste trabalho podem então ser descritos como:

- ❑ Revisão da literatura na classificação de imagens microscópicas de pólen, madeira e vírus;
- ❑ Estudar e comparar diferentes arquiteturas de redes neurais convolucionais (CNN);
- ❑ Propor metodologias de características pré-projetadas para comparação com CNNs;
- ❑ Verificar as vantagens e desvantagens de cada arquitetura CNN na classificação de imagens microscópicas;
- ❑ Avaliar diferentes cortes no tronco de árvores no processo de classificação;
- ❑ Propor um protocolo simplificado de aquisição de imagens de madeira em campo;
- ❑ Avaliar as correlações entre as características aprendidas pelas CNNs e as observadas por um especialista em anatomia de madeira;
- ❑ Desenvolver um dispositivo portátil para aquisição de imagens de madeira em campo.

## 1.3 Organização

A tese está organizada da seguinte forma: O Capítulo 2 apresenta os trabalhos relacionados, divididos na seção 2.1, que aborda os trabalhos recentes na classificação de pólen, na seção 2.2, que aborda os trabalhos recentes na classificação de vírus, e na seção 2.3, que aborda os trabalhos recentes na classificação de madeira. No Capítulo 3 são apresentados os métodos de classificação, extratores de características pré-projetados e arquiteturas de CNN usadas nos experimentos comparativos. No Capítulo 4 são apresentados os experimentos realizados e resultados obtidos nas metodologias desenvolvidas. Por fim, o Capítulo 5 sumariza as contribuições obtidas com base nos trabalhos desenvolvidos e publicados em periódicos e conferências da área.

---

## Trabalhos Relacionados

O reconhecimento de imagens baseado em computador depende de duas abordagens principais para a extração de características: a extração de características pré-projetadas e o aprendizado automático de características. Na extração de características pré-projetadas algumas informações visuais específicas são extraídas considerando os algoritmos que descrevem os atributos com base nas propriedades de cor, textura ou forma. Já o aprendizado automático de características utiliza algoritmos de aprendizado que extraem significado dos dados usando uma hierarquia de múltiplas camadas, que imitam as redes neurais do nosso cérebro. O melhor exemplo dessa categoria de algoritmos são as CNNs, que através de convoluções consecutivas tem se mostrado mais eficazes que as tradicionais perceptron multicamadas. A seguir, são apresentados os principais trabalhos para o reconhecimento de madeira, pólen e vírus por meio de imagens.

### 2.1 Classificação de pólen

A palinologia tem se mostrado importante em diversas aplicações, tendo como primordial o mapeamento de sua morfologia para o reconhecimento das espécies. Diversos estudos da literatura levantaram informações morfológicas sobre diferentes grupos de espécies analisando sua proximidade entre si. Entretanto, estudos computacionais surgiram nas últimas décadas investigando a automatização desse processo através de descritores matemáticos. De acordo com diferentes pesquisas presentes na literatura, é possível verificar que características de textura possuem grande importância na identificação do pólen (LI; FLENLEY, 1999; MARCOS et al., 2015; RONNEBERGER; BURKHARDT; SCHULTZ, 2002). Há também várias pesquisas que empregam essas características de textura combinadas com a informação de sua forma, a fim de aumentar a capacidade de discriminação das amostras (ALLEN et al., 2008; TRELOAR; TAYLOR; FLENLEY, 2004; TELLO-MIJARES; FLORES, 2016; LI et al., 2004; POZO-BAÑOS et al., 2015; HODGSON et al., 2005; ZHANG et al., 2004; PUNYASENA et al., 2012; REDONDO et al., 2015).

Na metodologia proposta em (MARCOS et al., 2015), por exemplo, uma nova base de imagens contendo 15 espécies de pólen foi introduzida contendo um total de 1800 imagens divididas igualmente entre as espécies. A Figura 1 ilustra um exemplo de cada classe.

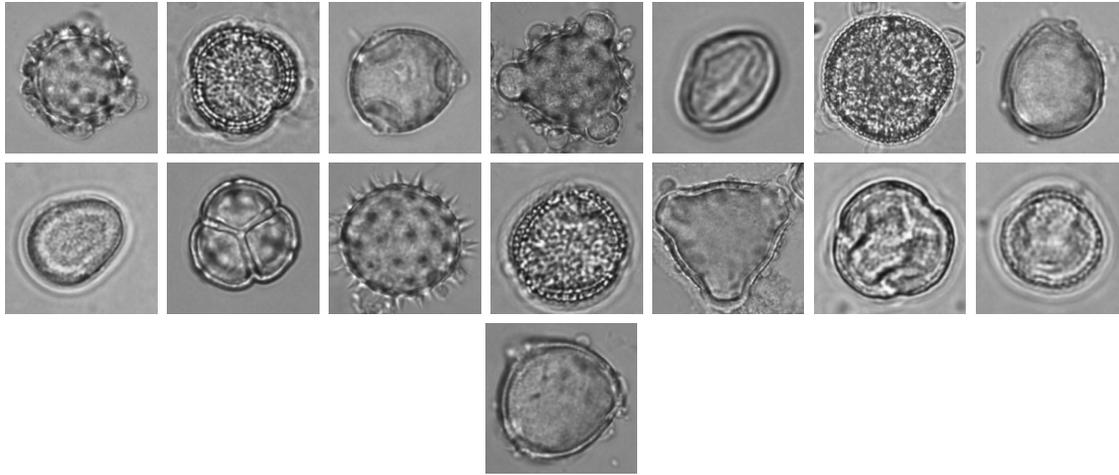


Figura 1 – Exemplo de uma amostra de cada espécie de pólen investigadas em (MARCOS et al., 2015).

Para a classificação das espécies, foram inicialmente realizados procedimentos para a remoção do plano de fundo das imagens, ilustrados resumidamente na Figura 2. Em seguida foram extraídas e combinadas as características de texturas obtidas com os métodos GLCM (do inglês, *Gray-level Co-occurrence Matrices*), LGF (do inglês, *Log-Gabor Filters*), LBP (do inglês, *Local Binary Patterns*) e DTM (do inglês, *Discrete Tchebichef Moments*), sendo o conjunto de atributos reduzidos pela aplicação do FDA (do inglês, *Fisher Discriminant Analysis*). A classificação das imagens foi realizada através do método kNN (do inglês, *K Nearest Neighbor*) com validação cruzada k-fold, atingindo 94,8% de acurácia.

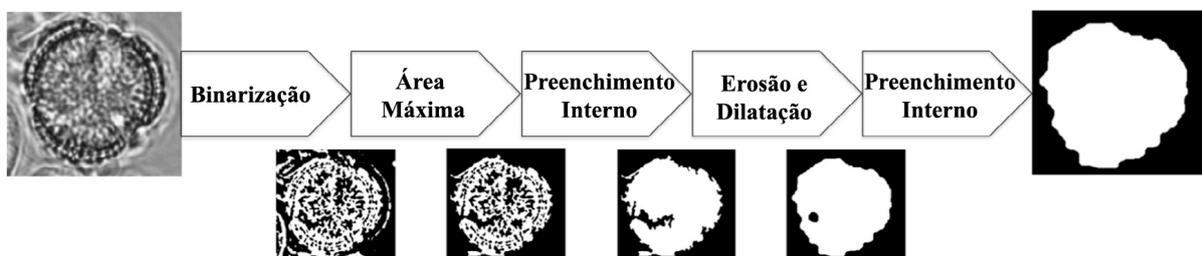


Figura 2 – Etapas de remoção do plano de fundo das imagens de pólen (MARCOS et al., 2015).

A alta acurácia apresentada no estudo é justificada pelas vantagens dos extratores de características de textura que evitam a dependência de posição dos grãos de pólen.

Entretanto, os autores destacam algumas limitações da metodologia: a baixa quantidade de amostras do conjunto, que impossibilita treinamento de classificadores mais promissores como redes neurais; a ausência de uma busca pelo valor ótimo de  $k$  no classificador kNN, que evitaria tendência nos resultados; e a ausência de uma metodologia para segmentação automática dos grãos nas imagens de lâminas microscópicas, que normalmente estão em meio a outros tecidos celulares e exigem seu recorte manual.

Em (REDONDO et al., 2015), o mesmo grupo de autores desenvolveu melhorias em seu trabalho anterior (MARCOS et al., 2015). A metodologia propõe um novo método de segmentação do contorno interno e remoção do plano de fundo, analisando o benefício de descritores morfológicos e estatísticos sobre as regiões segmentadas. Além disso, foram avaliados três classificadores: SVM (do inglês, *Support Vector Machine*), FDA e RF (do inglês, *Random Forest*). Este último apresentou melhor resultado, elevando a acurácia para 99.4%, quase 5% superior ao estudo anterior.

Um ponto interessante levantado pelos autores é a preocupação com a aplicabilidade prática da metodologia. Grãos de pólen frequentemente apresentam exemplos mal formados, oclusos e sobrepostos que dificultam sua classificação. Os autores sugerem como trabalho futuro uma abordagem para contornar esse problema. A inclusão de uma nova classe “duvidosa” para amostras fora do padrão, também chamados de *outliers*, fortaleceria a confiança do classificador. A ideia é minimizar os falsos-positivos, não se preocupando com os falso-negativos atribuídos à classe duvidosa. Isso se torna desejável, já que em uma amostra coletada centenas de grãos da mesma espécie estão presentes e podem ser testados.

(TELLO-MIJARES; FLORES, 2016) introduziram uma nova base de imagens de pólen com 12 espécies ilustradas na Figura 3. A base contém 390 imagens de lâminas microscópicas, compreendendo um total de 618 grãos de pólen, com número de amostras variado entre as espécies. Para a classificação foram usados descritores geométricos, estatísticas de textura de primeira ordem e estatísticas de textura baseadas em GLCM de segunda ordem obtidas do modelo de cores  $L^*a^*b^*$ . Resultados experimentais atingiram até 96,1% de precisão comparando os classificadores RF, BN (do inglês, *Bayesian Network*) e MLP (do inglês, *Multilayer Perceptron*).

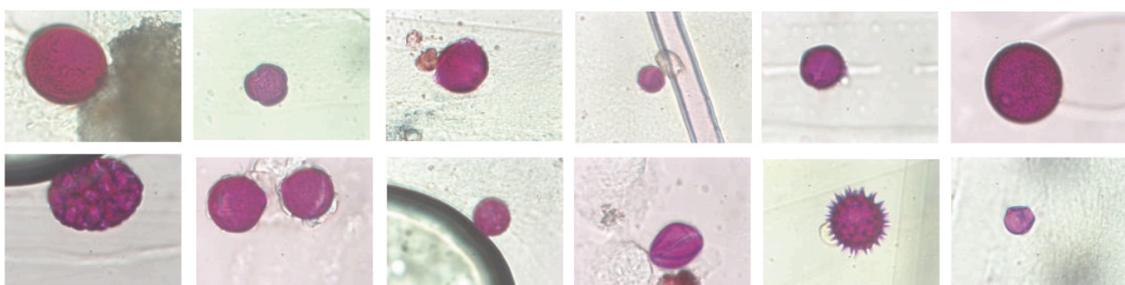


Figura 3 – Exemplo de uma amostra de cada espécie de pólen investigadas em (TELLO-MIJARES; FLORES, 2016)

O melhor resultado apresentado pelo classificador MLP demonstra seu potencial para atingir melhores resultados quando comparado a métodos de classificação clássicos utilizados em aplicações de bioinformática e medicina. Os autores também destacam a superioridade de seu método de segmentação proposto devido à inclusão de informações de cor no procedimento. Isto se justifica pelos diversos trabalhos na literatura que convertem as imagens em escala de cinza, perdendo o benefício das informações de cor.

(POZO-BAÑOS et al., 2015) combinaram características geométricas (área, centróide, área convexa, diâmetro equivalente, perímetro, entre outras) e de textura (GLCM, contraste, correlação, energia, entropia, entre outras), para classificação de um conjunto de imagens de pólen de 17 classes ilustradas na Figura 4. Além disso, foram experimentados os redutores de dimensionalidade LDA (do inglês, *Linear Discriminant Analysis*), PCA (do inglês, *Principal Component Analysis*) e ICA (do inglês, *Independent Component Analysis*). Para classificação, foram experimentados MLP e SVM com validação cruzada k-fold variando valor de  $k$ . O melhor resultado atingiu 95,27% de acurácia utilizando as características reduzidas por LDA e k-fold ( $k = 10$ ), combinando ambos classificadores.

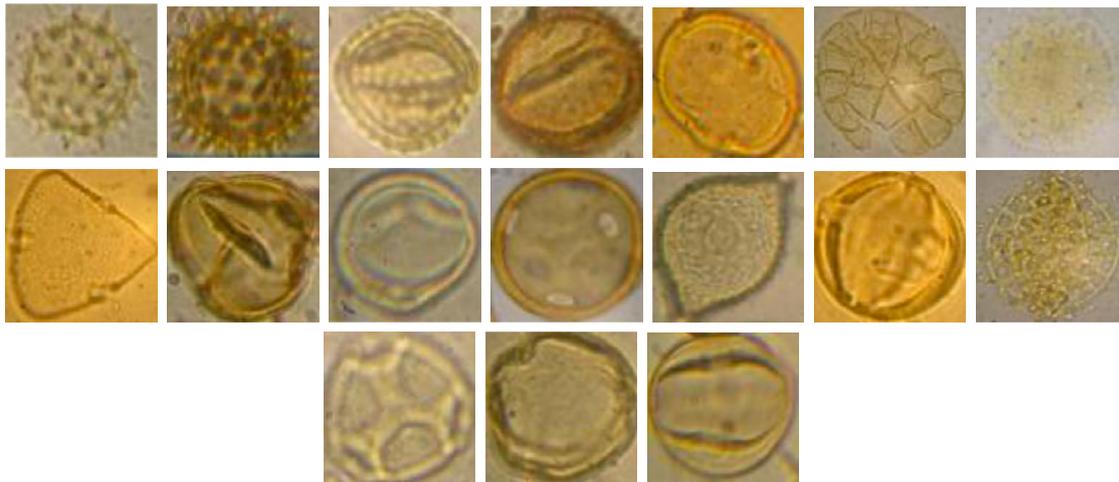


Figura 4 – Exemplo de uma amostra de cada espécie de pólen investigadas em (POZO-BAÑOS et al., 2015)

Focando o aprendizado automático de características, (DAOOD; RIBEIRO; BUSH, 2016) exploraram uma transferência de aprendizado utilizando uma CNN de 20 milhões de parâmetros, comparando seu desempenho a métodos de extração de características pré-projetadas. Para seus experimentos foram construídas duas bases de imagens de pólen contendo 30 espécies através de duas modalidades: utilizando um microscópio ótico e um microscópio eletrônico. As amostras de cada espécie utilizando um microscópio eletrônico estão ilustradas na Figura 5. A CNN proposta obteve um melhor resultado através das imagens do microscópio eletrônico, atingindo 93,99% de acurácia. Os autores relatam que apesar da metodologia atingir acurácia promissora, o treinamento de CNNs leva tempo considerável especialmente em computadores pessoais.

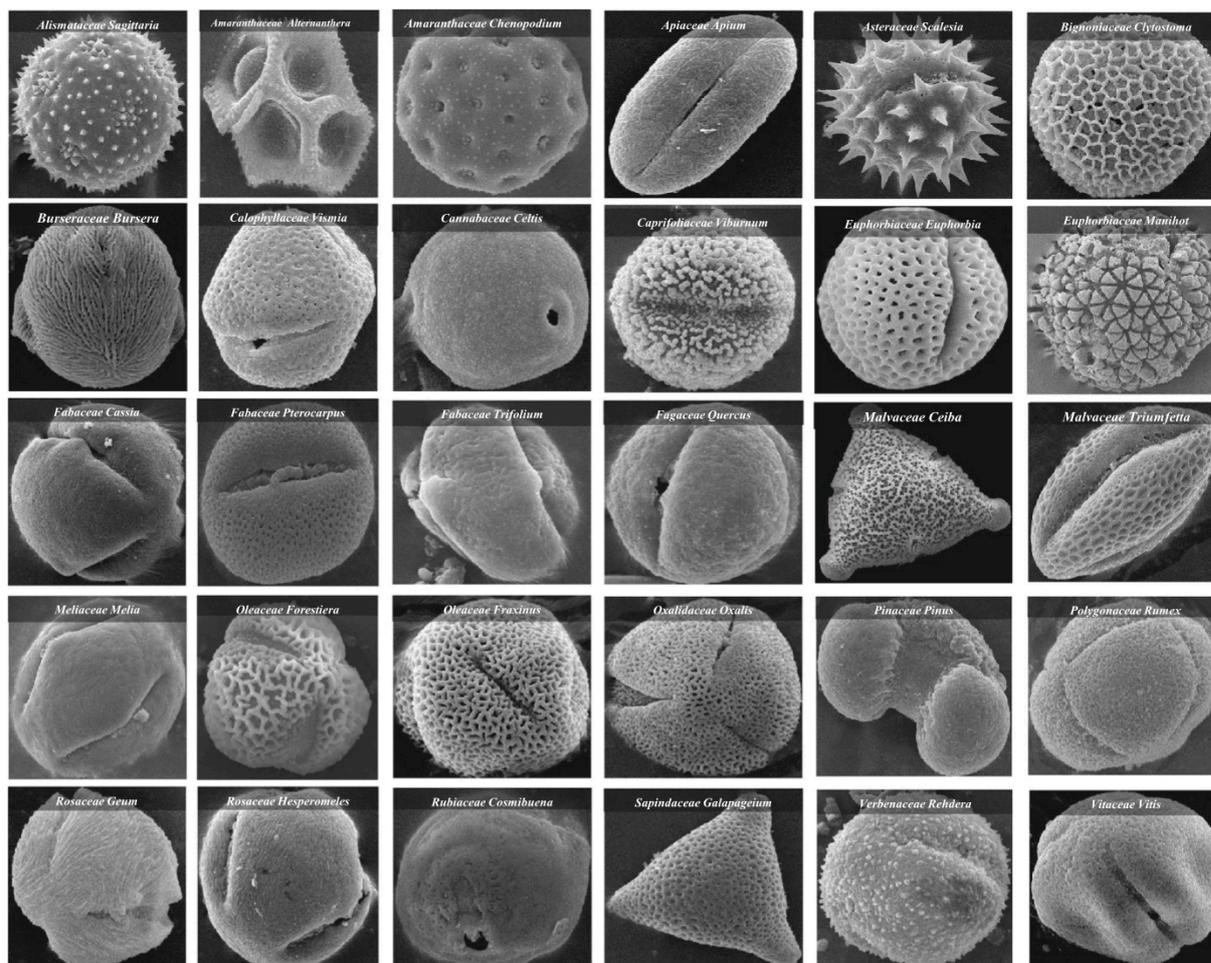


Figura 5 – Exemplo de uma amostra de cada espécie de pólen investigadas em (DAOOD; RIBEIRO; BUSH, 2016).

O estudo de (GONCALVES et al., 2016) introduziu um conjunto de imagens de pólen de 23 classes coletadas em amostras de mel no estado do Mato Grosso (veja Figura 6). As classes coletadas são importantes para apicultores e têm sido investigadas na literatura. A metodologia compara a classificação computacional dos grãos com o reconhecimento visual realizado por especialistas. Para classificação computacional são analisadas características de cor, formato, textura e suas combinações. Resultados experimentais foram capazes de atingir 66% de acurácia média utilizando uma variação do classificador SVM, contra média de 67% pelos especialistas.

Apesar das abordagens terem desempenho comparável em termos de acurácia, destaca-se que o tempo computacional é muito inferior quando comparado ao reconhecimento visual. A saber, para análise de 46 amostras os apicultores levaram em média 2 horas, enquanto os métodos computacionais levaram em média 10 segundos. Os autores sugerem que a baixa acurácia da metodologia quando comparada a outros estudos se dá pelo número elevado de classes introduzido. Isso indica que quanto maior o número de classes,

maior o desafio em classificação de pólen.

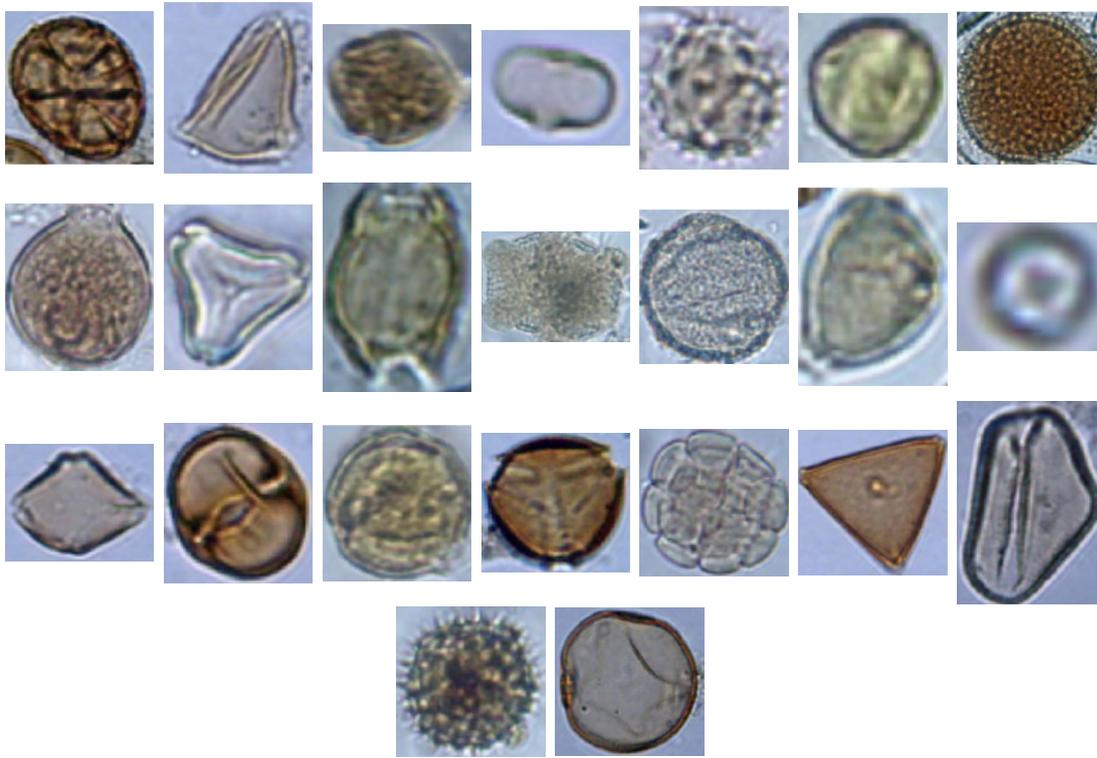


Figura 6 – Exemplo de uma amostra de cada espécie de pólen investigadas em (GONCALVES et al., 2016)

(SEVILLANO; AZNARTE, 2018) apresentaram três modalidades de transferência de aprendizado avaliando o conjunto de imagens introduzido em (GONCALVES et al., 2016). Nos experimentos foi utilizada a CNN AlexNet pré-treinada no conjunto de imagens ImageNet. Na primeira modalidade utilizou-se a CNN como extrator de características, classificando através do método LDA as características da camada imediatamente anterior às camadas de classificação. Na segunda, a camada de classificação foi alterada para se adaptar ao número de espécies do conjunto de pólen, treinando esta camada com uma alta taxa de aprendizado e nas demais camadas anteriores uma baixa taxa de aprendizado. Na terceira, uma abordagem híbrida é proposta, realizando treinamento da CNN como na segunda modalidade e classificando os dados com LDA como na primeira modalidade.

Resultados experimentais apontam que a terceira modalidade atinge melhores resultados, tendo acurácia de 97,22%. Entretanto, os autores sugerem que todas modalidades podem ser aplicadas, já que não tem diferenças significativas nos resultados. A saber, a primeira modalidade atingiu acurácia de 96,62%, enquanto a segunda modalidade 96,15%. Esse resultado sugere que características genéricas aprendidas em um conjunto de dados grande como o ImageNet podem ser utilizadas na classificação de pólen.

## 2.2 Classificação de vírus

A identificação de tipos de vírus é um problema essencial em medicina. Ela possibilita aos pesquisadores estudarem os atributos dos vírus, como sua morfologia, composição química e modos de replicação. Devido à sua importância, e o fato desta tarefa ser altamente dependente de um especialista qualificado, tem recebido alta atenção nos anos recentes. Através destes estudos, é possível determinar como o vírus interage com nossas células e descobrir medicamentos para bloqueá-los.

A microscopia convencional óptica não é capaz de visualizar diretamente as partículas virais, já que os vírus têm tamanho muito inferior ao de células. Entretanto, a partir da criação dos microscópios eletrônicos de transmissão, foi possibilitado que a comunidade de visão computacional aplicasse métodos para a classificação automática de vírus.

(KYLBERG; UPPSTRÖM; SINTORN, 2011) introduziu a base de imagens utilizada nos experimentos deste trabalho (seção 3.3.4), sendo em nosso conhecimento a maior base considerando a quantidade de tipos de vírus. Em seu estudo, a melhor configuração de características pré-projetadas geradas através do método LBP e LDP (do inglês, *Local Directional Patterns*) (JABID; KABIR; CHAE, 2010) foram aplicadas ao classificador RF e obtiveram 87% de acurácia nas 15 espécies de vírus.

Outros trabalhos também utilizaram esta base durante os últimos anos. (WEN et al., 2016) combinaram filtros PCA e CLBP (do inglês, *Completed Local Binary Patterns*), e utilizando SVM com kernel polinomial atingiram 86,2% de acurácia. (BACKES; JUNIOR, 2020) avaliaram o desempenho individual e a combinação de 10 métodos de características pré-projetadas realizando sua classificação através de SVM. O melhor resultado individual atingiu 73,53% de acurácia através do método RNN (do inglês, *Randomized Neural Network*) (JUNIOR; BACKES, 2016), enquanto a combinação de todos os métodos atingiu 87,27% de acurácia.

(WEN et al., 2018) propuseram a aplicação de filtros múltiplos de PCA seguidos de filtros Gaussianos de várias escalas para representar as informações em diferentes perspectivas. Em seus experimentos, foram avaliados 4 métodos de características pré-projetadas utilizando classificador SVM, os quais produziram acurácias entre 82% e 88%.

Outro desafio neste tipo de aplicação é a separação e a detecção dos objetos de interesse em imagens digitais de lâminas microscópicas, que assim como na classificação de pólen, apresenta outros tecidos em meio aos vírus. (ITO et al., 2018) propuseram uma estrutura de CNN para a detecção de partículas de vírus. Para tanto, construíram uma base de imagens composta de 35 imagens obtidas através de microscópio de transmissão, contendo em cada uma de 1 a 28 partículas de CRFK (do inglês, *Crandell-Rees Feline Kidney*). As imagens foram segmentadas manualmente por um especialista para realização do treinamento, indicando precisamente as partículas em cada imagem, como ilustra a Figura 7. A CNN proposta atingiu 99,8% F-score, superando entre 1,6% e 22,6% os estudos anteriores propostos.

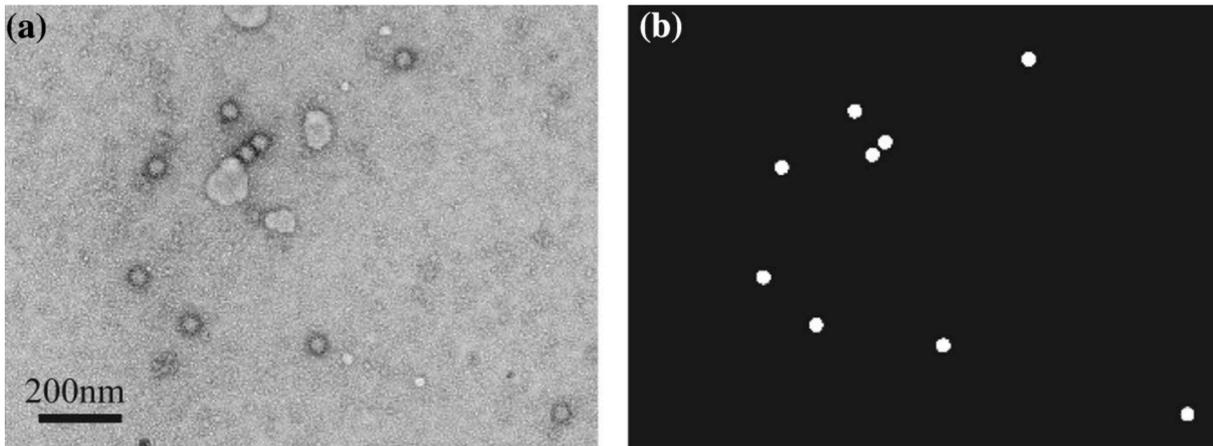


Figura 7 – (a) Exemplo de imagem microscópica analisada e (b) segmentação manual da imagem realizada por especialista (*ground truth*) (ITO et al., 2018).

De maneira semelhante, o estudo de (KYLBERG et al., 2011) desenvolveu um novo método para a segmentação de 22 tipos de vírus. A inovação proposta é baseada na análise estatística da vizinhança local seguida de uma discriminação pela largura dos objetos. Diante disso, para cada tipo de vírus é necessário informar como entrada do método o tamanho médio em nanômetros da partícula a ser analisada. Os resultados experimentais apontam 77% de sensibilidade, onde o maior valor (95%) foi atingido para Norovírus, Orf and Sapovírus e o pior (46%) para CCHF (do inglês, *Crimean-Congo haemorrhagic fever*). O pior resultado encontrado no tipo CCHF é justificado pela sua complexa estrutura que apresenta bordas irregulares (Figura 8).

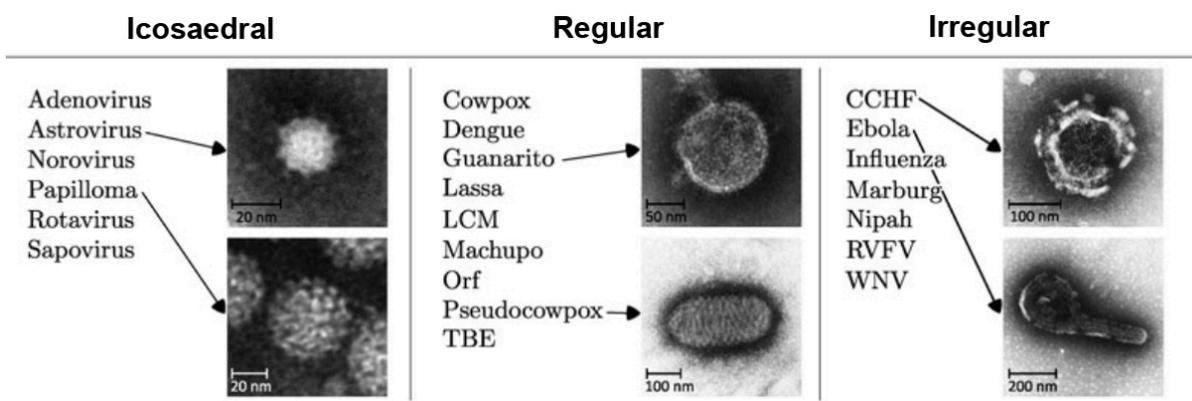


Figura 8 – Tipos de vírus agrupados nos formatos icosaedral, regular e irregular. (KYLBERG et al., 2011)

## 2.3 Classificação de madeira

O estado da arte na classificação de espécies de madeira é vasto, mas nesta seção, levantamos os trabalhos computacionais com uma análise de cortes no tronco usadas para a classificação de espécies de madeira. Dessa forma, os estudos visando a automação da classificação de espécies de madeira começaram na década de 1990, sendo as abordagens desenvolvidas categorizadas em quatro aspectos: 1) tipo de dados (imagem ou sinais), 2) quantidade de espécies, 3) características da imagem e 4) método de classificação.

Os dados de madeira mais amplamente utilizados são imagens, entretanto sinais ultrassônicos (JORDAN et al., 1998), curvas de refletância espectral (ZHAO; CAO, 2016) e ondas de espectroscopia no infravermelho (NISGOSKI; OLIVEIRA; NIZ, 2017) também foram utilizados. Quanto ao uso de imagens, a maioria das investigações utilizou imagens fotográficas; no entanto, estudos promissores têm usado imagens microscópicas. Entre os estudos utilizando as imagens microscópicas destaca-se o estudo de Hafemann et al. (HAFEMANN; OLIVEIRA; CAVALIN, 2014) que atingiram 97,32% de acurácia em um conjunto de imagens contendo 112 espécies.

Dentre os estudos que usaram imagens capturadas por simples câmeras fotográficas estão as pesquisas de (IBRAHIM et al., 2016; HU; LI; BAO, 2015; ZAMRI et al., 2016; SUNDARAM et al., 2015). (IBRAHIM et al., 2016) apresentaram um sistema de classificação de textura de madeira para 48 espécies de madeiras tropicais. Visando problemas de eficiência computacional, o sistema consiste em um classificador de dois níveis. Primeiro, um pré-classificador difuso é usado para classificar as imagens em quatro grandes categorias com base no tamanho dos poros. Em seguida, uma SVM é usada para realizar a classificação final da imagem com base nas características de textura da AM (do inglês, *Aura Matrix*) e estatísticas de distribuição de poros. A acurácia obtida foi de 98,5% para o conjunto de imagens com 48 espécies de madeira. (ZAMRI et al., 2016) propuseram um novo extrator de características de textura de imagem de madeira baseado na AM dos níveis de cinza. Segundo os autores, o extrator proposto é invariante à rotação e gera várias características menores que o método original, sendo consideradas apenas características discriminativas. O descritor foi utilizado para classificar 52 espécies de madeira utilizando uma MLP e atingiu 97,01% de acurácia sendo muito superior às matrizes de co-ocorrência e AM.

(SILVA et al., 2017) usa imagens microscópicas de seções transversais de madeira, obtidas de forma padronizada, com zoom de 2,5x. A caracterização das imagens é feita utilizando dois extratores de características de textura, a saber: LBP e LQP (do inglês, *Local Quinary Patterns*). Antes da classificação pelo método kNN, os autores aplicam a redução de dimensionalidade via LDA. O banco de dados usado nos testes contém 77 espécies de madeira africana; as amostras também foram categorizadas em gênero e família. A precisão da classificação obtida é de aproximadamente 88% no nível da espécie, 89% no nível do gênero e 90% no nível da família.

Focando o aprendizado automático, (HAFEMANN; OLIVEIRA; CAVALIN, 2014) investigaram o uso de *Deep Learning*, em particular, CNN para classificação de texturas em duas bases de imagens de espécies florestais brasileiras da literatura: uma base de imagens macroscópicas e outra com imagens microscópicas. A arquitetura proposta tem dois blocos consecutivos, compostos de convolução (64 filtros de  $5 \times 5$ ) e *pooling* ( $3 \times 3$  deslocamento 2), seguidos de duas camadas localmente conectadas (32 filtros de  $3 \times 3$ ) e uma camada totalmente conectada *softmax* para classificação. Na base de dados macroscópica, com 41 classes, obteve-se 95,77% de acurácia, enquanto para a base de dados microscópica com 112 classes, obteve-se uma acurácia de 97,32% (4,1% superior ao estado da arte). Em (KWON et al., 2017) foi desenvolvido um sistema automático para identificação de espécies de madeira usando modelos CNN como LeNet, MiniVGGNet e suas variantes. Resultados utilizando uma base de imagens de seção transversal macroscópica de cinco espécies de madeira de resinosas coreanas usando LeNet3 obtiveram acurácia de 99,3%.

As Tabelas 1 e 2 sumarizam os estudos para a classificação de madeira, focando no tipo de imagem usada, número de espécies, técnica de classificação e acurácia. Observa-se que a grande maioria dos estudos utiliza características pré-projetadas em seus experimentos, exceto dois trabalhos que usam aprendizado automático através de CNNs. Além disso, a quantidade de espécies entre os trabalhos varia bastante, sendo 112 o maior número investigado.

## 2.4 Considerações finais

Este capítulo levantou os principais trabalhos que abordam a identificação computacional de madeira, pólen e vírus. Destaca-se que um dos desafios destas aplicações é o número limitado de classes nos experimentos levantados, que possuem larga margem para expansão considerando o número real existente; também pela falta de padronização na aquisição das imagens, que leva em conta fatores como: equipamento de captura utilizado, metodologia de preparação do material, ambiente e iluminação. Além disso, as grandes variações das metodologias computacionais envolvem a extração de características pré-projetadas e o aprendizado automático de características, que embora apresentem resultados promissores nos conjuntos avaliados, podem apresentar desempenho diferente em conjuntos maiores e que não sigam a mesma padronização dos dados.

Tabela 1 – Sumarização de estudos para classificação de madeira.

Referência	Tipo de imagem	Espécies de madeira	Extractor de características	Classificador	Acurácia
Ibrahim al. (IBRAHIM et al., 2016)	Imagens fotográficas do corte transversal; 10x de zoom	48 espécies de madeira da Malasia	Tamanho, quantidade, distribuição dos poros	Lógica Fuzzy e SVM	98.5%
Ibrahim al. (IBRAHIM et al., 2017)	Imagens fotográficas do corte transversal; 10x de zoom	30 espécies de madeira da Malásia	24 características estatísticas	Lógica Fuzzy e MLP	89.3%
Hu et al. (HU; LI; BAO, 2015)	Imagens fotográficas do corte transversal; Zoom não especificado	28 espécies de madeira; Local não especificado	Bag-of-Words baseado em SIFT	MLP	90.2%
Khalid, Yusof and Khairuddin (KHALID; YUSOF; KHAIRUDIN, 2011)	Imagens fotográficas do corte transversal; 10x de zoom	52 espécies de madeira da Malasia	AM e estatísticas da distribuição dos poros	Classificador multi-level baseado em k-means, redução de dimensionalidade e kNN	96.92%
Zamri et al., (ZAMRI et al., 2016)	Imagens fotográficas do corte transversal; 10x de zoom	52 espécies de madeira da Malasia	Novo extrator de características baseado em AM	MLP	97.01%
Sundaram et al. (SUNDARAM et al., 2015)	Prospect Wood Database: imagens de 256 × 256 extraídas das amostras	10 espécies de madeira; Local não especificado	Características estatísticas e índices de saliência	MLP	90%
Nisgosi et al. (NISGOSKI; OLIVEIRA; NIZ, 2017)	Ondas de espectroscopia no infravermelho-próximo adquiridas com espectrômetro Bruker Tensor 37	4 espécies de madeira do Brasil	1500 números de onda de espectro	MLP	90%
Zhao and Cao (ZHAO; CAO, 2016)	Curvas de razão de reflectância espectral obtidas com espectrômetro portátil	5 espécies de madeira da China	Autovalores selecionados da matriz de dispersão	Função de distância	95%
Silva et al. (SILVA et al., 2017)	Imagens microscópicas do corte transversal; 2.5x de zoom	77 espécies de madeira da Africa	LPQ com redução de dimensionalidade através de LDA	kNN	88%

Tabela 2 – Sumarização de estudos para classificação de madeira.

Referência	Tipo de imagem	Espécies de madeira	Extrator de características	Classificador	Acurácia
Jordan et al. (JORDAN et al., 1998)	Sinais de ondas ultrassônicas	4 espécies de madeira; Zoom não especificado	Velocidade do pulso ultrassônico e amplitude do sinal	MLP	97.5%
Barmpoutis et al. (BARMPOUTIS et al., 2018)	Imagens fotográficas usando uma câmera Nikon D3300, 24 megapixels a 15-20cm de distância: três tipos de cortes: transversal (melhor resultado), radial and tangencial. Imagens foram cortadas em $400 \times 400$ pixels	12 espécies de madeira da Grécia	Bag-of-visual-words de características espaciais extraídas de cortes horizontais e verticais	SVM	91.47%
Hafemann et al. (HAFEMANN; OLIVEIRA; CAVALINI, 2014)	Imagens microscópicas; $100 \times$ de zoom	112 espécies de madeira do Brasil	-	Arquitetura de CNN treinada com cortes na imagem de tamanho $224 \times 224$ pixels	97.32%
Kwon et al. (KWON et al., 2017)	Imagens fotográficas do corte transversal usando iPhone 7	5 espécies de madeira da Korea	-	Arquitetura CNN LeNet3 (melhor resultado), MiniVGNet e suas variantes	99.3%
Hu et al. (HU; LI; BAO, 2015)	Imagens fotográficas	28 espécies de madeira. Local não especificado	Histograma de descritores SIFT	MLP (melhor resultado), SVM e kNN	90.2%
Peng (PENG, 2013)	Imagens fotográficas do corte transversal	5 espécies de madeira. Local não especificado	Descritor de variação de cor intrínseco e interesse específico baseado em <i>snake model</i>	Matching measure regarding resemblance and deformation	87.6 %
Tou et al. (TOU; TAY; LAU, 2009)	Imagens fotográficas do corte transversal	6 espécies de madeira da Malasia	Filtros Gabor (melhor resultado) e GLCM	kNN	85%
Yadav et al. (YADAV et al., 2013)	Imagens microscópicas; $100 \times$ de zoom	25 espécies de madeira do Brasil	GLCM extraídas de imagens aprimoradas por Filtros Gabor	MLP	92.5%

---

## Materiais e Métodos

Este capítulo apresenta uma breve introdução da teoria de CNNs e as arquiteturas recentes avaliadas neste trabalho de pesquisa, assim como métodos de extração de características pré-projetados utilizados para comparação. Também são apresentadas as bases de imagens usadas nos experimentos desta tese.

### 3.1 Redes Neurais Convolucionais

A Rede Neural Convolutiva, também conhecida como CNN, é um modelo de aprendizagem profunda comumente utilizado no processo de aprendizagem de características para classificação e segmentação de imagens (LUO et al., 2016; LONG; SHELHAMER; DARRELL, 2015). Basicamente, a CNN é uma rede bioinspirada que permite ao computador aprender classificações corretas a partir dos dados. Essa rede utiliza o conceito de campos receptivos para explorar correlações espaciais dentro da imagem, transformar e reduzir as informações contidas na imagem, obtendo assim uma representação mais significativa do seu conteúdo.

Sua composição básica consiste de três camadas: camada convolutiva, camada de subamostragem (*pooling*) e camada totalmente conectada. A camada convolutiva, conhecida como a parte principal das CNNs, usa a operação de convolução para emular um campo receptivo e sua resposta a um estímulo visual. Dada uma imagem bidimensional  $x$  como entrada, considere sua decomposição em uma entrada sequencial  $x = \{x_1, x_2, \dots, x_N\}$ . Para compartilhar o peso, a camada convolutiva é definida como:

$$y_j = f\left(\sum_i K_{ij} \otimes x_i + b_j\right) \quad (1)$$

onde  $y_j$  denota a  $j$ -ésima saída para a camada convolutiva e  $K_{ij}$  denota o kernel convolutivo com o  $i$ -ésimo mapa de entrada  $x_i$ .  $\otimes$  denota o operador de convolução discreta e  $b_j$  denota o *bias*. Além disso,  $f$  denota uma função de ativação tipicamente não linear, já que permite o aprendizado de características mais complexas dos dados (GOODFELLOW; BENGIO; COURVILLE, 2016). Dentre as funções mais utilizadas nos estudos

estão a função sigmóide e função tangente hiperbólica. Entretanto, a função ReLU (do inglês, *Rectified Linear Unit*) introduzida em (GLOT; BORDES; BENGIO, 2011), tem se mostrado mais eficaz em modelos de CNN profundas pela sua capacidade de lidar com o problema do desaparecimento do gradiente e seu menor custo computacional.

A camada de subamostragem visa reduzir a dimensão do mapa de características, assim como tornar a rede menos sensível às distorções e deslocamentos da imagem. Normalmente, ela pode ser implementada por uma operação de média ou máximo em uma janela de valores. Os parâmetros considerados são o tamanho da janela (ou filtro) e deslizamento utilizado. A forma mais comum utilizada é um filtro  $2 \times 2$  com deslizamento 2, que reduz a largura e altura do mapa de características pela metade, descartando 75% das ativações.

Em seguida, camadas totalmente conectadas, e uma camada *softmax* são normalmente colocadas nas camadas finais para classificação dos mapas de características obtidos pelas camadas anteriores. A camada totalmente conectada é uma estrutura de conexão entre os neurônios, que prevê que cada neurônio de saída tem ligação a todos os neurônios da camada anterior por um peso exclusivo entre eles, e um *bias* compartilhado. Já a camada *softmax* (KRIZHEVSKY; SUTSKEVER; HINTON, 2012) é apenas um caso especial de camada totalmente conectada onde o número de neurônios é igual ao número de classes, tendo como saída a probabilidade normalizada para cada classe.

Uma arquitetura CNN voltada à classificação inclui camadas convolucionais e camadas de subamostragem para o aprendizado de características em imagens, exigindo tempo computacional elevado para o treinamento. Nos últimos anos várias arquiteturas foram propostas, redefinindo o estado da arte em problemas de classificação (HOWARD et al., 2017; SZEGEDY et al., 2016). Seu grande sucesso foi possibilitado principalmente pelo processamento realizado em GPU, reduzindo consideravelmente o seu tempo de treinamento.

Devido ao elevado número de camadas, e conseqüentemente pesos na rede, é fundamental um número elevado de imagens de treinamento para convergência da rede sem causar super ajuste, fenômeno chamado de *overfitting*. Esse problema já foi investigado em diversas pesquisas, motivando o uso de técnicas como o *dropout* (SRIVASTAVA et al., 2014), que consiste em desativar alguns neurônios aleatoriamente durante o treinamento; transferência de aprendizado, que utiliza um grande conjunto de imagens de treinamento para inicialização dos pesos; e *data augmentation*, que realiza operações nas imagens, como rotação, cortes, mudança na iluminação, para gerar um conjunto de imagens maior e mais abrangente.

### 3.1.1 Transferência de Aprendizado

Em diversos problemas, são raros conjuntos de imagens de tamanho suficiente para treinar uma CNN a partir da inicialização dos pesos aleatória. Por isso, é comum pré-

treinar essas arquiteturas utilizando um grande conjunto de imagens, visando aproveitar características genéricas aprendidas nas camadas iniciais da rede (YOSINSKI et al., 2014). Um conjunto frequentemente utilizado nessa abordagem é o ImageNet 2012 (RUSSAKOVSKY et al., 2015), que foi idealizado para possibilitar a pesquisadores fácil acesso à uma base de imagens em larga-escala, disponibilizando 1,2 milhões de imagens agrupadas em 1000 classes. Como as modernas CNNs levam de 2 a 3 semanas para treinar em várias GPUs no ImageNet, é comum o compartilhamento de pesos treinados em CNNs para o benefício de outras pessoas. A partir do modelo pré-treinado existem dois principais cenários em transferência de aprendizado:

**CNN como extrator de característica** (DONAHUE et al., 2014): O processo de extração de características consiste em remover da arquitetura pré-treinada a última camada totalmente conectada utilizada para classificação e utilizar a saída dos neurônios como características. A partir dessas características um novo classificador pode ser treinado ou pode-se optar por treinar uma nova camada *softmax* para o novo conjunto de classes.

**Ajuste-fino** (GIRSHICK et al., 2014): A estratégia visa não apenas substituir e treinar uma nova camada *softmax* para o novo conjunto de imagens, mas também ajustar os pesos da rede pré-projetada. É possível fazer o ajuste fino de todas as camadas da CNN, ou manter algumas das camadas anteriores fixas (devido a problemas de *overfitting*) e apenas ajustar camadas de nível superior da rede. Isso é motivado pelo fato de que os filtros das primeiras camadas de uma CNN estimam características mais genéricas (e.g. detectores de borda) que podem ser úteis para diferentes tarefas. Já as camadas posteriores da CNN se tornam mais específicas para os detalhes das classes contidas nos dados.

### 3.1.2 Arquiteturas de CNN

Existem várias arquiteturas de CNN propostas na literatura que apresentam bons resultados de classificação em diversos conjuntos de imagens. Dentre as de maior destaque estão:

**ResNet:** É um dos modelos de aprendizado residual profundo para reconhecimento de imagens (HE et al., 2016) proposto pela equipe de pesquisa da Microsoft. Os autores criaram uma ideia simples e elegante. Por meio de uma CNN profunda padrão, foram adicionadas conexões de atalho (veja Figura 9) que contornam poucas camadas convolucionais de cada vez. As conexões de atalho criam blocos residuais, onde as saídas das camadas de convolução são adicionadas ao vetor de entrada do bloco. Por exemplo, o modelo ResNet50 é composto por 50 camadas de blocos semelhantes com conexões de atalho. Essas conexões mantêm o cálculo baixo e, ao mesmo tempo, fornecem recursos avançados de combinação.

O modelo ResNet50 tem uma camada convolucional seguida por uma camada de normalização, e possui duas camadas de agrupamento entre as quais há um total de 16 módulos residuais. Dois tipos de módulos são alternados, um que tem 4 camadas convolucionais e outro com 3 camadas. Cada camada convolucional é seguida por uma etapa de normalização. O bloco residual com 4 camadas convolucionais é o primeiro usado, seguido de dois ou mais blocos residuais com 3 camadas convolucionais e assim por diante.

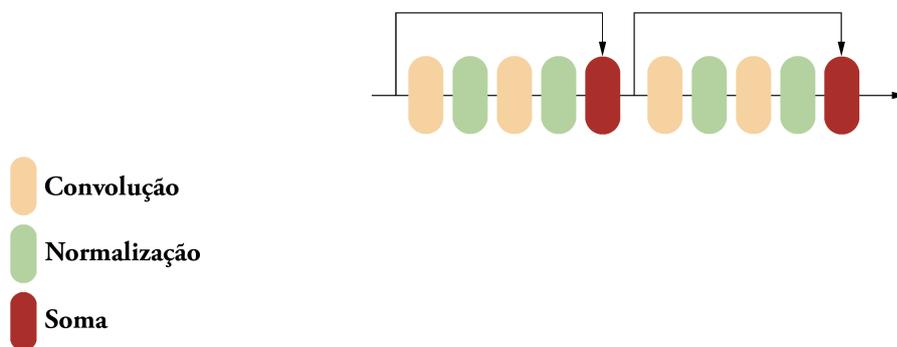


Figura 9 – Modelo de arquitetura ResNet.

**Inception:** é um modelo produzido pela equipe de pesquisa do Google, concentrando-se principalmente na redução da carga computacional das CNNs, e buscando manter o mesmo nível de desempenho. A inovação surgiu do novo módulo chamado *inception* que, na maioria das vezes, pode ser descrito como 4 caminhos paralelos de filtros de convolução  $1 \times 1$ ,  $3 \times 3$  e  $5 \times 5$ , como ilustra a Figura 10. Por causa de sua implementação paralela, além das camadas de subamostragem em cada bloco, o tempo gasto para execução do modelo é maior que o da ResNet.

A equipe de pesquisa propôs muitos modelos ao longo dos anos que são cada vez mais complexos. A CNN InceptionV3 (SZEGEDY et al., 2016) foi introduzida próxima a ResNet e construída com alguns novos princípios de design, por exemplo, a utilização de convoluções  $3 \times 3$  em vez de  $5 \times 5$  ou  $7 \times 7$  nos módulos de inception, também a expansão da largura em cada camada para aumentar a combinação de recursos para a próxima camada. Essas decisões foram tomadas com o objetivo de construir uma rede neural com um desempenho computacional balanceado entre sua profundidade e largura.

**SqueezeNet:** Trata-se de uma arquitetura (IANDOLA et al., 2016) projetada para ter menor tamanho, mas atingindo resultados comparáveis a arquiteturas de maior tamanho. Os módulos *Fire* introduzidos pela arquitetura, conforme Figura 11, possuem

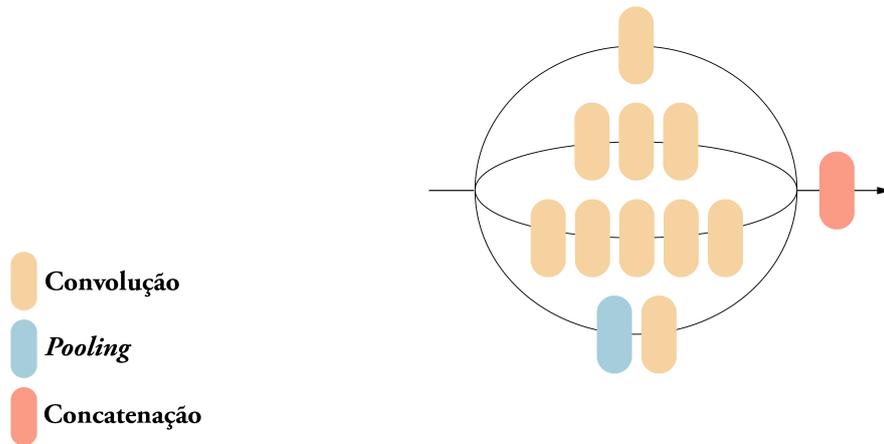


Figura 10 – Módulo *inception* de arquitetura InceptionV3.

camadas de convolução de compressão (que possuem apenas filtros de  $1 \times 1$ ), alimentando uma camada de expansão que possuem uma mistura de filtros de convolução de  $1 \times 1$  e  $3 \times 3$ . Ela tem  $50\times$  menos parâmetros que a AlexNet, oferecendo pelo menos três vantagens principais: menor comunicação entre servidores durante o treinamento distribuído, menor largura de banda para exportar um novo modelo da nuvem para um carro autônomo e maior viabilidade para implantar em FPGAs e outros hardwares com memória limitada.

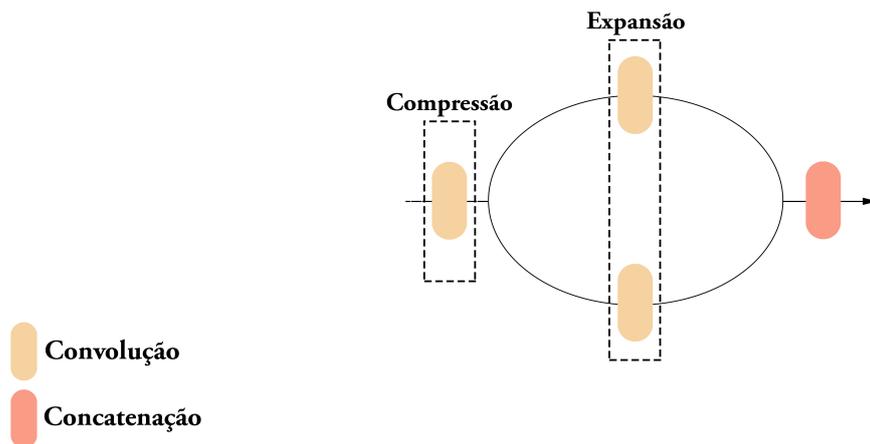


Figura 11 – Módulo *Fire* da arquitetura SqueezeNet.

**DenseNet:** É uma extensão lógica da ResNet, tendo como bloco de construção fundamental o conceito de conexões de resíduos. Em contraste com a ResNet, a DenseNet propõe concatenar as camadas anteriores em vez de usar um somatório. A DenseNet (HU-

ANG et al., 2017) tem seu módulo Dense Block ilustrado na Figura 12, que conecta cada camada a todas as outras camadas em um modo *feed-forward*. Enquanto as CNNs tradicionais com  $L$  camadas têm grupos de conexões  $L$  (um grupo entre cada camada e sua camada subsequente), a DenseNet tem  $L(L + 1)/2$  grupos de conexões. Para cada camada, os mapas de características das camadas predecessoras são usados como entrada e seu mapa de saída é usado como entrada para as camadas subseqüentes. DenseNet considera-se uma pequena CNN com 7 milhões de parâmetros.

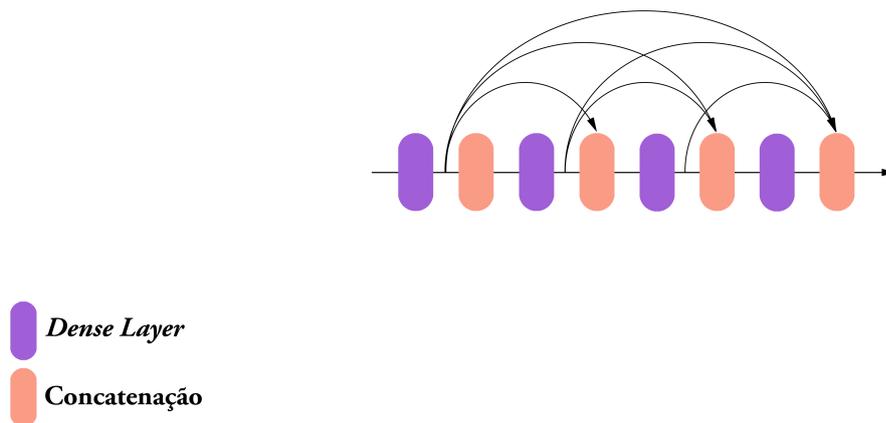


Figura 12 – Módulo *Dense Block* de arquitetura DenseNet.

## 3.2 Extratores de características pré-projetados

Nesta seção são apresentados os métodos de extração de características utilizados para reproduzir trabalhos recentes e utilizados para propor metodologias a serem comparadas com o aprendizado automático de características por CNN.

### 3.2.1 LBP

Inicialmente proposto por (OJALA; PIETIKÄINEN; HARWOOD, 1996), o LBP (do inglês, *Local Binary Patterns*) tem sido empregado em diversos estudos na literatura que demonstram seu potencial para representação e classificação de imagens de textura em escala de cinza. Sua abordagem estatística é capaz de descrever mudanças de pequenas escalas presentes em diferentes texturas, tais como áreas planas, cantos, bordas, etc. Além disso, o LBP oferece invariância contra mudanças monotônicas do nível de cinza em imagens.

De forma resumida, o LBP calcula as diferenças das intensidades de um pixel em uma certa região. Considere  $g_c$  sendo a intensidade associada a um pixel central e  $g_p$ ,

$p = 1, \dots, P$  a intensidade em cada um dos seus  $P$  vizinhos. O LBP produz um código binário baseado nas variações de intensidades e posições da vizinhança de pixels em relação ao seu pixel central. O código binário é calculado da seguinte forma:

$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c)2^p \quad (2)$$

$$s(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (3)$$

onde  $R$  é o tamanho do raio da vizinhança.

A literatura tem investigado diversas maneiras para definir a vizinhança utilizada no cálculo do código binário que impactam diretamente no desempenho do método. Alguns autores utilizam uma vizinhança circular, enquanto outros preferem uma máscara  $3 \times 3$ . Independentemente da abordagem, os pixels vizinhos são sempre visitados no sentido horário. Um valor de 1 é atribuído a esse pixel se sua intensidade for maior que a do pixel central; caso contrário, um valor 0 é atribuído. Isso resulta em um código binário que é convertido em um valor decimal associado ao pixel central. Esse processo é calculado para todos os pixels da imagem, resultando no mapa LBP da imagem. Tradicionalmente, um histograma é obtido deste mapa e usado para representar o padrão de imagem original.

### 3.2.2 LPQ

O método LPQ (do inglês, *Local Phase Quantization*), proposto em (OJANSIVU; HEIKKILÄ, 2008), é semelhante ao LBP, que calcula uma descrição para cada pixel considerando o conjunto de vizinhos em um círculo de raio  $R$ . Entretanto, para cada pixel, é gerada uma descrição quantizando a transformada de Fourier nas vizinhanças locais. Em seguida, os histogramas da quantização local sobre uma região  $P$  são calculados e concatenados para compor o descritor.

### 3.2.3 RI-LPQ

O método RI-LPQ (do inglês, *Rotation Invariant Local Phase Quantization*), é uma melhoria proposta em (RAHTU et al., 2012), que basicamente torna o método LPQ invariante à rotação. O processo tem forma semelhante ao LPQ, porém inicialmente estima-se a orientação das características. O procedimento para o cálculo das características é semelhante ao LPQ, mas a vizinhança em cada região é rotacionada de acordo com a direção da orientação da característica.

### 3.2.4 Filtro de Gabor 2D

O Filtro de Gabor 2D (DAUGMAN, 1985) é um descritor de textura vastamente utilizado em aplicações nas últimas décadas, principalmente em segmentação e classificação de imagens (COSTA, 2009). Sua popularidade vem principalmente pela sua inspiração na teoria da filtragem multicanal para o processamento de informações visuais nos estágios iniciais do sistema visual humano (DAUGMAN, 1980), além de sua seletividade para frequência e orientação (LEE; WANG, 2001).

Um filtro de Gabor 2D é uma onda senoidal modulada por um envelope gaussiano, que é dada por:

$$G_{\sigma,\mu,\theta}(x, y) = g_{\sigma}(x, y) \exp[2\pi j u(x \cos \theta + y \sin \theta)] \quad (4)$$

$$g_{\sigma}(x, y) = \frac{1}{2\pi\sigma^2} \exp[-(x^2 + y^2)/2\sigma^2] \quad (5)$$

onde  $j = (-1)^{1/2}$ . A frequência da onda senoidal é dada por  $\mu$  e sua orientação especificada como  $\theta$ .  $g_{\sigma}(x, y)$  é a função gaussiana e  $\sigma$  é o desvio padrão do envelope gaussiano. A Figura 13 ilustra um conjunto de 40 filtros de Gabor.

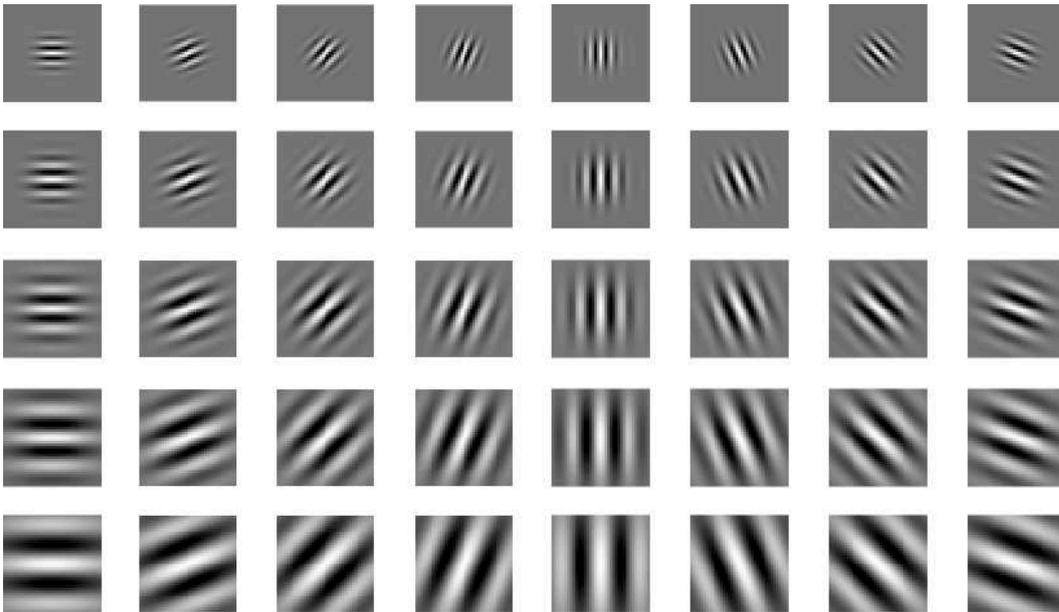


Figura 13 – Exemplo de um conjunto de 40 filtros de Gabor.

### 3.2.5 HOG

O método HOG (do inglês, *Histogram of Oriented Gradients*) (DALAL; TRIGGS, 2005) é uma ferramenta matemática de grande potencial para reconhecimento de objetos em diversos escopos (XIAO et al., 2010; DADI; PILLUTLA, 2016; LLORCA; ARROYO; SOTELO, 2013). A ideia principal por trás do método é que a forma local de um objeto

em uma imagem pode ser caracterizada pela intensidade da distribuição dos gradientes, isto é, pelas direções de seus contornos.

O procedimento de caracterização é realizado pela divisão da imagem em pequenas regiões, também chamadas células, utilizadas para calcular um histograma das direções dos gradientes. Cada pixel na célula contribui com um voto ponderado para um histograma orientado com base nos valores dos gradientes computados. Os gradientes horizontal e vertical são computados da seguinte maneira:

$$dx = I(x + 1, y) - I(x - 1, y) \quad (6)$$

$$dy = I(x, y + 1) - I(x, y - 1) \quad (7)$$

Destes gradientes, a magnitude e orientação, respectivamente, são calculadas como:

$$m(x, y) = \sqrt{dx^2 + dy^2} \quad (8)$$

$$\theta(x, y) = \arctan(dy/dx) \quad (9)$$

onde  $I(x, y)$  corresponde ao valor de intensidade do pixel na posição  $x$  e  $y$ . A fim de obter melhor invariância a iluminação, as células são organizadas em blocos, e para cada bloco é feita normalização dos valores.

## 3.3 Bases de Imagens

### 3.3.1 Base de Pólen

Em parceria com o grupo RCPol (Rede de Catálogos Polínicos Online) foi construída uma nova base de imagens contendo 134 espécies de grão de pólen. Em sua construção foram usados diferentes microscópios e reagentes, que resultam em diferentes cores de plano de fundo (veja Figura 14).

O número de imagens de lâmina por classe é de aproximadamente 25, contendo um número variável de pólen que geralmente se sobrepõem. A fim de manter a integridade de cada pólen para posterior redimensionamento, regiões de interesse quadradas foram extraídas para isolar ao máximo um grão de pólen em cada imagem, como mostra a Figura 15. Os grãos de pólen sobrepostos ou com partes oclusas não foram extraídos das imagens, entretanto algumas amostras contém partes de outros grãos (veja Figura 15b).

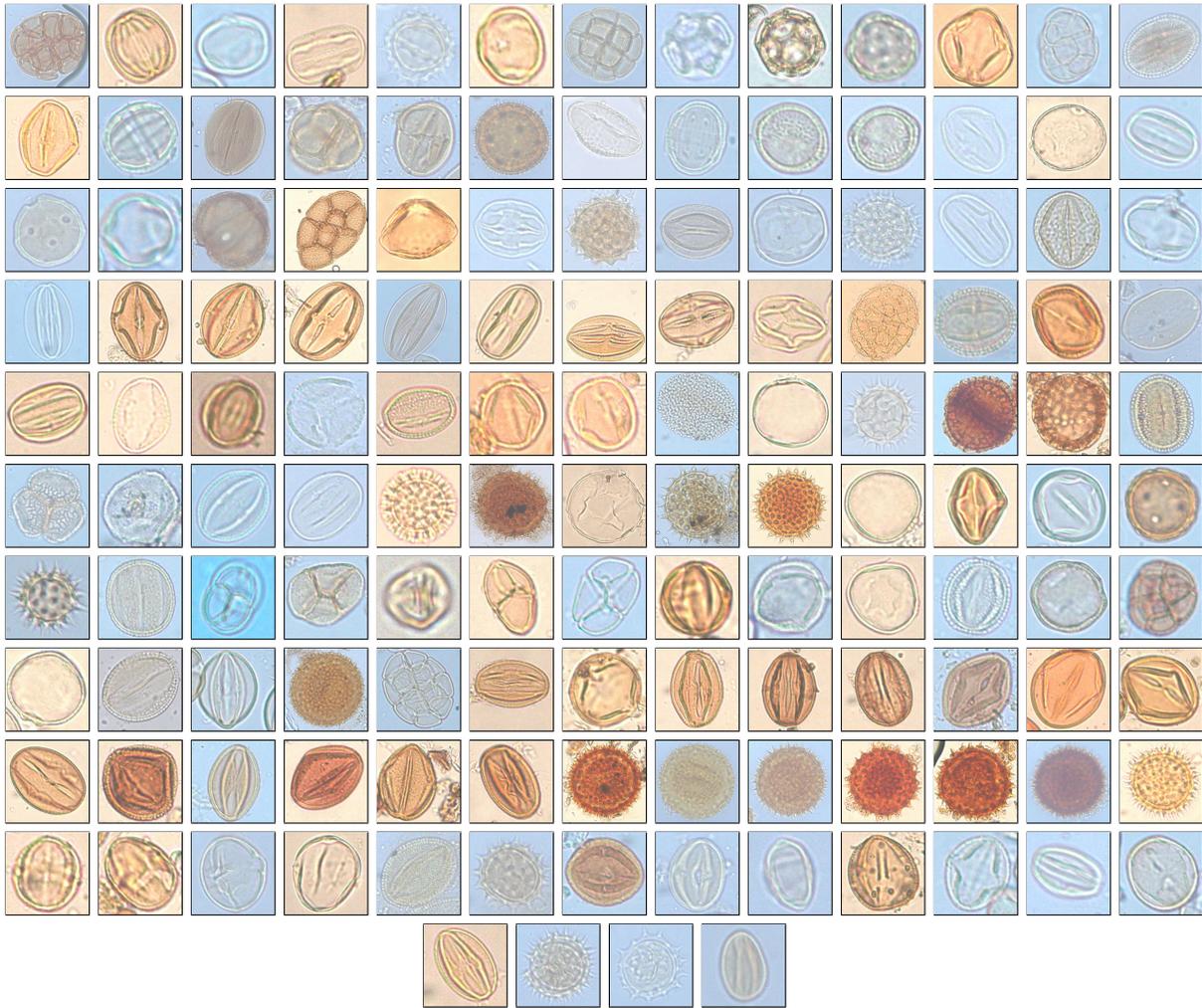


Figura 14 – Exemplos de amostras de cada tipo de pólen da base de imagens.

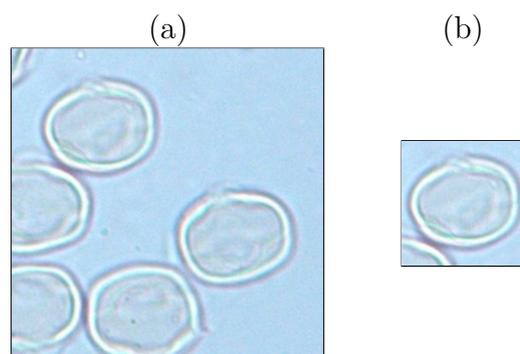


Figura 15 – Amostra de pólen: (a) Lâmina Original; (b) Região de interesse.

A Tabela 3 mostra as espécies de planta correspondente a base de imagens de pólen, o número de amostras extraídas das imagens de lâminas e o ID da espécie. A Figura 14 ilustra uma amostra de cada espécie de pólen, em que o ID é incremental percorrendo sequencialmente linha a linha do canto superior esquerdo até o canto inferior direito.

Tabela 3 – Espécie, ID e número de amostras (#) da base de imagens de pólen.

ID	Espécie	#	ID	Espécie	#
0	<i>Adenanthera pavonina</i> L.	25	67	<i>Euphorbia hyssopifolia</i>	28
1	<i>Aeschynomene americana</i>	25	68	<i>Euploca Polyphyllum</i>	22
2	<i>Aeschynomene benthamii</i>	28	69	<i>Froelichia humboldtiana</i>	31
3	<i>Aeschynomene viscidula</i>	22	70	<i>Gossypium hirsutum</i>	21
4	<i>Ageratina asclepiadea</i>	26	71	<i>Guadua trinii</i>	26
5	<i>Agonandra brasiliensis</i>	29	72	<i>Ipomoea asarifolia</i>	25
6	<i>Albizia lebeck</i>	11	73	<i>Ipomoea bahiensis</i>	25
7	<i>Alternanthera tenella</i>	22	74	<i>Ischaemum minus</i>	27
8	<i>Alternanthera brasiliensis</i> var. <i>villosa</i>	25	75	<i>Laguncularia racemosa</i>	30
9	<i>Amaranthus viridis</i>	26	76	<i>Macroptilium lathyroides</i>	27
10	<i>Amburana cearensis</i>	27	77	<i>Malpighia emarginata</i>	27
11	<i>Anadenanthera colubrina</i>	25	78	<i>Melanthera latifolia</i>	25
12	<i>Arcytophyllum nitidum</i>	26	79	<i>Mesosphaerum suaveolens</i>	26
13	<i>Azadirachta indica</i>	25	80	<i>Mimosa arenosa</i>	34
14	<i>Asemeia violacea</i>	27	81	<i>Mimosa candollei</i>	24
15	<i>Bauhinia monandra</i>	25	82	<i>Mimosa hirsutissima</i> var. <i>hirsutissima</i>	37
16	<i>Bejaria resinosa</i>	23	83	<i>Mimosa pigra</i>	30
17	<i>Bixa orellana</i>	23	84	<i>Mimosa tenuiflora</i>	26
18	<i>Boerhavia coccinea</i>	25	85	<i>Mouriri guianensis</i>	25
19	<i>Bomarea hirsuta</i>	27	86	<i>Myracrodrum urundeuva</i>	28
20	<i>Borago officinalis</i>	27	87	<i>Olyra latifolia</i>	29
21	<i>Borreria spinosa</i>	37	88	<i>Parkinsonia aculeata</i>	29
22	<i>Borreria verticillata</i>	34	89	<i>Petiveria alliacea</i>	27
23	<i>Brachyotum strigosum</i>	31	90	<i>Piptadenia stipulacea</i>	22
24	<i>Bromus catharticus</i>	26	91	<i>Piptochaetium montevidense</i>	25
25	<i>Bucquetia glutinosa</i>	47	92	<i>Poincianella bracteosa</i>	26
26	<i>Bunchosia</i> aff. <i>acuminata</i>	25	93	<i>Prosopis juliflora</i>	31
27	<i>Byrsonima sericea</i>	29	94	<i>Pseudomalachra ciliaris</i>	22
28	<i>Caesalpinia pulcherrima</i>	28	95	<i>Senegalia riparia</i>	25
29	<i>Calliandra sessilis</i>	24	96	<i>Senna acuruensis</i>	28
30	<i>Canavalia brasiliensis</i>	25	97	<i>Senna alata</i>	33
31	<i>Capsicum annuum</i>	28	98	<i>Senna cana</i> var. <i>hypoleuca</i>	25
32	<i>Carduus acanthoides</i>	31	99	<i>Senna chrysocharpa</i>	25
33	<i>Cassia fistula</i>	31	100	<i>Senna gardneri</i>	28
34	<i>Cedrela odorata</i>	28	101	<i>Senna macranthera</i>	23
35	<i>Centratherum punctatum</i>	30	102	<i>Senna occidentalis</i>	32
36	<i>Centrolobium tomentosum</i>	23	103	<i>Senna pendula</i>	30
37	<i>Centrosema virginianum</i>	24	104	<i>Senna pilifera</i>	28
38	<i>Chaetocalyx scandens</i> var. <i>pubescens</i>	25	105	<i>Senna quinquangulata</i>	28
39	<i>Chamaecrista calycioides</i>	23	106	<i>Senna spectabilis</i>	27
40	<i>Chamaecrista diphylla</i>	25	107	<i>Senna trachypus</i>	35
41	<i>Chamaecrista ensiformis</i>	26	108	<i>Senna uniflora</i>	31
42	<i>Chamaecrista flexuosa</i>	26	109	<i>Senna velutina</i>	27
43	<i>Chamaecrista hispidula</i>	20	110	<i>Sida angustissima</i>	24
44	<i>Chamaecrista nictitans</i>	26	111	<i>Sida ciliaris</i>	16
45	<i>Chamaecrista ramosa</i>	29	112	<i>Sida cordifolia</i>	24
46	<i>Chamaecrista repens</i> var. <i>multijuga</i>	34	113	<i>Sida galheirensis</i>	26
47	<i>Chamaecrista serpens</i>	23	114	<i>Sida linifolia</i>	26
48	<i>Chloroleucon foliosum</i>	24	115	<i>Sida santaremensis</i>	25
49	<i>Citrus x limonia</i>	34	116	<i>Sida spinosa</i>	25
50	<i>Coccoloba latifolia</i>	26	117	<i>Simarouba amara</i>	38
51	<i>Cocos nucifera</i>	29	118	<i>Simarouba versicolor</i>	32
52	<i>Combretum lanceolatum</i>	34	119	<i>Solanum paniculatum</i>	31
53	<i>Commelina benghalensis</i>	31	120	<i>Spartina ciliata</i>	35
54	<i>Conocarpus erectus</i>	26	121	<i>Spathodea campanulata</i>	28
55	<i>Cordia oncocalyx</i>	29	122	<i>Sphagneticola trilobata</i>	36
56	<i>Curatella americana</i>	29	123	<i>Spondias macrocarpa</i>	26
57	<i>Crateva tapia</i>	35	124	<i>Swartzia simplex</i> var. <i>grandiflora</i>	35
58	<i>Crotalaria spectabilis</i> Roth	25	125	<i>Syzygium malaccense</i>	23
59	<i>Croton hirtus</i>	25	126	<i>Tanaecium seloi</i>	26
60	<i>Cynodon dactylon</i>	26	127	<i>Tarenaya spinosa</i>	25
61	<i>Dahlia imperialis</i>	23	128	<i>Terminalia catappa</i>	29
62	<i>Dalechampia scandens</i>	26	129	<i>Tephrosia purpurea</i>	21
63	<i>Delonix regia</i>	28	130	<i>Trischidium molle</i>	32
64	<i>Dicliptera mucronifolia</i>	24	131	<i>Tridax procumbens</i>	25
65	<i>Drimys granadensis</i>	26	132	<i>Wedelia paludosa</i>	30
66	<i>Echinodorus lanceolatus</i>	25	133	<i>Zornia latifolia</i>	24

Observando a Tabela 3 pode ser visto que a base de imagens contém múltiplas espécies do mesmo gênero, apresentando frequentemente características visuais semelhantes, como pode ser visto na Figura 16. Essa alta similaridade entre as classes torna o problema desafiador, já que as características extraídas computacionalmente precisam ser capazes de diferenciar estas pequenas variações.



Figura 16 – Exemplos de amostras do gênero *Chamaecrista*; da esquerda para direita IDs 40 (*diphylla*), 41 (*ensiformis*) e 42 (*flexuosa*).

### 3.3.2 Base de Madeira

Em parceria com o grupo LPF (Laboratório de Produtos Florestais) foi construída uma nova base de imagens microscópicas contendo 281 espécies de madeira. Essa base contém amostras das mais valiosas e ameaçadas espécies de árvores, principalmente do Bioma Amazônico. Para cada espécie existem três conjuntos de 20 amostras, correspondentes a diferentes cortes feitos no tronco: transversal, radial e tangencial. A Figura 17 ilustra os cortes realizados. Considerando o eixo longitudinal interceptando a medula de um tronco de madeira, então o corte transversal é o plano perpendicular ao eixo longitudinal. O corte radial é o plano vertical desde a medula até a casca. O corte tangencial é o plano tangencial dos anéis de crescimento.

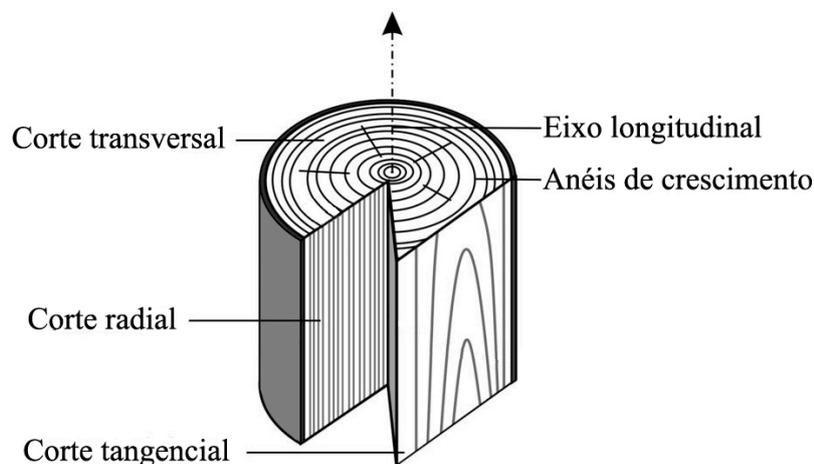


Figura 17 – Os três principais planos de um tronco: transversal, radial, and tangencial. Adaptada de (FRANKE; QUENNEVILLE, 2011)

Todas as madeiras selecionadas foram serradas em pedaços menores, medindo cerca de  $3 \times 3\text{cm}$ , usando uma serra de bancada. As amostras de madeira foram então submetidas a um processo de polimento e limpeza que busca destacar as características anatômicas da madeira. Os cortes passam por um processo de polimento em três níveis: primeiro, uma lixa grossa de 180 grãos, indo para uma lixa intermediária de 600 grãos e, finalmente, uma lixa de água de 1200 grãos. Nesse processo, é importante evitar o entupimento de poros artificiais, portanto são utilizados jatos de água e escovas para limpar o pó liberado pela madeira. Ambas as superfícies radial e tangencial estão rachadas, de modo que os raios podem ser vistos naturalmente. Esse polimento é fundamental para expor e contrastar as diferentes fisionomias anatômicas apresentadas pelas espécies. Todas as imagens foram obtidas utilizando um estereomicroscópio Olympus SZX7 de maneira padronizada, utilizando uma ampliação de  $20\times$ .

As Figuras 19, 20 e 21 ilustram uma imagem de cada espécie de madeira considerando um ID incremental percorrendo sequencialmente linha a linha do canto superior esquerdo até o canto inferior direito. As Tabelas 4 e 5 associam o ID ao nome científico de cada espécie da base de imagens.

Por meio de uma breve análise visual é fácil se confundir na diferenciação das espécies, até mesmo no corte transversal utilizado pelos especialistas. A Figura 18 ilustra duas espécies nos três cortes. Percebe-se que praticamente não há diferenças nos cortes tangencial e radial, justificando a necessidade do corte transversal utilizado na maioria dos estudos da literatura. A comparação destes cortes ajuda a determinar se métodos computacionais são capazes de distinguir pequenas diferenças imperceptíveis visualmente.

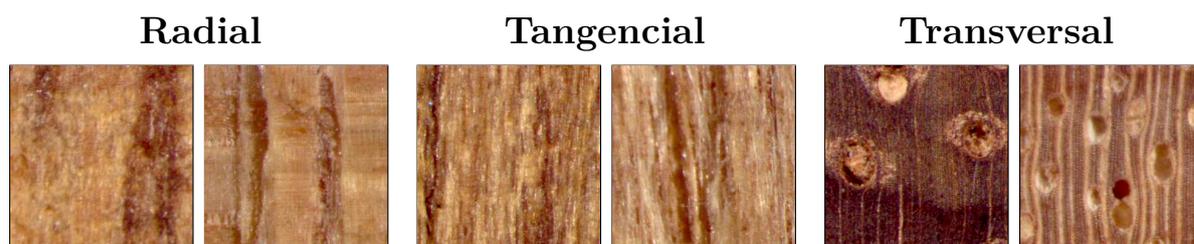


Figura 18 – Exemplos de amostras nos três cortes: à esquerda ID 254 (*Tachigali poeppigiana*) e à direita ID 269 (*Virola michelli*).

Destaca-se a importância deste conjunto de imagens pela padronização na aquisição e preparação das amostras. Isso é fundamental para evitar alguma tendência no desempenho das metodologias computacionais. Entretanto, os procedimentos realizados levam tempo considerável para serem executados, enfatizando um dos desafios na construção deste conjunto de imagens. Além disso, a disponibilidade de um elevado número de espécies enriquece o estudo desta aplicação, superando em 2,5 vezes o maior conjunto estudado na literatura.

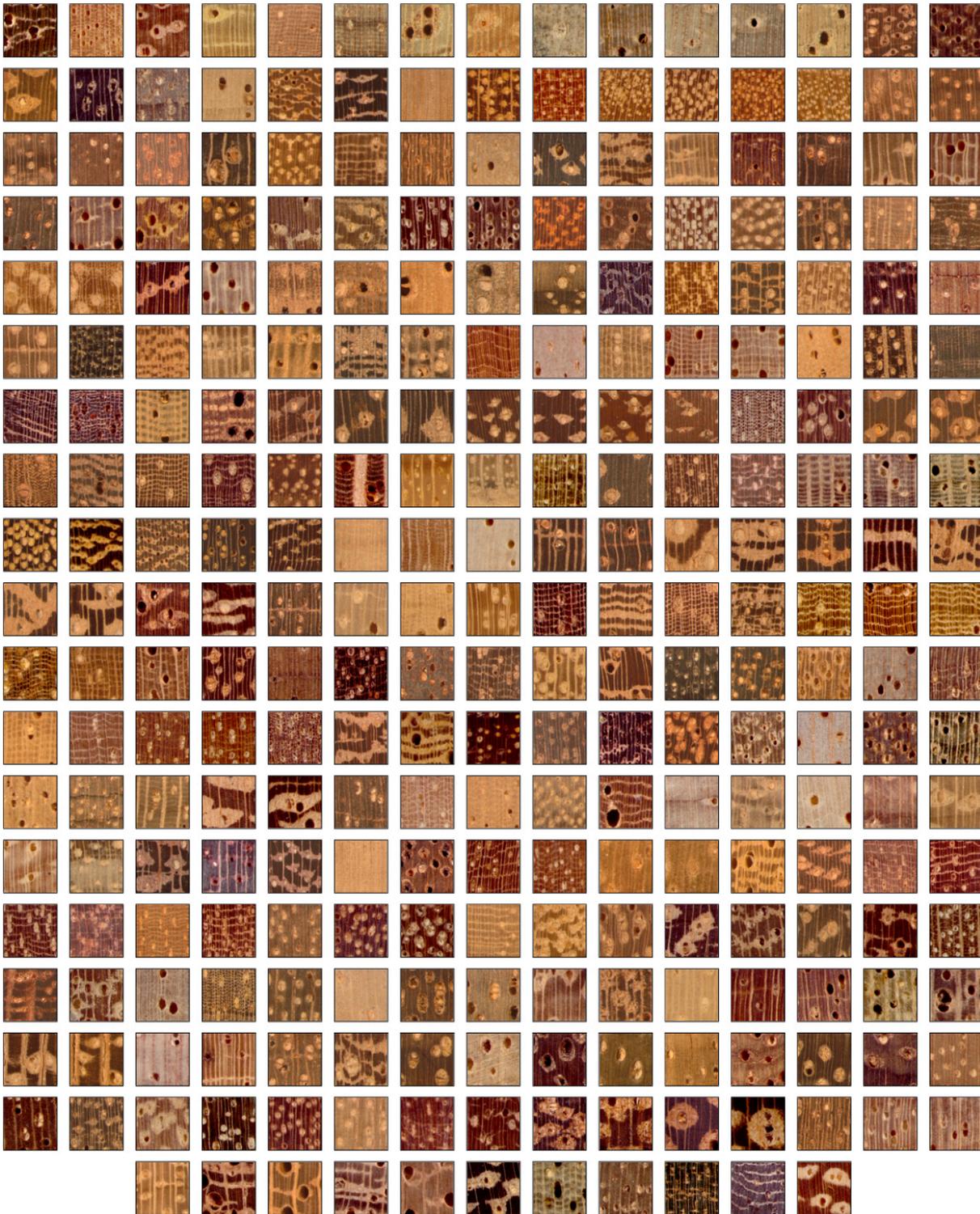


Figura 19 – Exemplo de uma amostra de cada espécie de madeira no corte transversal.



Figura 20 – Exemplo de uma amostra de cada espécie de madeira no corte radial.

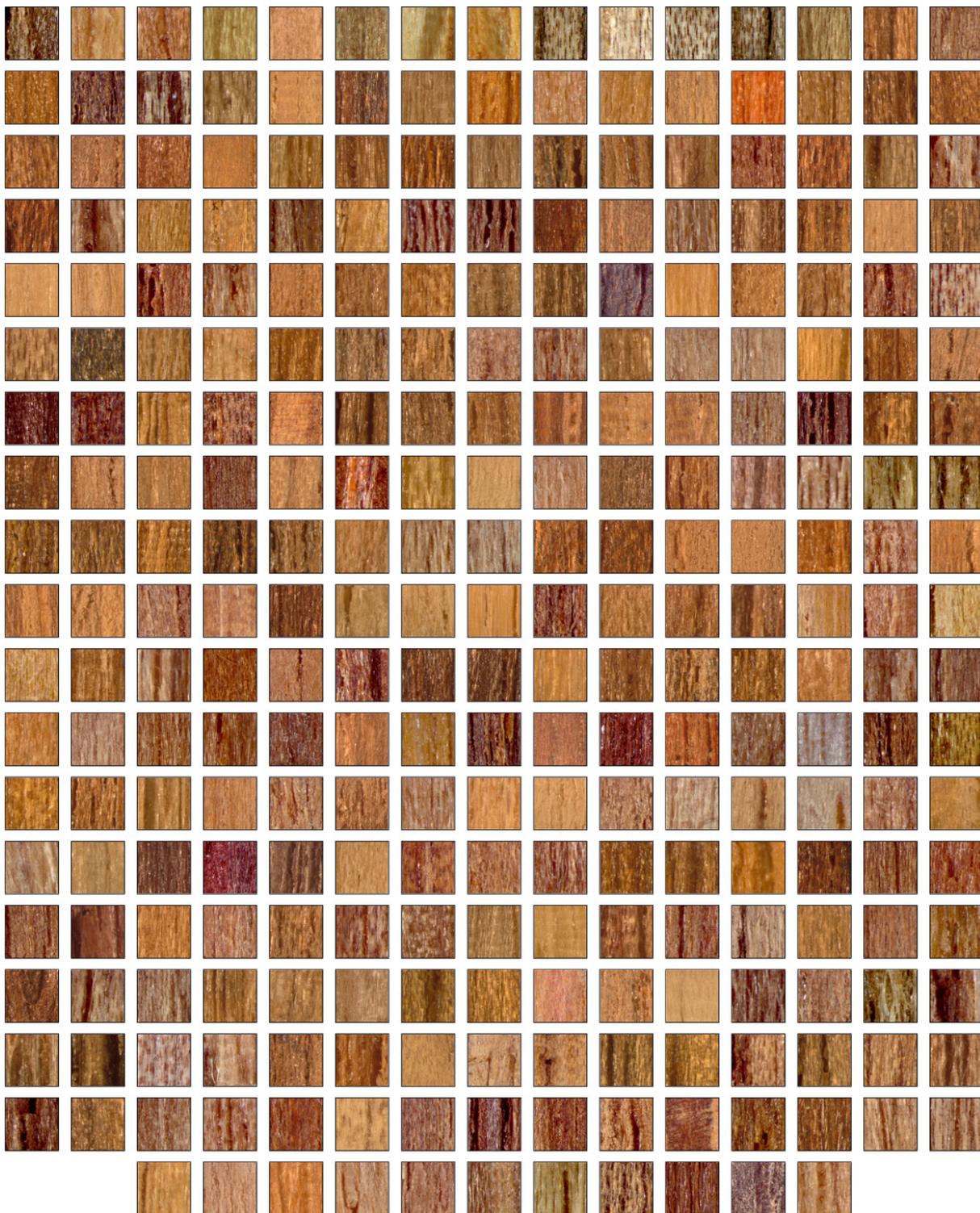


Figura 21 – Exemplo de uma amostra de cada espécie de madeira no corte tangencial.

Tabela 4 – Espécies e ID da base de imagens de madeira.

ID	Espécie	ID	Espécie	ID	Espécie	ID	Espécie
1	<i>Acioa edulis</i>	37	<i>Bicuiba oleifera</i>	73	<i>Commiphora leptophloeos</i>	109	<i>Eschweilera grandiflora</i>
2	<i>Agonandra brasiliensis</i>	38	<i>Bixa arborea</i>	74	<i>Copaifera duckei</i>	110	<i>Eucalyptus grandis</i>
3	<i>Albizia pedicellaris</i>	39	<i>Bowdichia nitida</i>	75	<i>Copaifera multijuga</i>	111	<i>Euplassa pinnata</i>
4	<i>Alexa grandiflora</i>	40	<i>Bowdichia virgilioides</i>	76	<i>Copaifera reticulata</i>	112	<i>Euxylophora paraensis</i>
5	<i>Allantoma decandra</i>	41	<i>Brosimum acutifolium</i>	77	<i>Cordia americana</i>	113	<i>Gallesia integrifolia</i>
6	<i>Allantoma lineata</i>	42	<i>Brosimum gaudichaudii</i>	78	<i>Cordia americana alb</i>	114	<i>Glycydendron amazonicum</i>
7	<i>Amburana acraea</i>	43	<i>Brosimum paraense</i>	79	<i>Cordia bicolor</i>	115	<i>Goniorrhachis marginata</i>
8	<i>Amburana cearensis</i>	44	<i>Brosimum parinarioides</i>	80	<i>Cordia goeldiana</i>	116	<i>Goupia glabra</i>
9	<i>Anacardium giganteum</i>	45	<i>Brosimum potabile</i>	81	<i>Cordia oncocalyx</i>	117	<i>Guarea guidonia</i>
10	<i>Anacardium microsepalum</i>	46	<i>Brosimum rubescens</i>	82	<i>Cordia sagotii</i>	118	<i>Guatteria olivacea</i>
11	<i>Anacardium occidentale</i>	47	<i>Brosimum utile</i>	83	<i>Couepia robusta</i>	119	<i>Guatteria poeppigiana</i>
12	<i>Anacardium parvifolium</i>	48	<i>Buchenavia grandis</i>	84	<i>Couma macrocarpa</i>	120	<i>Guatteria procera</i>
13	<i>Anacardium spruceanum</i>	49	<i>Buchenavia huberi</i>	85	<i>Couratari guianensis</i>	121	<i>Handroanthus impetiginosus</i>
14	<i>Anadenanthera colubrina</i>	50	<i>Buchenavia macrophylla</i>	86	<i>Couratari macrosperma</i>	122	<i>Handroanthus incanus</i>
15	<i>Anaueria brasiliensis</i>	51	<i>Buchenavia parvifolia</i>	87	<i>Couratari multiflora</i>	123	<i>Handroanthus ochraceus</i>
16	<i>Andira surinamensis</i>	52	<i>Byrsonima crista</i>	88	<i>Couratari oblongifolia</i>	124	<i>Handroanthus serratifolius</i>
17	<i>Aniba canelilla</i>	53	<i>Byrsonima stipulacea</i>	89	<i>Curatella americana</i>	125	<i>Handroanthus umbellatus</i>
18	<i>Aniba roseadora</i>	54	<i>Caesalpinia echinata</i>	90	<i>Dalbergia nigra</i>	126	<i>Hebea brasiliensis</i>
19	<i>Apeiba echinata</i>	55	<i>Calophyllum brasiliense</i>	91	<i>Dalbergia spruceana</i>	127	<i>Huberodendron swietenoides</i>
20	<i>Apuleia leiocarpa</i>	56	<i>Calycophyllum spruceanum</i>	92	<i>Dialium guianense</i>	128	<i>Hura crepitans</i>
21	<i>Arapatiella psilophylla</i>	57	<i>Caraipa densifolia</i>	93	<i>Dicliananona calycina</i>	129	<i>Hymenaea courbaril</i>
22	<i>Araucaria angustifolia</i>	58	<i>Carapa guianensis</i>	94	<i>Dicoryna paraensis</i>	130	<i>Hymenaea parvifolia</i>
23	<i>Aspidosperma desmanthum</i>	59	<i>Cariniana legalis</i>	95	<i>Dinizia excelsa</i>	131	<i>Hymenolobium excelsum</i>
24	<i>Aspidosperma discolor</i>	60	<i>Cariniana micrantha</i>	96	<i>Diplotropis purpurea</i>	132	<i>Hymenolobium heterocarpum</i>
25	<i>Aspidosperma dispernum</i>	61	<i>Caryocar glabrum</i>	97	<i>Diplotropis racemosa</i>	133	<i>Hymenolobium modestum</i>
26	<i>Aspidosperma macrocarpon</i>	62	<i>Caryocar villosum</i>	98	<i>Dipteryx alata</i>	134	<i>Hymenolobium nitidum</i>
27	<i>Aspidosperma polyneuron</i>	63	<i>Cassia fastuosa</i>	99	<i>Dipteryx magnifica</i>	135	<i>Hymenolobium petraeum</i>
28	<i>Aspidosperma pyriforme</i>	64	<i>Castilla ulei</i>	100	<i>Dipteryx odorata</i>	136	<i>Hymenolobium pulcherrimum</i>
29	<i>Astronium gracile</i>	65	<i>Cedrela fissilis</i>	101	<i>Dipteryx polyphylla</i>	137	<i>Hymenolobium sericeum</i>
30	<i>Astronium graveolens</i>	66	<i>Cedrela odorata</i>	102	<i>Drypetes variabilis</i>	138	<i>Inga alba</i>
31	<i>Astronium lecontei</i>	67	<i>Cedrela cateniformis</i>	103	<i>Endopleura uchi</i>	139	<i>Inga paraensis</i>
32	<i>Astronium ulei</i>	68	<i>Ceiba pentandra</i>	104	<i>Enterolobium maximum</i>	140	<i>Iryanthera grandis</i>
33	<i>Astronium urundeuva</i>	69	<i>Centrolobium robustum</i>	105	<i>Enterolobium schomburgkii</i>	141	<i>Jacaranda copaia</i>
34	<i>Bagassa guianensis</i>	70	<i>Chamaecrista scleroxylon</i>	106	<i>Eriotheca longipedicellata</i>	142	<i>Joannesia heveoides</i>
35	<i>Balfourodendron riedelianum</i>	71	<i>Chrysophyllum lucentifolium</i> subsp. pachycarpum	107	<i>Erisma uncinatum</i>	143	<i>Laetia procera</i>
36	<i>Bertholletia excelsa</i>	72	<i>Clarisia racemosa</i>	108	<i>Eschweilera coriacea</i>	144	<i>Lecythis idatimon</i>

Tabela 5 – Espécies e ID da base de imagens de madeira.

ID	Espécie	ID	Espécie	ID	Espécie	ID	Espécie
145	<i>Lecythis lurida</i>	179	<i>Ocotea aciphylla</i>	213	<i>Pouteria obscura</i>	247	<i>Tachigali aurea</i>
146	<i>Lecythis pisonis</i>	180	<i>Ocotea cymbarum</i>	214	<i>Pouteria sect franchetella</i>	248	<i>Tachigali cavipes</i>
147	<i>Lecythis poiteaui</i>	181	<i>Ocotea fragrantissima</i>	215	<i>Protium heptaphyllum</i>	249	<i>Tachigali chrysophylla</i>
148	<i>Licania gracilipes</i>	182	<i>Ocotea porosa</i>	216	<i>Protium tenuifolium</i>	250	<i>Tachigali glauca</i>
149	<i>Licania heteromorpha</i>	183	<i>Onychopetalum amazonicum</i>	217	<i>Pseudopiptadenia suaveolens</i>	251	<i>Tachigali myrmecophila</i>
150	<i>Licania kunthiana</i>	184	<i>Ormosia coccinea</i>	218	<i>Pterocarpus rohrii</i>	252	<i>Tachigali paraensis</i>
151	<i>Licania octandra</i>	185	<i>Ormosia paraensis</i>	219	<i>Pterodon pubescens</i> (Benth.)	253	<i>Tachigali pilgeriana</i>
152	<i>Lueheopsis duckyana</i>	186	<i>Osteophloeum platyspermum</i>	220	<i>Pterogyne nitens</i>	254	<i>Tachigali poeppigiana</i>
153	<i>Lueheopsis rosea</i>	187	<i>Parahancornia amapa</i>	221	<i>Qualea brevipedicellata</i>	255	<i>Tapirira guianensis</i>
154	<i>Maclura tictoria</i>	188	<i>Parahancornia fasciculata</i>	222	<i>Qualea dinizii</i>	256	<i>Tectona grandis</i>
155	<i>Macrobium acaciifolium</i>	189	<i>Paratecoma peroba</i>	223	<i>Qualea gracilior</i>	257	<i>Terminalia amazonia</i>
156	<i>Manilkara bidentata</i>	190	<i>Parinari excelsa</i>	224	<i>Qualea paraensis</i>	258	<i>Terminalia argentea</i>
157	<i>Manilkara cavalcantei</i>	191	<i>Parkia gigantocarpa</i>	225	<i>Rauvolfia paraensis</i>	259	<i>Tetragastris altissima</i>
158	<i>Manilkara huberi</i>	192	<i>Parkia multijuga</i>	226	<i>Roupala montana</i>	260	<i>Tetragastris panamensis</i>
159	<i>Maquira sclerophylla</i>	193	<i>Parkia nitida</i>	227	<i>Ruizterania albiflora</i>	261	<i>Trattinnickia burserifolia</i>
160	<i>Martiodendron elatum</i>	194	<i>Parkia paraensis</i>	228	<i>Sapium marmieri</i>	262	<i>Trichilia lecoinei</i>
161	<i>Melanoxylon brauna</i>	195	<i>Parkia pendula</i>	229	<i>Savia dictyocarpa</i>	263	<i>Vantanea parviflora</i>
162	<i>Mezilaurus itauba</i>	196	<i>Parkia velutina</i>	230	<i>Schefflera morototoni</i>	264	<i>Vatairea guianensis</i>
163	<i>Mezilaurus lindaviana</i>	197	<i>Paulownia tomentosa</i>	231	<i>Schizolobium parahyba</i> var. <i>amazonicum</i>	265	<i>Vatairea paraensis</i>
164	<i>Micrandra minor</i>	198	<i>Peltogyne maranhensis</i>	232	<i>Sclerolobium aureum</i>	266	<i>Vatairea sericea</i>
165	<i>Micrandra rossiana</i>	199	<i>Peltogyne paniculata</i>	233	<i>Sclerolobium pilgerianum</i>	267	<i>Vataireopsis speciosa</i>
166	<i>Micropholis melinoniana</i>	200	<i>Peltogyne subsessilis</i>	234	<i>Senegalia polyphylla</i>	268	<i>Virola biculhyba</i>
167	<i>Micropholis mensalis</i>	201	<i>Pinus eliottii</i>	235	<i>Sextonia rubra</i>	269	<i>Virola michelli</i>
168	<i>Micropholys guyanensis</i>	202	<i>Piptadenia gonoacantha</i>	236	<i>Simarouba amara</i>	270	<i>Virola surinamensis</i>
169	<i>Micropholys venulosa</i>	203	<i>Planchonella oblanceolata</i>	237	<i>Sloanea nitida</i>	271	<i>Vitex polygama</i>
170	<i>Minquartia guianensis</i>	204	<i>Planchonella obscura</i>	238	<i>Sloanea terniflora</i>	272	<i>Vochysia guianenses</i>
171	<i>Mora paraensis</i>	205	<i>Plathymenia foliolosa</i>	239	<i>Spondias mombin</i>	273	<i>Vochysia maxima</i>
172	<i>Moronobea coccinea</i>	206	<i>Plathymenia reticulata</i>	240	<i>Sterculia apeibophylla</i>	274	<i>Vochysia obidensis</i>
173	<i>Mouriri collocarpa</i>	207	<i>Platonia insignis</i>	241	<i>Sterculia excelsa</i>	275	<i>Vochysia tomentosa</i>
174	<i>Myracrodruon urundeuva</i>	208	<i>Platymiscium trinitatis</i> var. <i>duckei</i>	242	<i>Sterculia speciosa</i>	276	<i>Vouacapoua americana</i>
175	<i>Myrcarpus frondosus</i>	209	<i>Pouteria anomala</i>	243	<i>Stryphnodendron pulcherrimum</i>	277	<i>Xylopia nitida</i>
176	<i>Myroxylon balsamum</i>	210	<i>Pouteria caimito</i>	244	<i>Swartzia recurva</i>	278	<i>Zanthoxylum rhoifolium</i>
177	<i>Nectandra cuspidata</i>	211	<i>Pouteria gongrijpii</i>	245	<i>Swietenia macrophylla</i>	279	<i>Ziziphus cinnamomum</i>
178	<i>Ochroma pyramidalis</i>	212	<i>Pouteria oblanceolata</i>	246	<i>Symphonia globulifera</i>	280	<i>Zollemia paraensis</i>
						281	<i>Zygia racemosa</i>

### 3.3.3 Base de Madeira - Dispositivo portátil

A maioria dos trabalhos tem focado no reconhecimento de madeiras, os quais realizam procedimentos meticulosos que envolvem vários níveis de polimento e limpeza da superfície (IBRAHIM et al., 2016). Embora estes procedimentos auxiliem a melhor expor as características anatômicas da madeira, eles levam tempo considerável para serem realizados e se mostram impraticáveis fora de um laboratório.

Focando em construir uma metodologia que possa ser usada por um especialista na fiscalização do contrabando de madeira em campo, então foi proposto uma simplificação do procedimento de captura das imagens. Em parceria com o grupo LPF construímos um conjunto de imagens contendo amostras de 11 espécies de madeira com alto valor comercial.

Para a aquisição das imagens, as amostras do tronco não foram submetidas a nenhum procedimento de lixamento. A única preparação realizada foi um corte através de um estilete que ajuda a expor as características anatômicas da madeira. Todas as imagens foram capturadas por um microscópio portátil de baixo custo com resolução  $640 \times 480$  conectado a um *smartphone*. A Figura 22 mostra o microscópio portátil e uma amostra de corte do tronco.



Figura 22 – Captura das imagens com o microscópio portátil de baixo custo.

A base de imagens é composta de 440 imagens contendo 40 imagens para cada uma das 11 espécies. A Figura 23 ilustra uma imagem de cada espécie considerando um ID incremental percorrendo sequencialmente linha a linha do canto superior esquerdo até o canto inferior direito. A Tabela 6 associa o ID ao nome científico de cada espécie da base de imagens. Esta base de imagens está disponível para o desenvolvimento de pesquisas no link <[http://www.facom.ufu.br/~backes/wood\\_dataset.html](http://www.facom.ufu.br/~backes/wood_dataset.html)>.

Vale ressaltar que semelhante à base de vírus analisada nos experimentos deste trabalho, esta base também apresenta alta variabilidade intra-classe, demonstrando um dos

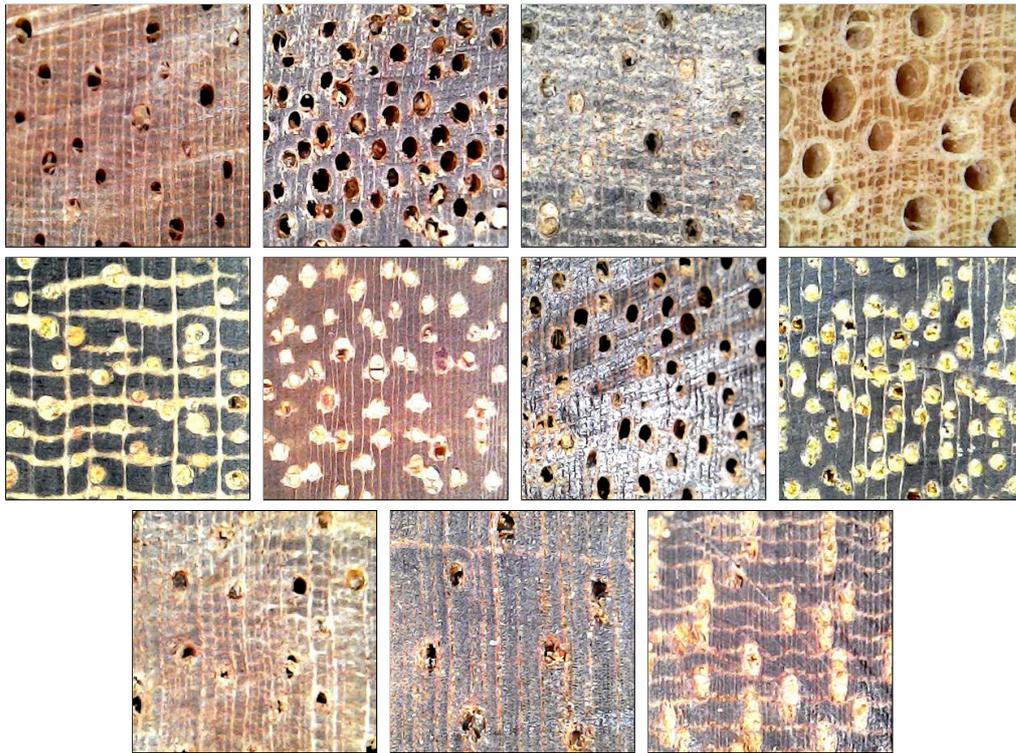


Figura 23 – Exemplo de uma amostra de cada espécie de madeira.

Tabela 6 – Espécies e ID da base de imagens de madeira.

ID	Família	Espécie
1	<i>Lecythidaceae</i>	<i>Allantoma decandra</i>
2	<i>Calophyllaceae</i>	<i>Caraipa densifolia</i>
3	<i>Lecythidaceae</i>	<i>Cariniana micrantha</i>
4	<i>Caryocaraceae</i>	<i>Caryocar villosum</i>
5	<i>Moraceae</i>	<i>Clarisia racemosa</i>
6	<i>Fabaceae</i>	<i>Dipteryx odorata</i>
7	<i>Goupiaceae</i>	<i>Goupia glabra</i>
8	<i>Bignoniaceae</i>	<i>Handroanthus incanus</i>
9	<i>Malvaceae</i>	<i>Lueheopsis duckeana</i>
10	<i>Myristicaceae</i>	<i>Osteophloeum platyspermum</i>
11	<i>Sapotaceae</i>	<i>Pouteria caimito</i>

desafios na aplicação proposta. A primeira linha da Figura 24 ilustra 5 amostras da espécie *Allantoma decandra*. Percebe-se que as amostras apresentam diferentes colorações, apresentando tons acinzentados que indicam principalmente degradação fotoquímica causada pela ação de raios ultravioletas (AYADI et al., 2003). Outro fator observado nas amostras de *Osteophloeum platyspermum* (Figura 24) são os marcantes anéis de crescimento com diferentes larguras, que tem variabilidade devido à fatores como baixa fertilidade, ataques de praga e doenças (JAYAWICKRAMA et al., 1997).

Um dos grandes desafios desse tipo de aplicação é a criação de conjuntos de dados grandes o suficiente, tanto para cobrir o número de espécies existentes, quanto para cobrir

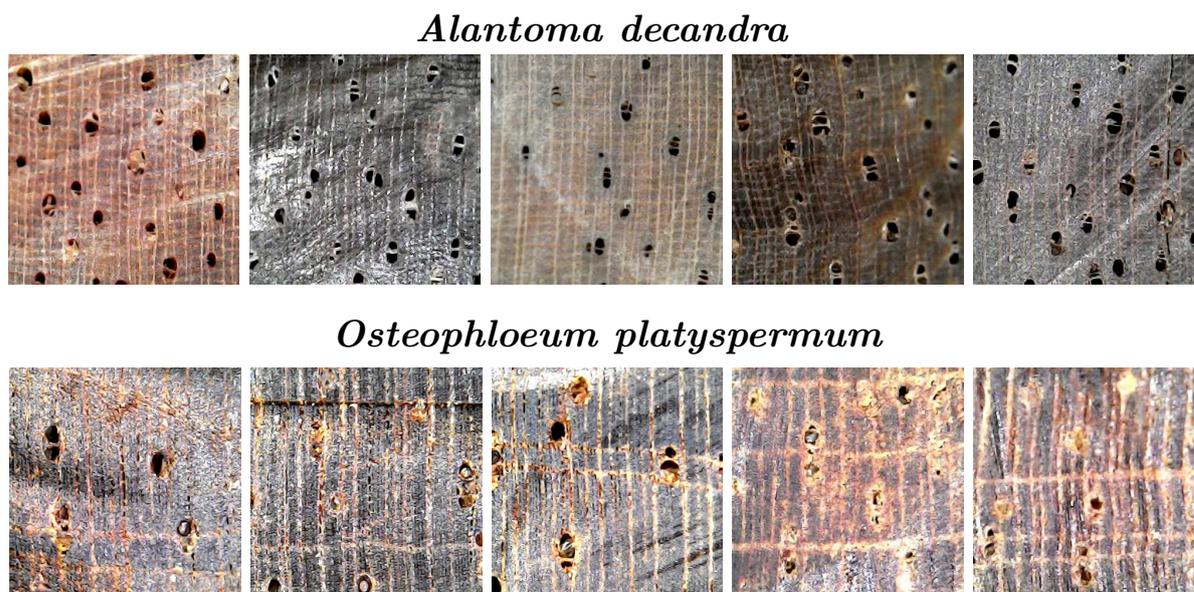


Figura 24 – Na primeira linha amostras da espécie do ID 1 (*Alantoma decandra*), enquanto na segunda linha amostras da espécie do ID 10 (*Osteophloeum platyspermum*).

o número de amostras por espécie necessárias para representar toda sua variabilidade. Para tanto, é desejável a colaboração de outros centros de pesquisa em anatomia de madeira para fornecer imagens de amostras.

Diante dessas colaborações, faz-se necessário padronizar o protocolo de aquisição para que o banco de imagens construído não apresente nenhuma tendência que prejudique o processo de classificação. Devido ao microscópio portátil utilizado ser de propósito geral, ele fornece ajustes que dão liberdade no processo de aquisição (por exemplo, o ajuste de zoom e luminosidade), e portanto, introduziriam tendência no banco colaborativo construído.

Buscando o desenvolvimento de um dispositivo de baixo custo mais adequado ao problema, foi desenvolvido um microscópio portátil em carcaça de aço inoxidável (Inox 304) e plástico ABS impresso em 3D. Para melhor visualização de seus componentes a Figura 25 ilustra a vista explodida do equipamento. O módulo de captura de imagens, composto da câmera em sua placa eletrônica, lente com zoom ajustável e circuito de leds ao seu redor, foi adquirido do mesmo fabricante do microscópio utilizado anteriormente.

Esta versão, além de ser mais robusta a danos externos devido a resistência do Inox, proporciona vantagens que beneficiam a padronização na aquisição das imagens. A pirâmide com plano cortado em seu topo, limita a área de interesse em 4,3mm x 2,9mm (veja Figura 26), evitando o desafio de cortar uma área plana maior na madeira. Além disso, a altura da pirâmide define uma distância padrão até o objeto, eliminando a necessidade de calibração da distância focal. O vidro temperado *extra clear* protege a lente da câmera e facilita sua limpeza caso o pequeno pó de serragem da madeira adentre o equipamento. Por fim, com a distância da lente ao objeto definida, foi possível substituir a regulagem

de intensidade de luz por uma resistência fixa que mantém a iluminação padrão.

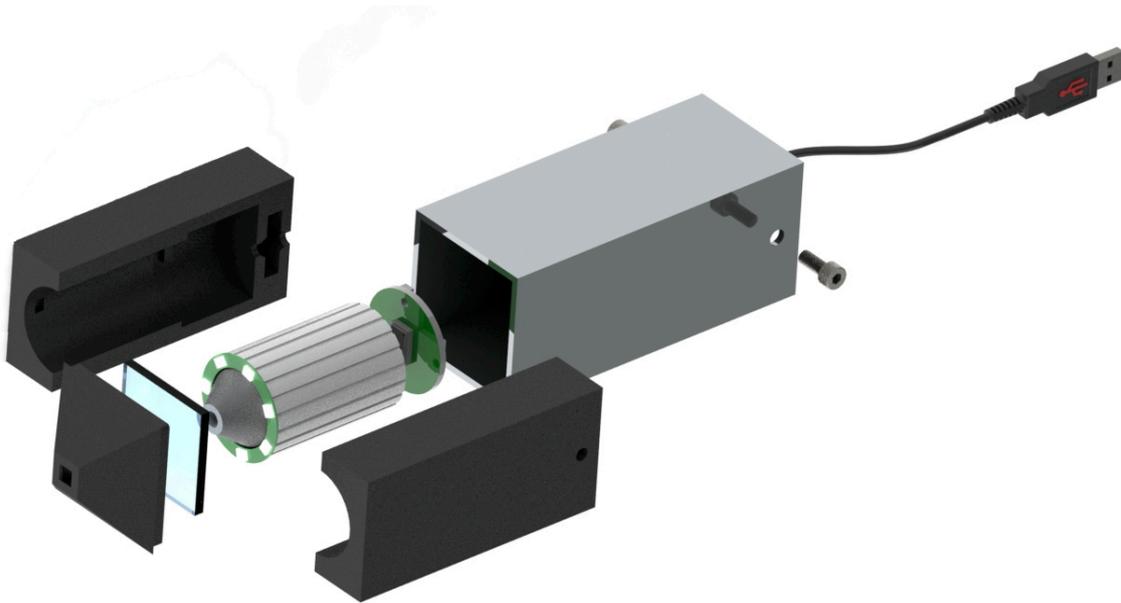


Figura 25 – Vista explodida do microscópio portátil

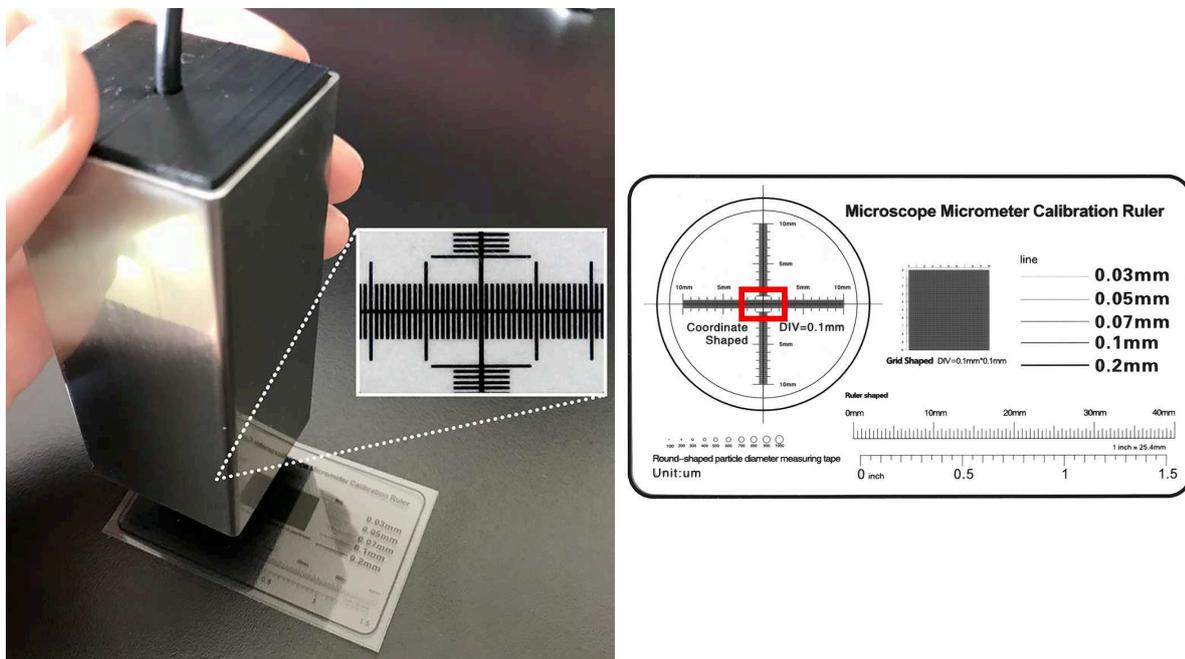


Figura 26 – Microscópio portátil focado na área em vermelho indicada no cartão de calibração, onde cada divisão equivale a 0,1mm.

### 3.3.4 Base de Vírus

Essa base é composta de 1500 imagens de vírus adquiridas por microscópio eletrônico de transmissão. Essas imagens representam 15 diferentes tipos de vírus: Adenovirus,

Astrovirus, CCHF, Cowpox, Dengue, Ebola, Influenza, Lassa, Marburg, Norovirus, Orf, Papilloma, Rift Valley, Rotavirus, WestNile. No procedimento de captura foram usadas diferentes magnificações com tamanho de pixel variando entre 0,5nm e 5nm. Para que tenham texturas comparáveis, as amostras foram redimensionadas através de interpolação bilinear para que o raio da partícula de vírus seja representada por 20 pixels.

Cada tipo de vírus contém 100 imagens de resolução  $41 \times 41$  pixels em tons de cinza. Uma amostra de cada tipo pode ser vista na Figura 27. As imagens podem ser acessadas em [www.cb.uu.se/~gustaf/virustexture/](http://www.cb.uu.se/~gustaf/virustexture/) e estão disponíveis nos formatos de 8-bits e 16-bits.

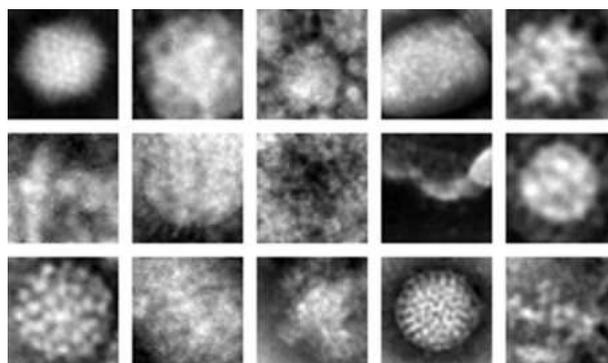


Figura 27 – Base de imagens de vírus. Do canto superior esquerdo ao canto inferior direito: Adenovirus, Astrovirus, CCHF, Cowpox, Dengue, Ebola, Influenza, Lassa, Marburg, Norovirus, Orf, Papilloma, Rift Valley, Rotavirus and West-Nile

### 3.4 Considerações finais

Este capítulo introduziu os conceitos fundamentais que envolvem o aprendizado automático de características através das CNNs, destacou as inovações das arquiteturas avaliadas nesse trabalho, assim como os métodos de características pré-projetadas utilizados para comparação. Ademais, foi descrita a base de 15 tipos de vírus introduzida anteriormente na literatura, e outras três construídas durante esse trabalho: imagens microscópicas de 134 espécies de pólen; imagens microscópicas de 281 espécies de madeira em 3 cortes; e imagens de 11 espécies de madeira amazônica capturadas por dispositivo portátil com protocolo de aquisição simplificado.



## Experimentos e Análise dos Resultados

Dentre os principais objetivos desta tese, foi proposto o estudo e desenvolvimento de metodologias para a classificação de novos conjuntos de imagens microscópicas de pólen e madeira. Além disso, foi selecionada também uma base de imagens de vírus que se mostra desafiadora devido à baixa acurácia reportada em estudos anteriores.

Em todos os experimentos foram realizados estudos comparativos entre extratores pré-projetados de característica e aprendizado automático de características utilizando CNNs. A seguir, são descritos os procedimentos e parâmetros utilizados para treinamento das CNNs, assim como os resultados obtidos em cada uma das bases de imagens.

### 4.1 Treinamento das CNNs

A escolha da função de perda e método de otimização no treinamento de CNNs tem influência direta na convergência e acurácia da arquitetura. Em todos os casos foi utilizado como função de perda o cálculo da entropia-cruzada  $L$ , comumente utilizada em problemas de classificação (BOSMAN; ENGELBRECHT; HELBIG, 2020):

$$L = - \sum_i p_i \log(f_i) \quad (10)$$

onde  $p$  é a distribuição de saída ideal e  $f$  é a distribuição estimada pela rede.

A distribuição ideal  $p$  é dada para cada imagem do conjunto de dados. Considerando um conjunto de  $n$  amostras, contendo um número  $c$  de classes, a distribuição é dada por:

$$p_i^n = (p_1^n, \dots, p_c^n) \quad (11)$$

$$p_i = \begin{cases} 1, & \text{se } i = \text{classe da imagem} \\ 0, & \text{caso contrário} \end{cases}$$

Para atualizar os pesos das CNNs minimizando o erro da entropia-cruzada foi utilizado o método do SGD (do inglês, *Stochastic Gradient Descent*). Para tanto, foram separadas aleatoriamente 10% das amostras para teste, 10% para validação e 80% para treinamento.

Quando realizado o treinamento a partir dos pesos aleatórios e como extrator de características (treinando apenas as camadas de classificação) foi selecionada uma taxa de aprendizado inicial de  $10^{-2}$ , sendo a mesma dividida por 10 quando não houver melhoria da acurácia por 3 épocas consecutivas. Para o ajuste-fino, o procedimento tem taxa inicial de  $10^{-4}$ , já que é desejável que os pesos sejam modificados minimamente a fim de não perder os pesos aprendidos anteriormente.

Dado que os conjuntos de imagens contêm um pequeno número de amostras por classe, é essencial o uso de técnicas de aumento de dados para aplicar CNNs. Basicamente, essa é uma técnica para aumentar o tamanho dos conjuntos de treinamento, aplicando transformações que preservam os rótulos das classes (CIRESAN et al., 2010), (KRIZHEVSKY; SUTSKEVER; HINTON, 2012), (SAJJADI; JAVANMARDI; TASHIZEN, 2016). O aumento de dados pode ser visto como uma maneira de inserir conhecimento sobre variações em uma tarefa ou domínio. As transformações de imagem mais comuns são rotação, translação, espelhamento, redimensionamento, adição de ruído e mudança de iluminação. O processo de aumento de dados também se justifica pela necessidade de um grande volume de dados para treinar a rede sem causar *overfitting* (SAJJADI; JAVANMARDI; TASHIZEN, 2016). Em nossos experimentos, aumentamos o conjunto de dados de treinamento executando rotações de um ângulo variando entre  $-15^\circ$  and  $15^\circ$  graus em cada imagem.

## 4.2 Classificação de pólen

Com o intuito de avaliar o desempenho de arquiteturas recentes de CNN em um conjunto elevado de classes com imagens de alta similaridade intra-classe, foram considerados três tipos de treinamento, o treinamento com pesos aleatórios e duas modalidades de transferência de aprendizado: ajuste-fino e CNN como extrator de características. As arquiteturas avaliadas foram InceptionV3, DenseNet e ResNet, sendo todas pré-treinadas no conjunto de imagens ImageNet 2012.

A Figura 28 mostra as acurácias das arquiteturas de CNN InceptionV3, DenseNet-121 e ResNet-50 considerando os 3 tipos de treinamento. Quando considerado o treinamento da CNN como extrator de características, uma nova camada *softmax* foi criada e apenas a estrutura de classificação foi retreinada. Como pode ser visto, utilizar uma CNN como extrator de características não produziu resultados satisfatórios, sugerindo que características genéricas aprendidas em um grande conjunto de imagens não se aplicam à classificação de pólen. Por outro lado, o ajuste-fino sobre uma arquitetura pré-treinada em um grande conjunto de imagens produziu um ganho médio de acurácia de 6,65% quando comparado à inicialização aleatória de pesos.

Para comparar as arquiteturas de CNN com métodos de extração de característica pré-projetados foram selecionados métodos que apresentam resultados promissores em diversos estudos da literatura. Além disso, foram avaliados suas combinações e a redução

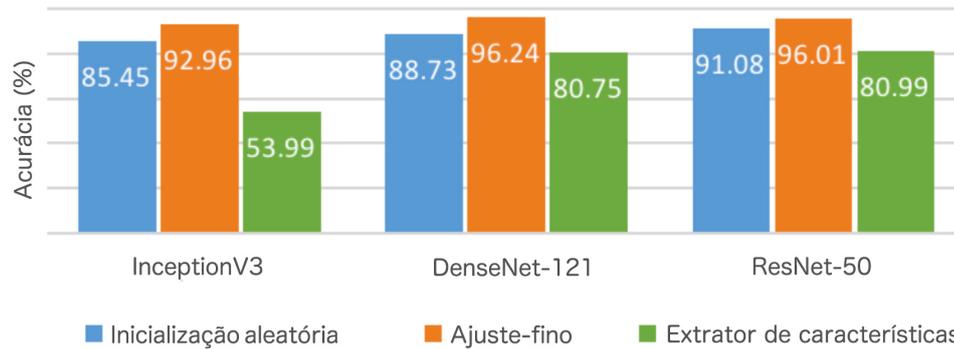


Figura 28 – Acurácia comparando a inicialização aleatória dos pesos, ajuste-fino e CNN como extrator de características.

de dimensionalidade por LDA. Em todos os experimentos foram utilizados kNN ( $k = 1$ ) como classificador, com validação cruzada realizada através de k-fold ( $k = 10$ ). A Tabela 7 resume as configurações utilizadas nos experimentos, onde #carac refere ao número de características produzidas.

Tabela 7 – Descrição dos métodos de característica pré-projetados. Não foram experimentados Filtros Gabor individualmente devido ao número elevado de características, que resultam em um alto custo computacional para classificação.

Método	Descrição/Configuração	#carac
HOG	usa tamanho de célula $32 \times 32$ pixels e tamanho de bloco $4 \times 4$ pixels	1296
LBP	usa tamanho de célula $32 \times 32$ pixels	4779
Gabor/LDA	aplica 40 filtros, realiza subamostragem de $4 \times 4$ e redução de dimensionalidade por LDA	134
HOG+LBP	concatena as características de HOG e LBP	6075
HOG+(Gabor/LDA)	concatena características HOG e Gabor/LDA	1430
LBP+(Gabor/LDA)	concatena características LBP e Gabor/LDA	4913
HOG+LBP+(Gabor/LDA)	concatena características HOG, LBP e Gabor/LDA	6209
(HOG+LBP+Gabor)/LDA	aplica LDA na concatenação das características produzidas por cada método	134
HOG/LDA+LBP/LDA+Gabor/LDA	aplica LDA nas características produzidas por cada método e concatena os conjuntos de características reduzidos	402

A Tabela 8 mostra os resultados de acurácia obtido para métodos de extração de características pré-projetadas versus aprendizado de características com ajuste fino. Pode-se observar que CNNs com ajuste-fino produzem resultados muito superiores aos de características pré-projetadas. Também foram experimentados diferentes tamanhos de ResNet, variando o número de camadas de 18 a 152. Pode ser observado que o número de camadas de ResNet não resultou em aumento da acurácia. Em relação as arquiteturas da

CNN, nota-se que o melhor resultado foi obtido pela DenseNet, que é a menor arquitetura em termos de número de parâmetros. Ao comparar a CNN tradicional Inception com CNNs baseadas em resíduos (ResNet e DenseNet), nota-se que as CNNs baseadas em resíduos alcançaram melhores resultados. Esse resultado pode ser justificado devido às conexões de atalho da DenseNet que minimizam o conhecido problema de gradientes explodindo/desvanecendo, que ocorre quando um grande número de camadas é empilhado.

Tabela 8 – Resultados obtidos para métodos de extração de características pré-projetadas comparados a aprendizado de características com ajuste-fino em diferentes arquiteturas de CNN.

		Método/Arquitetura	Acurácia
Características pré-projetadas		HOG	35,39%
		LBP	58,18%
		Gabor/LDA	49,09%
		HOG+LBP	56,26%
		HOG+(Gabor/LDA)	42,84%
		LBP+(Gabor/LDA)	<b>59,96%</b>
		HOG+LBP+(Gabor/LDA)	57,98%
		(HOG+LBP+Gabor)/LDA	48,68%
		HOG/LDA+LBP/LDA+Gabor/LDA	57,08%
Aprendizado de características		InceptionV3 (23,8M param)	92,96%
		DenseNet-121 (8,0M param)	<b>96,24%</b>
		Resnet-18 (11.2M param)	95,54%
		Resnet-34 (21.3M param)	93,66%
		Resnet-50 (25,6M param)	96,01%
		Resnet-101 (44,6M param)	94,84%
		Resnet-152 (60,3M param)	95,07%

### 4.3 Classificação de madeira

A maioria dos trabalhos em classificação automática de madeiras utiliza o corte transversal no tronco para suas amostras (SILVA et al., 2017; IBRAHIM et al., 2016; KWON et al., 2017). Isso ocorre principalmente pela sua disponibilidade, já que os especialistas utilizam este corte para o reconhecimento visual. No entanto, neste experimento serão avaliados os três cortes que compõe cada uma das 281 espécies de madeira da base de imagens, a fim de determinar o melhor corte para o reconhecimento automático.

Assim como no experimento realizado com a base de pólen, o conjunto de imagens de madeira analisado possui um elevado número de classes com características visuais semelhantes. Portanto, nesse experimento todas as arquiteturas avaliadas foram pré-treinadas no conjunto de imagens ImageNet 2012. Entretanto, apenas a Resnet-50 dentre suas variadas profundidades foi avaliada, devido a ter apresentado melhor resultado no experimento

anterior e por ser mais frequentes nos estudos da literatura. Além das três arquiteturas avaliadas no experimento anterior, foi avaliada também a arquitetura SqueezeNet, de tamanho muito menor quando comparada as outras arquiteturas. O procedimento de treinamento das CNNs realizado foi replicado do experimento anterior.

Além disso, para avaliar o reconhecimento automático utilizando características pré-projetadas, foi reproduzida a metodologia proposta em (SILVA et al., 2017), que por meio de imagens de corte transversal foi capaz de atingir 88% de acurácia em 77 espécies de madeira Africana. A metodologia utiliza o descritor LPQ seguido de redução de características por LDA, classificando por meio do método kNN com validação cruzada k-fold ( $k = 5$ ). Além da metodologia proposta, foram avaliados o descritor padrão LBP e a versão invariante a rotação do LPQ, nomeada RI-LPQ.

A Tabela 9 mostra os resultados de acurácia de características pré-projetadas versus aprendizado de características. Os melhores resultados foram encontrados no conjunto de imagens do corte transversal em ambas as metodologias. Este resultado está de acordo com especialistas, que relatam que a maioria das estruturas anatômicas pode ser visualizada no corte transversal.

Em termos de acurácia, as arquiteturas de CNN se mostram superiores quando comparadas aos métodos de característica pré-projetados, obtendo uma acurácia pelo menos 13% superior. A respeito dos experimentos de características pré-projetadas, os parâmetros  $P$  (célula) e  $R$  (raio) foram analisados no intervalo de 4 a 16 e 1 a 2, respectivamente. A melhor combinação de  $P = 12$  e  $R = 1$  alcançaram 55,23% de acurácia no corte transversal. Entretanto, sem a redução por LDA, a acurácia aumentou para 84,23%. O descritor RI-LPQ mostrou melhor performance apenas no corte transversal, sugerindo que a diferente organização das estruturas prioritárias (vasos, parênquima axial e fibras) podem beneficiar um descritor invariante a rotação.

A Figura 29 mostra 16 amostras de espécies da base de imagens, em que as amostras na mesma posição em cada bloco (Radial, Transversal e Tangencial) correspondem a mesma espécie. Nela é possível perceber que as características mais marcantes estão presentes no corte transversal. Considerando as arquiteturas de CNN o benefício se mostra pequeno, aproximadamente 2,5% maior em acurácia. Entretanto, nos métodos de características pré-projetadas esse aumento é de aproximadamente 25% e 43% para os cortes tangencial e radial, respectivamente. Isso sugere que as CNNs são capazes de extrair características que melhor representam as sutis mudanças na textura quando comparadas aos métodos de características pré-projetados e análise visual.

Também foi estimada a complexidade das arquiteturas medindo o tempo necessário para passar uma imagem pela rede, também chamada de tempo de inferência. Cada tempo registrado é a média de 100 execuções em CPU em uma máquina Intel Core i7, 2.3Ghz, 16GB de RAM executando macOS. A Tabela 10 mostra o número de parâmetros e o tempo de inferência para cada arquitetura. Dado que a arquitetura SqueezeNet possui

Tabela 9 – Métodos de características pré-projetados comparados a aprendizado de características.

Método/Arquitetura		Radial Acc.	Tang. Acc.	Trans. Acc.
Características pré-projetadas	LBP	26,37%	45,55%	41,17%
	LBP/LDA	7,90%	10,71%	23,90%
	LPQ	<b>42,79%</b>	<b>60,05%</b>	84,23%
	LPQ/LDA (SILVA et al., 2017)	14,20%	15,89%	55,23%
	RiLPQ	35,73%	49,80%	<b>85,93%</b>
	RiLPQ/LDA	6,57%	10,64%	36,26%
Aprendizado de características	InceptionV3	93,59%	93,41%	98,57%
	DenseNet	<b>96,44%</b>	96,26%	<b>98,75%</b>
	ResNet	95,37%	<b>96,44%</b>	98,22%
	SqueezeNet	88,96%	92,17%	97,86%



Figura 29 – Exemplos de amostras dos cortes radial, tangencial e transversal de 16 espécies dentre as 281 espécies da base de imagens de madeira.

um tempo menor de inferência e precisão comparável, isso sugere que, apesar de menor, essa CNN pode ser aplicada com sucesso ao problema de classificação de espécies de madeira. Além disso, o tempo de 113ms indica que esse tipo de rede pode ser incorporada a dispositivos móveis, como smartphones e tablets, com tempo de resposta aceitável, dado as limitações de *hardware* presentes nestes dispositivos.

Tabela 10 – Número de parâmetros em cada arquitetura de CNN e resultados de desempenho em tempo de inferência.

Arquitetura	Número de parâmetros	Tempo de inferência
InceptionV3	23,8M	717ms
ResNet-50	25,6M	525ms
DenseNet	8,0M	503ms
SqueezeNet	1,2M	113ms

## 4.4 Classificação de vírus

Diversas bases de imagens disponíveis na literatura não têm conjuntos específicos de treino e teste. Portanto, para melhor estimar a acurácia das CNNs foi investigado o uso de validação cruzada através do método k-fold.

Para treinar adequadamente todas as arquiteturas de CNN e evitar qualquer tendência, deve-se garantir que as diferentes arquiteturas sejam treinadas usando o mesmo subconjunto de amostras. Para tanto, foi dividido o conjunto de dados em 10 conjuntos separados ( $k = 10$ ) para que o mesmo subconjunto de amostras seja usado para treinar cada rede. Neste experimento, ao invés de treinar as redes com uma taxa adaptativa de aprendizado, foi utilizado uma taxa fixa de  $10^{-3}$ , empiricamente escolhido.

A Figura 30 mostra a evolução da acurácia de cada arquitetura no conjunto de teste durante seu treinamento. A Tabela 11 apresenta a acurácia média e desvio padrão das arquiteturas avaliadas após o treinamento de 50 épocas. Independente da arquitetura analisada, é possível notar que existe uma pequena variação da acurácia entre os *folds*. Embora as arquiteturas apresentem uma acurácia alta e um desvio padrão muito pequeno, existem subconjuntos de imagens que são mais difíceis de testar do que outros. Isso é esperado, pois o conjunto de dados usado nos experimentos apresenta baixa similaridade intra-classe, como mostra a Figura 31.

Tabela 11 – Acurácia média das arquiteturas de CNN.

Arquitetura	#Épocas	Acurácia
<b>DenseNet-121</b>	<b>50</b>	<b>89% ± 2,3</b>
ResNet-50	50	88,6% ± 2,0
InceptionV3	50	87,2% ± 2,8
SqueezeNet	50	82,6% ± 2,4
SqueezeNet	100	86,5% ± 1,7

Embora essa variação entre os *folds* esteja presente, a alta precisão indica que os mapas de características aprendidos pelas arquiteturas são robustos. Como resultado, essas características são capazes de identificar a classe de vírus mesmo com muitas distorções

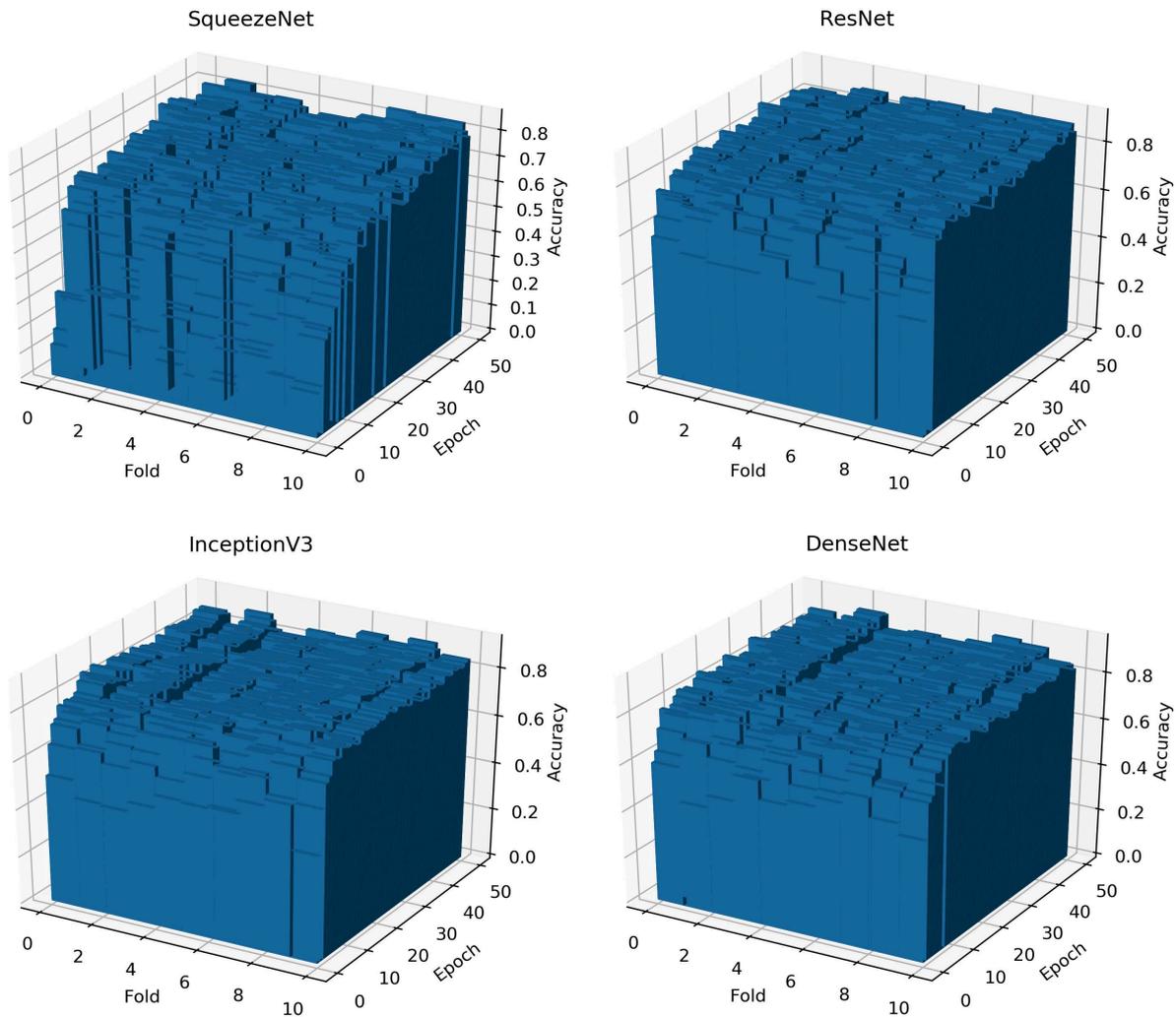


Figura 30 – Gráficos de barra da acurácia sobre épocas para cada arquitetura.

presentes nas imagens (por exemplo, translação e oclusão de partes do vírus), superando o problema de baixa similaridade intra-classe das imagens.

Outro ponto essencial a ser observado é a menor acurácia apresentada pela SqueezeNet, como mostrado na Tabela 11. Essa arquitetura apresentou a menor acurácia (82,6%), que é 4,6% menor que a obtida pela InceptionV3 (87,2%). Para investigar seu baixo desempenho, a Figura 32 apresenta a evolução da função de perda de cada arquitetura durante seu treinamento. Após 50 épocas, nota-se que, exceto na SqueezeNet, todas as outras arquiteturas de CNN apresentam uma função de perda próxima de zero e uma variação muito pequena entre os *folds*.

A SqueezeNet é a arquitetura de CNN com a menor quantidade de filtros convolucionais. Embora pareça contraditório, a SqueezeNet exige um número maior de épocas para treinar sua arquitetura reduzida de modo a obter um resultado semelhante a CNNs mais profundas, como mostra a Figura 33. Uma explicação para esse desempenho pode estar na própria estrutura do SqueezeNet, que usa poucas camadas de subamostragem para ob-

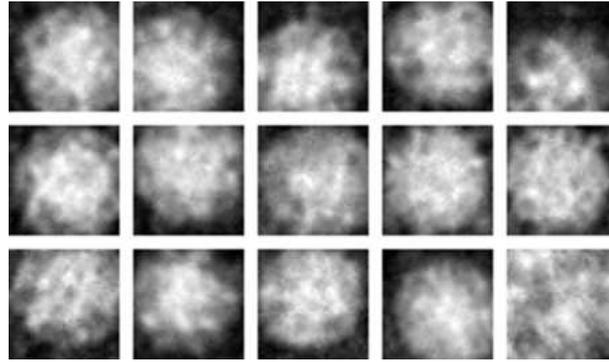


Figura 31 – Exemplo de variações em uma classe de vírus. Fonte: (KYLBERG; UPPSTRÖM; SINTORN, 2011)

ter um grande mapa de características. As camadas de subamostragem ajudam a reduzir a dimensionalidade da imagem e a sensibilidade da CNN a distorções e deslocamentos de imagem, problemas presentes na base de imagens de vírus. Ao usar 100 épocas para treinar essa arquitetura (Tabela 11), é possível reduzir sua função de perda média. No entanto, ainda existe uma variação significativa entre os *folds*. Como resultado, a acurácia média foi elevada de 82,6% para 86,5%, mas ainda assim, o menor valor de acurácia entre todas as CNNs avaliadas.

Por fim, quando comparados nossos resultados com outros trabalhos que abordaram o mesmo problema (veja Tabela 12), nosso melhor resultado (obtido pela DenseNet) é capaz de superar as mais altas acurácias de duas abordagens baseadas em características pré-projetadas, apresentando uma acurácia de 2,8% maior que o melhor método comparado (SANTOS et al., 2015). Este resultado demonstra a eficácia dos métodos de aprendizado de características em problemas de imagem que mostram alta similaridade entre as classe e baixa semelhança dentro da classe.

Tabela 12 – Comparação da acurácia média de métodos aplicados à classificação da base de imagens de vírus.

Métodos	Acurácia média
<b>DenseNet</b>	<b>89% ± 2,3</b>
Maior acurácia em (SANTOS et al., 2015)	85,7%
Maior acurácia em (WEN et al., 2016)	86,2% ± 2,0

## 4.5 Classificação de madeira - Dispositivo portátil

De maneira semelhante ao experimento anterior, avaliamos o desempenho das arquiteturas de CNN no conjunto de imagens de madeira portátil com interesse em descobrir se há variações no comportamento das CNNs relacionadas ao conjunto de treinamento. Para tanto, foi treinado as arquiteturas em validação cruzada *k*-fold ( $k = 10$ ).

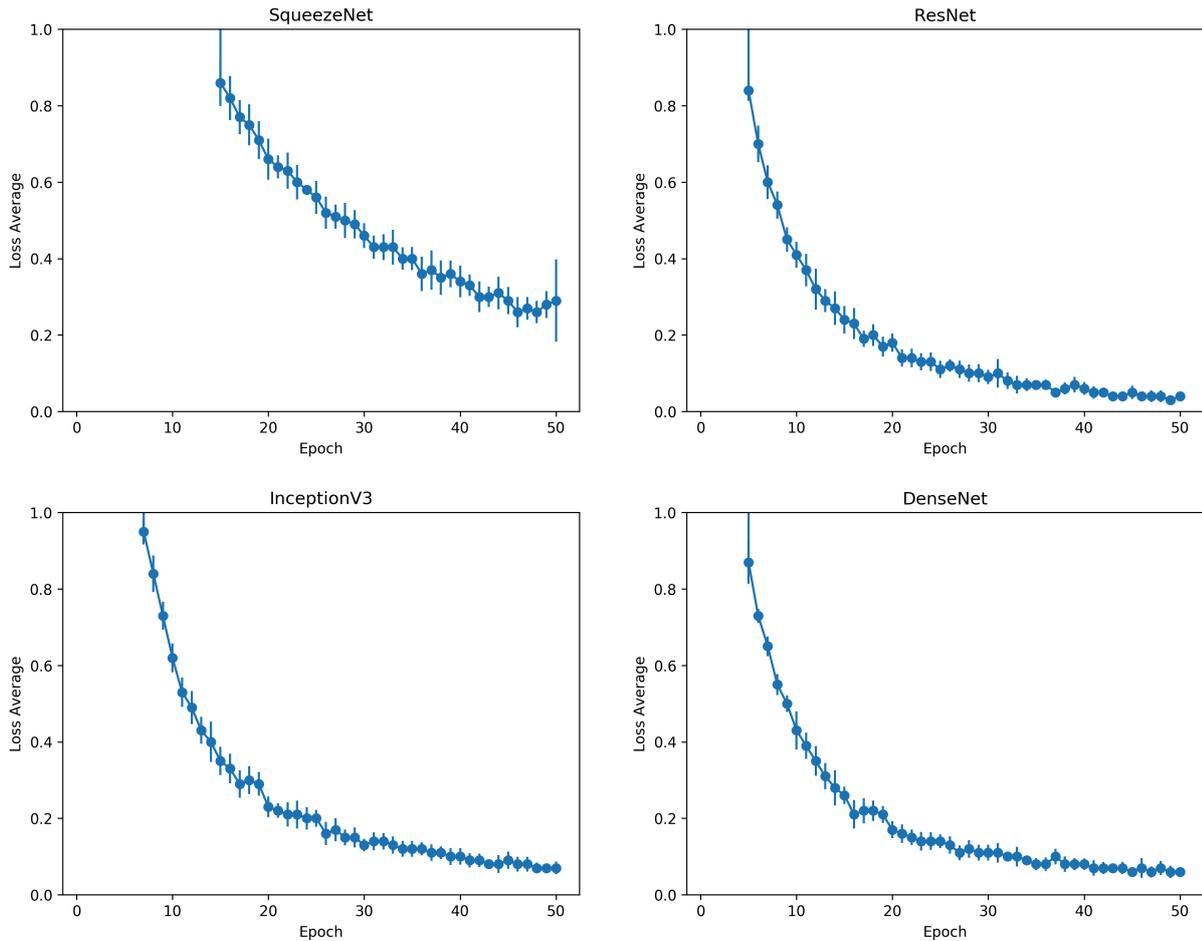


Figura 32 – Gráficos da perda média e desvio padrão dos *folders* em cada época para cada arquitetura.

A Figura 34 mostra a evolução da acurácia de cada arquitetura de CNN entre os diferentes *folders*. Note que estes resultados se referem apenas ao *fold* usado como conjunto de teste. Este resultado mostra uma pequena variação entre os *folders* independente da arquitetura de CNN. Isto é natural e esperado, já que diferentes espécies de madeira apresentam estruturas anatômicas semelhantes. Entretanto, o fato de existirem subconjuntos de imagens que são mais difíceis de classificar que outros, não compromete o desempenho da CNN, o qual ainda apresenta alta acurácia e pequeno desvio padrão. É importante ressaltar que esta variação é maior para algumas CNNs, como SqueezeNet e InceptionV3, sendo a última a que apresenta a maior variação de acurácia entre cada *fold*, indicando que as características aprendidas pela CNN podem se diferenciar entre os *folders*.

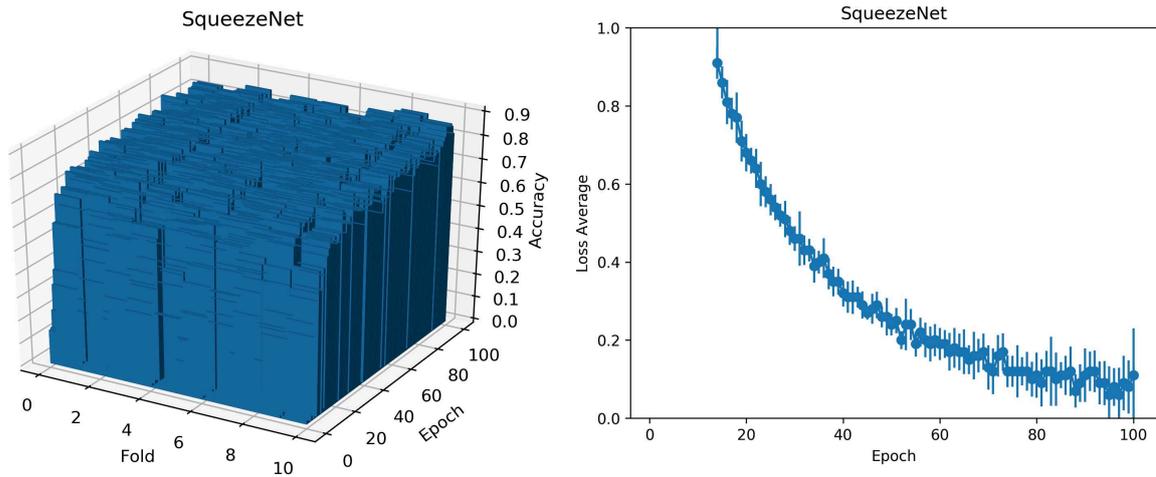


Figura 33 – Gráficos da perda média e desvio padrão dos *fold*s em cada época da arquitetura SqueezeNet usando 100 épocas.

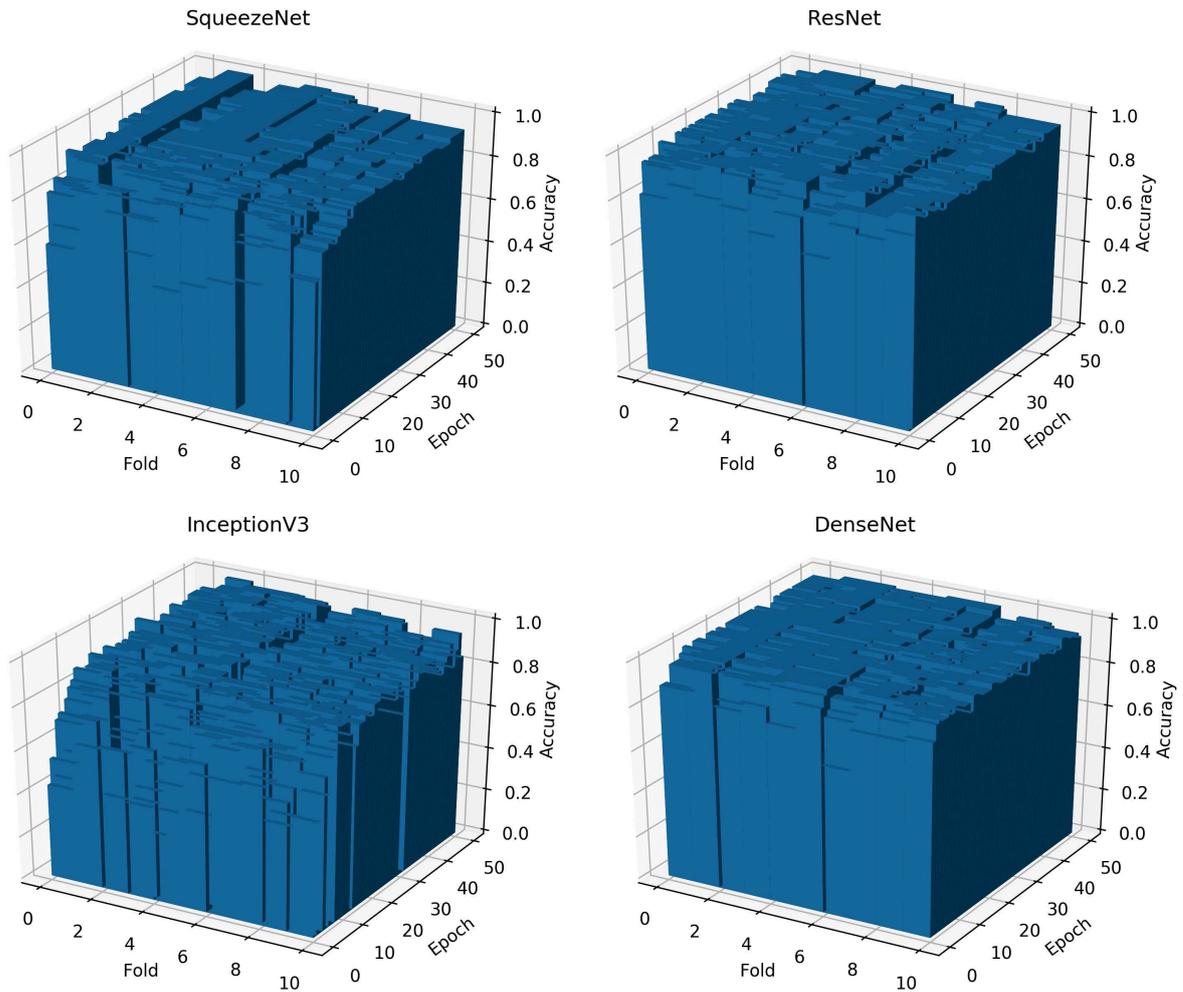


Figura 34 – Gráficos de barra da acurácia sobre épocas para cada arquitetura.

Estas variações ilustradas na Figura 34 indicam a presença de diferenças entre alguns conjuntos de treino e teste, por exemplo, a combinação de alguns *folds* resulta em um conjunto de treino que não contempla todas as características presentes no conjunto de teste. Outra possível explicação está no fato que esta base de imagens apresenta baixa similaridade intra-classe (Figura 24) e algum nível de similaridade inter-classe dado que as amostras de diferentes espécies apresentam estruturas semelhantes.

Embora alguns *folds* são mais desafiadores para treinar, como mostra a Figura 34, a alta acurácia média e pequeno desvio padrão demonstra que os mapas de características aprendidos pelas arquiteturas de CNN são robustos (veja Tabela 13). A arquitetura InceptionV3 apresenta a menor acurácia (93,59%) e a maior variância (desvio padrão de 3,84), enquanto a maior acurácia foi obtida pela DenseNet (98,13%).

Tabela 13 – Métodos de características pré-projetados comparados a aprendizado de características.

	Método/Arquitetura	Acurácia
Características pré-projetadas	Primeira ordem	43,86% ± 7,88
	Fourier	34,32% ± 6,81
	Gabor	81,59% ± 9,37
	Wavelet descriptors	63,41% ± 8,47
	Haralick	57,50% ± 11,93
	Tamura	40,00% ± 8,31
	DF+LBP	74,77% ± 7,30
	Turista	47,73% ± 6,06
	LBP <sub>r=1</sub>	55,23% ± 9,47
	LBP <sub>r={1,2,3}</sub>	80,68% ± 6,63
	PSO (Combinação)	<b>90,23% ± 5,14</b>
Aprendizado de características	InceptionV3	93,59% ± 3,84
	DenseNet	<b>98,13% ± 1,96</b>
	ResNet	97,31% ± 1,85
	SqueezeNet	96,36% ± 2,53

Para aprimorar a análise, os resultados das arquiteturas de CNN foram comparados com métodos promissores de características pré-projetadas encontradas na literatura. Os resultados são dispostos na Tabela 13. É importante enfatizar que, mesmo tendo informação de cor nas amostras de madeira, os métodos analisados trabalham apenas com imagens em escala de cinza, levando em consideração apenas sua luminância no cálculo.

É possível perceber que os métodos de características pré-projetadas selecionados englobam várias abordagens para extrair textura, como histogramas (exemplo, Primeira ordem), espectrais (exemplo, Fourier e Gabor), estatísticos (exemplo, Haralick e Turista) e baseados em análise fractal (DF+LBP). Esta variedade de métodos resulta em uma

larga variedade de acurácia, variando de 34,32% (Fourier) até 81,59% (Gabor). Entretanto, todas as arquiteturas de CNN apresentam resultados que superam mesmo o melhor dos métodos de características pré-projetadas, sendo o resultado mais baixo das CNNs (InceptionV3) 12% superior ao melhor método de características projetadas. Além de apresentar menor desempenho, estes descritores também mostram uma grande variação de resultados entre os folds, sendo o mínimo e máximo, respectivamente, 6,81 (Fourier) and 11,93 (Haralick). Estes desvios padrão são maiores que a máxima variação encontrada nas CNNs (3,84). Por outro lado, é interessante notar que o descritor Fourier apresenta a menor taxa de acerto, e o menor desvio padrão, indicando que ele classifica mal as amostras independente dos conjuntos de treino e teste. Adicionalmente, os filtros Gabor apresentam a maior acurácia, mas também o terceiro maior desvio padrão entre os métodos de características pré-projetados, isto é, mesmo apresentando uma alta acurácia, estes resultados são muito dependentes nos conjuntos de treino e teste do que nos outros métodos.

Todas as arquiteturas de CNN são construídas utilizando filtros que atuam como métodos de extração de características. Pela combinação de sequências de filtros convolucionais, uma CNN é capaz de extrair características em diferentes níveis semânticos. Portanto, para realizar uma comparação justa, os métodos de características pré-projetadas foram combinados em um único vetor. Entretanto, como são analisados 10 métodos, seria muito oneroso avaliar todas as possíveis combinações de métodos ( $2^{10}$  combinações). Ao invés disso, foi utilizado o algoritmo PSO (do inglês, *Particle Swarm Optimization*) (KENNEDY; EBERHART, 1995) para determinar a melhor combinação dos métodos. De forma simplificada, no algoritmo PSO, o problema é representado por uma partícula na posição  $[X(1), X(2), \dots, X(D)]$ , onde  $D = 10$  é o número de dimensões (métodos de características) no espaço de busca, com  $0 \leq X(i) \leq 1$ . É considerado que um método de característica  $i\#$  é selecionado pelo PSO para compor a combinação de métodos quando  $X(i) \geq 0,5$ .

Os resultados mostram que a combinação de características melhora consideravelmente a taxa de acerto na base de imagens. (Tabela 13). Através da combinação de métodos, atingiu-se uma acurácia de 90,23%, um aumento de 8,64% quando comparado ao melhor método sozinho (81,59%). Esta combinação de métodos apresentou o menor desvio padrão, indicando desempenho similar entre os folds. No entanto, estes resultados ainda são 3,36% inferiores ao pior resultado atingido pelas CNNs (InceptionV3).

Adicionalmente, é apresentado na Figura 35 a acurácia média por espécie entre os folds. DenseNet tem melhor desempenho que o PSO em todas as espécies, tendo como pior resultado 90% na espécie *Dipteryx odorata* (ID # 6). Como comparação, o pior resultado (77,50%) do PSO foi obtido na espécie *Handroanthus incanus* (ID # 8). Coincidentemente, esta é a espécie de madeira onde a DenseNet e o PSO apresentam a maior diferença em desempenho (22,50%). Isto corrobora para a superioridade em eficiência das CNNs e seu

uso como uma abordagem adequada ao problema de classificação de madeira.

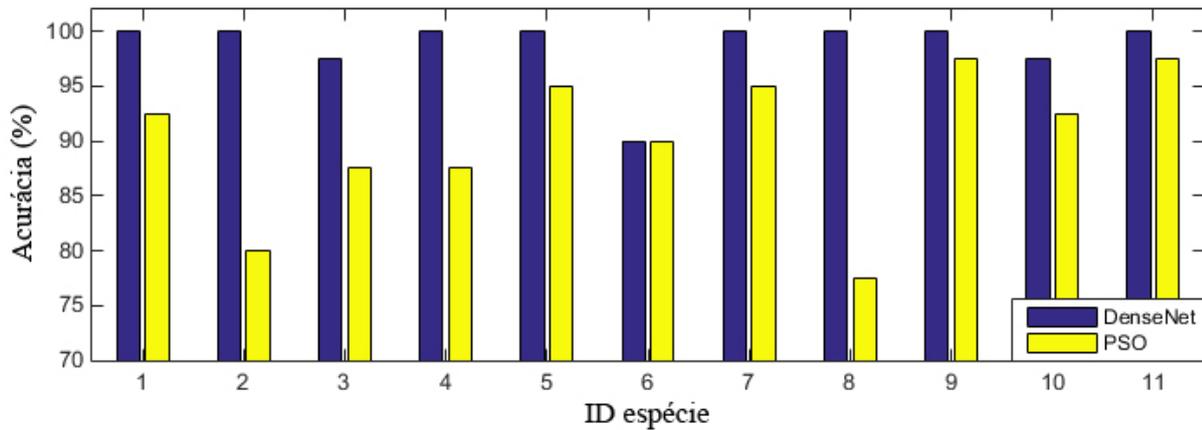


Figura 35 – Acurácia média por espécie de madeira.

Outro fator analisado, foi uma busca de correlações entre o aprendizado de características realizado pelas CNNs e as características observadas por um especialista em anatomia da madeira. Para tanto, foram utilizados mapas de calor gerados pelo método Grad-CAM++ (CHATTOPADHAY et al., 2018), que utiliza informações do gradiente e sua interação entre as camadas da CNN para produzir as regiões de maior atenção na imagem.

A Figura 36 mostra alguns exemplo de amostras de madeira e seus respectivos mapas de calor gerados pelo método Grad-CAM++ utilizando as arquiteturas SqueezeNet e DenseNet. Enquanto a SqueezeNet mostra mapas de calor concentrados em regiões específicas da imagem, como vasos, a DenseNet apresenta mapas de calor mais difusos, avaliando o aspecto global da amostra. Essa diferença nos mapas de calor é esperada uma vez que essas CNNs apresentam estruturas diferentes, como o uso de informações residuais pela DenseNet e o uso de módulos de compressão e expansão pela SqueezeNet.

Devido ao seu tamanho reduzido (isto é, menor número de parâmetros treináveis), a SqueezeNet se concentra em regiões específicas da imagem que sejam mais discriminantes para aquela classe em particular. Por esse motivo, optamos por fazer as análises seguintes com base nos mapas gerados por essa rede, mesmo ela não apresentando a maior taxa de acertos dentre as CNNs avaliadas.

As espécies *Alantoma decandra* e *Cariniana micrantha* são pertencentes à mesma família (*Lecytidaceae*) e apresentam grande similaridade anatômica (veja Figura 37), tornando sua identificação um desafio até mesmo para anatomistas experientes. Uma das estruturas mais importantes para a determinação destas espécies é o parênquima em axial em faixas do tipo reticulado, que se distribui de forma uniforme por toda a extensão da superfície transversal dessa madeira. Entretanto, essa característica abrange um grupo muito grande de espécies, que se torna insuficiente para realizar sua distinção.

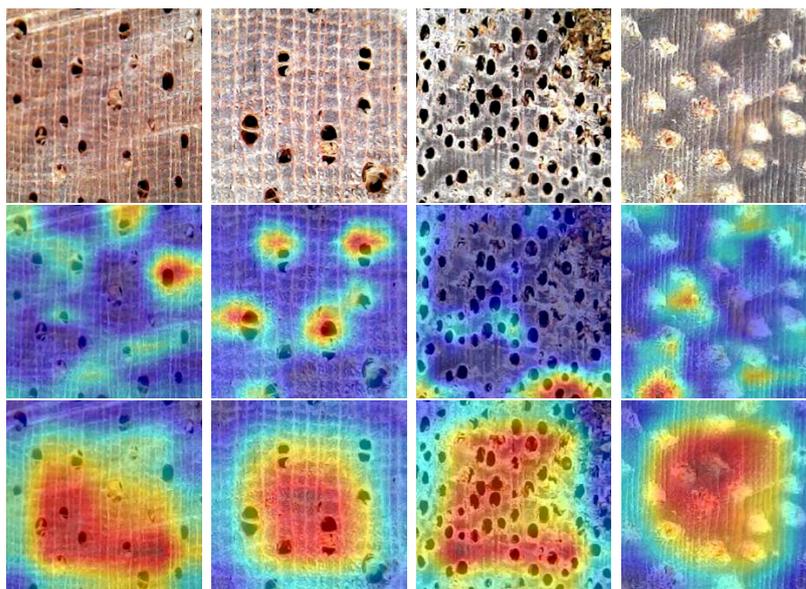


Figura 36 – Exemplos de amostras de madeira e seus respectivos mapas de calor gerados pelo método Grad-CAM++ utilizando as arquiteturas SqueezeNet (segunda linha) e DenseNet (terceira linha).

Nesse sentido, o mapa de calor indica tendência a ignorar tal estrutura e concentra as zonas de calor nos vasos, em especial nos vasos desobstruídos. É interessante notar que os vasos são estruturas importantes como caracteres taxonômicos, onde suas variações de tamanho, frequência e distribuição trazem informações importantes para a determinação das espécies, porém, definir tais parâmetros com uma simples observação a olho nu é extremamente difícil, muitas vezes exigindo a utilização de equipamentos de microscopia adequados para essa mensuração.

Nas espécies *Caraipa densifolia*, *Caryocar villosum*, *Dipteryx odorata* e *Lueheopsis duckeana* as áreas de calor também se sobrepõem preferencialmente nos vasos, como nas espécies levantadas anteriormente (Figura 37). Porém, na espécie *Caraipa densifolia* elas aparentam levar em consideração o arranjo desses vasos entre si, apresentando sobreposição com uma característica anatômica importante para essa espécie denominada “distribuição em cadeias diagonais”.

Já na espécie *Caryocar villosum* as áreas em destaque se concentram nos *clusters* de vasos com obstruções. Tais estruturas recebem a denominação anatômica de “vasos agrupados em múltiplos radiais” e são fundamentais como característica taxonômica. De forma semelhante, na espécie *Dipteryx odorata*, as áreas de maior calor se concentram principalmente nos vasos obstruídos. Interessante notar que a obstrução dos vasos nessa espécie se dá pelo acúmulo de extrativos bastante comuns nesse grupo taxonômico, sendo uma de suas características principais.

Diferente das espécies anteriores, na espécie *Clarisia racemosa* (veja Figura 38) os mapas de calor se concentram principalmente ao parênquima, que nessa espécie apresenta

linhas distribuídas de maneira não uniforme pela superfície observada, apresentando desenhos bastante variantes. Tais faixas são formadas por parênquima paratraqueal (relacionado aos vasos) que se estendem lateralmente e se conectam. Tais estruturas recebem o nome de “Paratraqueal aliforme linear de extensão longa”.

Na espécie *Handroanthus incanus* (Figura 38) há uma correlação semelhante a observada na espécie *Dipterix odorata*, o que é esperado por serem da mesma família e muito parecidas entre si. Já na espécie *Osteophloeum platyspermum*, as zonas de calor se sobrepõem nos vasos agrupados em múltiplos radiais, que representam uma característica taxonômica bastante importante para seu reconhecimento. Por fim, na espécie *Goupia glabra*, embora não haja correlação com as características observadas por um anatomista, percebe-se que as zonas de calor se localizam entre os vasos, que visualmente possuem tonalidade bem mais marcante que as demais espécies.

## 4.6 Considerações finais

Este capítulo apresentou os experimentos realizados e seus resultados. Através dele, foi possível constatar a superioridade em termos de acurácia do aprendizado automático de características por meio das CNN sobre os métodos de características pré-projetadas. Mais especificamente, em todas as bases de imagens avaliadas, a arquitetura DenseNet apresentou melhor desempenho. Além disso, foi possível avaliar os aspectos mais específicos na aplicação de classificação de madeiras, como a influência do corte no tronco no desempenho dos métodos computacionais; a influência da simplificação do protocolo de aquisição das imagens; e a correlação entre as características aprendidas através de uma CNN e um especialista em anatomia da madeira.

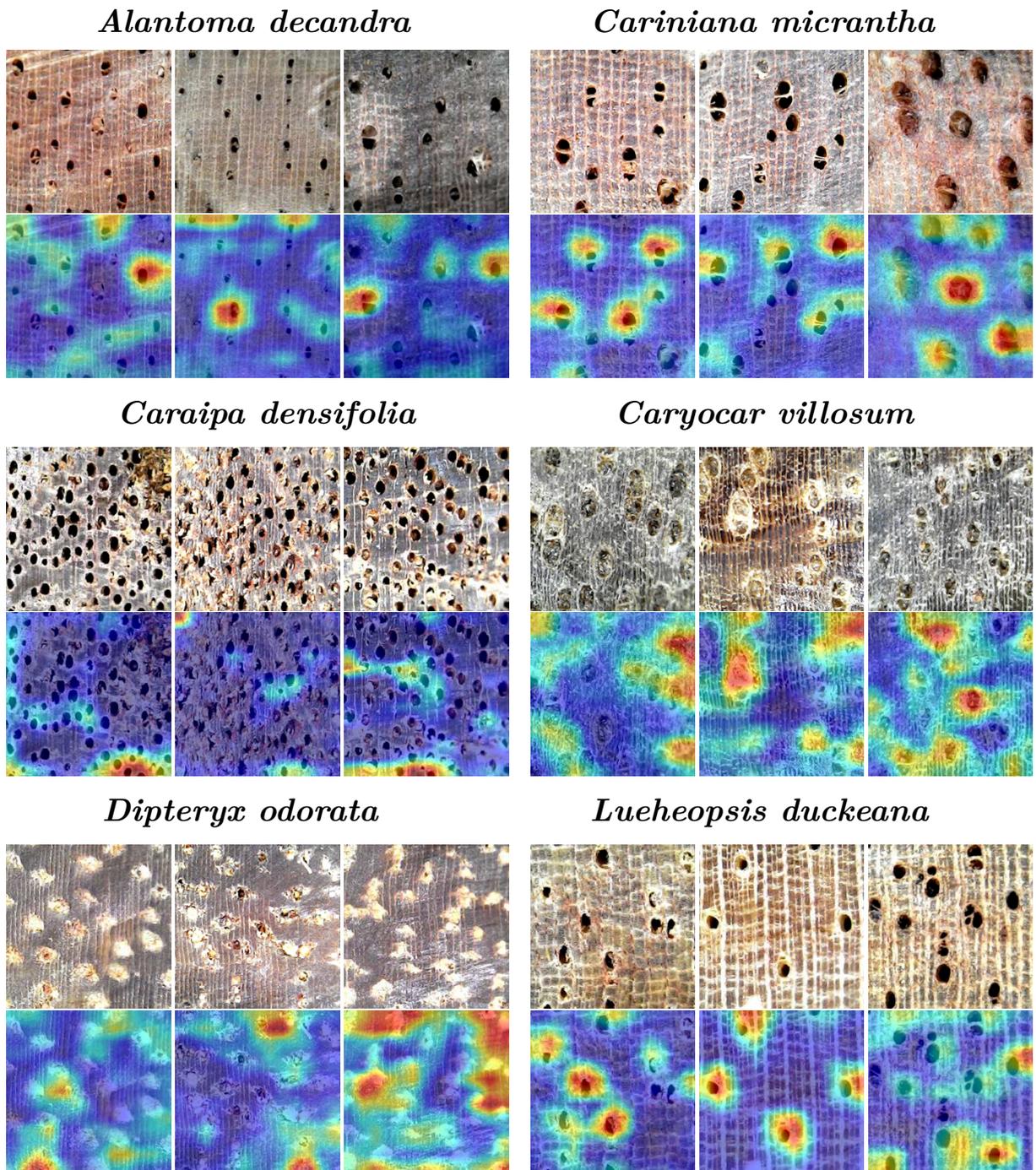


Figura 37 – Amostras das espécies de ID 1 (*Alantoma decandra*), ID 2 (*Caraipa densifolia*), ID 3 (*Cariniana micrantha*), ID 4 (*Caryocar villosum*), ID 6 (*Dipteryx odorata*) e ID 9 (*Lueheopsis duckeana*); e abaixo de cada uma a sobreposição do mapa de calor gerado pelo método Grad-CAM++.

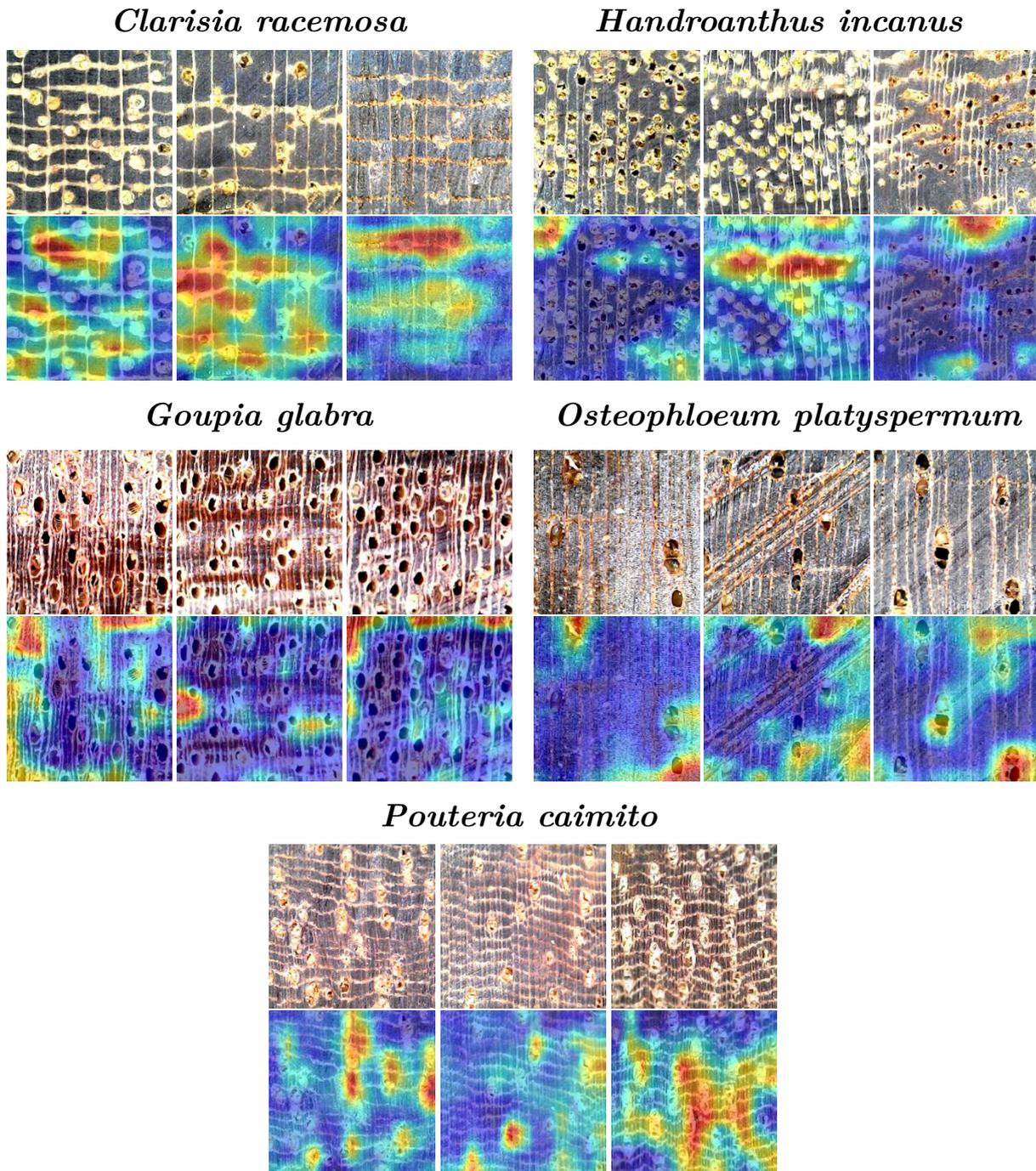


Figura 38 – Amostras das espécies de ID 5 (*Clarisia racemosa*), ID 7 (*Goupia glabra*), ID 8 (*Handroanthus incanus*), ID 10 (*Osteophloeum platyspermum*) e ID 11 (*Pouteria caimito*); e abaixo de cada uma a sobreposição do mapa de calor gerado pelo método Grad-CAM++.

---

## Conclusão

Encontrar uma solução para a classificação automática de imagens de pólen, vírus e madeira implica em benefícios para diversas áreas e aplicações. A extração e comercialização ilegal de madeiras é uma prática frequente e difícil de ser combatida, já que o reconhecimento visual das espécies é realizado por profissionais que são escassos e não conseguem cobrir todo o território nacional. De maneira semelhante, estudos palinológicos e virológicos que dependem de reconhecimento visual, também dependem de profissionais altamente qualificados para realizá-los. Tais estudos têm se inserido em diversos escopos, contribuindo para investigações criminais, estudos climáticos, certificando a produção de mel, estudos da interação entre células e vírus, e na criação de medicamentos.

Como resultado desse trabalho foram introduzidas duas novas bases de imagens microscópicas de pólen e madeira, contendo 134 e 281 classes, respectivamente. Ambas as bases possuem um número maior de classes do que as bases levantadas na revisão bibliográfica. Por meio delas, foram realizados estudos comparativos de classificação de imagens entre métodos pré-projetados de extração de características e métodos de aprendizado automático de características, mais especificamente, as CNNs. Além disso, foram realizados experimentos para avaliar como as diferentes arquiteturas de CNNs se comportam sob diferentes conjuntos de treinamento e teste. Para tanto, foi selecionada uma base de imagens de vírus, o qual se mostra desafiadora devido à sua alta variabilidade intra-classe e baixa acurácia reportada em estudos anteriores.

Os resultados experimentais apontam que o aprendizado automático de características tem melhor desempenho em acurácia que características pré-projetadas em todas as bases de imagens. A saber, a arquitetura de CNN DenseNet apresentou os melhores resultados, inclusive quando comparada a outras arquiteturas. Verifica-se que, mesmo com elevado número de classes das bases de imagens, as metodologias propostas são capazes de atingir acurácia satisfatória para sua classificação. Os melhores resultados obtidos atingiram 96,24% na classificação de pólen e 98,75% na classificação de madeira, considerando avaliação no corte transversal do tronco.

Por meio do protocolo simplificado de aquisição das imagens de madeira, foi possí-

vel validar a utilização da metodologia computacional em campo, evitando a realização dos procedimentos tradicionais de laboratório para a preparação das amostras do tronco. Além disso, um novo dispositivo portátil foi construído com propósito de padronizar a aquisição das imagens e possibilitar colaboração entre centros de pesquisa. Os resultados experimentais revelaram a superioridade das arquiteturas de CNN na base de imagens de 11 espécies construída através do protocolo. A arquitetura DenseNet atingiu a melhor acurácia de 98,13%, superando em quase 8% (90,23%) o melhor resultado de características pré-projetadas obtido pela combinação do método PSO.

## 5.1 Contribuições

As contribuições oriundas desse estudo podem ser sumarizadas da seguinte maneira:

- ❑ Revisão bibliográfica dos estudos de classificação de imagens de pólen, vírus e madeira;
- ❑ Reprodução e desenvolvimento de metodologias utilizando características pré-projetadas;
- ❑ Análise da influência da inicialização aleatória de pesos e transferência de aprendizado em CNNs na classificação de pólen. EUSIPCO 2019 (GEUS et al., 2019);
- ❑ Análise de diferentes cortes no tronco de árvores comparando a classificação por meio de características pré-projetadas e transferência de aprendizado em CNNs. Multimedia Tools and Applications (GEUS et al., 2020);
- ❑ Análise da variabilidade de CNNs na seleção dos conjuntos de treino e teste aplicados à classificação de imagens de vírus. VISAPP 2020 (GEUS; BACKES; SOUZA, 2020);
- ❑ Simplificação do protocolo de aquisição de imagens de madeira comparando sua classificação por meio de características pré-projetadas e transferência de aprendizado em CNNs. Wood Science and Technology (GEUS et al., 2021).

## 5.2 Trabalhos futuros

Pretende-se expandir as bases de imagens das três aplicações e avaliar o desempenho das metodologias em um maior número de espécies/tipo. Destaca-se também pontos específicos para as aplicações:

- ❑ Escrita de um *survey* dos trabalhos de reconhecimento computacional de madeira para evidenciar as metodologias de aquisição de imagens e seus desempenhos;

- Aprimorar o dispositivo portátil desenvolvido, para eliminar a necessidade de sua conexão a um celular ou computador pessoal;
- Desenvolvimento de métodos automáticos de segmentação dos grãos de pólen em imagens de lâminas microscópicas completas.

## 5.3 Contribuições em Produção Bibliográfica

A seguir estão listadas as publicações desenvolvidas durante o período de doutorado:

- GEUS, A. R.; BATISTA, M. A.; RABELO, M. N.; BARCELOS, C. Z.; DA SILVA, S. F.. Maize Insects Classification Through Endoscopic Video Analysis. Lecture Notes in Computer Science: Springer International Publishing, 2019. v. 11489, p. 251-262. Proceedings of CAIAC 2019.
- GEUS, A. R.; BARCELOS, C. Z.; BATISTA, M. A.; DA SILVA, S. F.. Large-scale Pollen Recognition with Deep Learning. Proceedings of EUSIPCO 2019 - 27th European Signal Processing Conference, 2019.
- KITANO, B. T.; MENDES, C. T.; GEUS, A. R.; OLIVEIRA, H. C.; SOUZA, J. R.. Corn Plant Counting Using Deep Learning and UAV Images. IEEE Geoscience and Remote Sensing Letters, 2019.
- GEUS, A. R.; BACKES, A. R.; SOUZA, J. R.. Variability evaluation of CNNs using cross-validation on Viruses images. Proceedings of VISAPP 2020 - 15th International Conference on Computer Vision Theory and Applications, 2020.
- SPOLTI A.; GUIZILINI V. C.; MENDES, C. T.; CROCE M. D.; GEUS, A. R.; OLIVEIRA H. C.; BACKES, A. R.; SOUZA, J. R.. Application of U-Net and Auto-Encoder to the Road/Non-road Classification of Aerial Imagery in Urban Environments. Proceedings VISAPP 2020 - 15th International Conference on Computer Vision Theory and Applications, 2020.
- DA SILVA, C. J. P.; SOUZA, J. R.; PESSIN, G.; GEUS, A. R.. Estimativa do tamanho de bolhas em espuma de flotação utilizando Visão Computacional. Proceedings of XXIII Congresso Brasileiro de Automática, 2020.
- GEUS, A. R.; DA SILVA, S. F.; GONTIJO, A. B.; SILVA, F. O.; BATISTA, M. A.; SOUZA, J. R.. An analysis of timber sections and deep learning for wood species classification. Multimedia Tools and Applications: Springer International Publishing, 2020.

- GEUS, A. R.; BACKES, A. R.; GONTIJO, A. B.; ALBUQUERQUE G. H. Q.; SOUZA, J. R.. Amazon wood specie classification: a comparison between deep learning and pre-designed features. Wood Science and Technology: Springer International Publishing, 2021.

---

## Referências

- ALJOHAR, H. I. et al. Physical and chemical screening of honey samples available in the Saudi market: An important aspect in the authentication process and quality assessment. **Saudi Pharmaceutical Journal**, 2018. <<https://doi.org/10.1016/j.jsps.2018.04.013>>.
- ALLEN, G. P. et al. Machine vision for automated optical recognition and classification of pollen grains or other singulated microscopic objects. In: **International Conference on Mechatronics and Machine Vision in Practice (M2VIP)**. [S.l.: s.n.], 2008. <<https://doi.org/10.1109/MMVIP.2008.4749537>>.
- AUGUSTYN, A. et al. Palynology. **Encyclopedia Britannica**, 2017.
- AYADI, N. et al. Color stability of heat-treated wood during artificial weathering. **European Journal of Wood and Wood Products**, 2003. <<https://doi.org/10.1007/s00107-003-0389-2>>.
- BACKES, A. R.; JUNIOR, J. J. de M. S. Virus classification by using a fusion of texture analysis methods. In: **International Conference on Systems, Signals and Image Processing (IWSSIP)**. [S.l.: s.n.], 2020. <<https://doi.org/10.1109/IWSSIP48289.2020.9145325>>.
- BANERJEE, N.; MUKHOPADHYAY, S. Viral glycoproteins: biological role and application in diagnosis. **VirusDisease**, 2016. <<https://doi.org/10.1007/s13337-015-0293-5>>.
- BAO, W.; YUE, J.; RAO, Y. A deep learning framework for financial time series using stacked autoencoders and long-short term memory. **PLOS ONE**, 2017. <<https://doi.org/10.1371/journal.pone.0180944>>.
- BARMPOUTIS, P. et al. Wood species recognition through multidimensional texture analysis. **Computers and Electronics in Agriculture**, 2018. <<https://doi.org/10.1016/j.compag.2017.12.011>>.
- BENVENUTO, D. et al. The 2019-new coronavirus epidemic: evidence for virus evolution. **Journal of Medical Virology**, 2020. <<https://doi.org/10.1002/jmv.25688>>.
- BOSMAN, A. S.; ENGELBRECHT, A.; HELBIG, M. Visualising basins of attraction for the cross-entropy and the squared error neural network loss functions. **Neurocomputing**, 2020. <<https://doi.org/10.1016/j.neucom.2020.02.113>>.

- CHATTOPADHAY, A. et al. Grad-cam++: Generalized gradient-based visual explanations for deep convolutional networks. In: **IEEE Winter Conference on Applications of Computer Vision (WACV)**. [S.l.: s.n.], 2018. <<https://doi.org/10.1109/WACV.2018.00097>>.
- CIRESAN, D. C. et al. Deep, big, simple neural nets for handwritten digit recognition. **Neural Computation**, 2010. <[https://doi.org/10.1162/NECO\\_a\\_00052](https://doi.org/10.1162/NECO_a_00052)>.
- COSTA, R. M. C. J. Luciano da F. **Shape Classification and Analysis - Theory and Practice**. [S.l.]: CRC Press, 2009. (Image Processing Series).
- DADI, H. S.; PILLUTLA, G. K. M. Improved face recognition rate using hog features and svm classifier. **Journal of Electronics and Communication Engineering**, 2016. <<http://doi.org/10.9790/2834-1104013444>>.
- DALAL, N.; TRIGGS, B. Histograms of oriented gradients for human detection. In: **Conference on Computer Vision and Pattern Recognition (CVPR)**. [S.l.: s.n.], 2005. <<https://doi.org/10.1109/CVPR.2005.177>>.
- D'AMATO, G. et al. Allergenic pollen and pollen allergy in europe. **Allergy**, 2007. <<https://doi.org/10.1111/j.1398-9995.2007.01393.x>>.
- DAOOD, A.; RIBEIRO, E.; BUSH, M. Pollen grain recognition using deep learning. In: **International Symposium on Visual Computing (ISVC)**. [S.l.: s.n.], 2016. <[https://doi.org/10.1007/978-3-319-50835-1\\_30](https://doi.org/10.1007/978-3-319-50835-1_30)>.
- DAUGMAN, J. G. Two-dimensional spectral analysis of cortical receptive field profiles. **Vision research**, 1980. <[https://doi.org/10.1016/0042-6989\(80\)90065-6](https://doi.org/10.1016/0042-6989(80)90065-6)>.
- \_\_\_\_\_. Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. **Journal of the Optical Society of America A**, 1985. <<https://doi.org/10.1364/JOSAA.2.001160>>.
- DONAHUE, J. et al. Decaf: A deep convolutional activation feature for generic visual recognition. **International Conference on Machine Learning (ICML)**, 2014.
- FRANKE, B.; QUENNEVILLE, P. Numerical modeling of the failure behavior of dowel connections in wood. **Journal of Engineering Mechanics**, 2011. <[https://doi.org/10.1061/\(ASCE\)EM.1943-7889.0000217](https://doi.org/10.1061/(ASCE)EM.1943-7889.0000217)>.
- FUKUSHIMA, K. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. **Biological Cybernetics**, 1980. <<https://doi.org/10.1007/BF00344251>>.
- GEUS, A. R. et al. Amazon wood specie classification: a comparison between deep learning and pre-designed features. **Wood Science and Technology**, 2021. <<https://doi.org/10.1007/s00226-021-01282-w>>.
- GEUS, A. R.; BACKES, A. R.; SOUZA, J. R. Variability evaluation of cnns using cross-validation on viruses images. In: **International Conference on Computer Vision Theory and Applications (VISAPP)**. [S.l.: s.n.], 2020. <<http://doi.org/10.5220/0009352106260632>>.

- GEUS, A. R. et al. Large-scale pollen recognition with deep learning. In: **European Signal Processing Conference (EUSIPCO)**. [S.l.: s.n.], 2019. <<https://doi.org/10.23919/EUSIPCO.2019.8902735>>.
- \_\_\_\_\_. An analysis of timber sections and deep learning for wood species classification. **Multimedia Tools and Applications**, 2020. <<https://doi.org/10.1007/s11042-020-09212-x>>.
- GIRSHICK, R. et al. Rich feature hierarchies for accurate object detection and semantic segmentation. In: **Conference on Computer Vision and Pattern Recognition (CVPR)**. [S.l.: s.n.], 2014. <<https://doi.org/10.1109/CVPR.2014.81>>.
- GLOROT, X.; BORDES, A.; BENGIO, Y. Deep sparse rectifier neural networks. In: **International Conference on Artificial Intelligence and Statistics (AISTATS)**. [S.l.: s.n.], 2011.
- GONCALVES, A. B. et al. Feature extraction and machine learning for the classification of brazilian savannah pollen grains. **PLOS ONE**, 2016. <<https://doi.org/10.1371/journal.pone.0157044>>.
- GOODFELLOW, I. J.; BENGIO, Y.; COURVILLE, A. **Deep Learning**. [S.l.]: MIT Press, 2016.
- HAFEMANN, L. G.; OLIVEIRA, L. S.; CAVALIN, P. Forest species recognition using deep convolutional neural networks. In: **International Conference on Pattern Recognition (ICPR)**. [S.l.: s.n.], 2014. <<https://doi.org/10.1109/ICPR.2014.199>>.
- HE, K. et al. Deep residual learning for image recognition. **Conference on Computer Vision and Pattern Recognition**, 2016. <<https://doi.org/10.1109/CVPR.2016.90>>.
- HODGSON, R. M. et al. Progress towards a system for the automatic recognition of pollen using light microscope images. In: **International Symposium on Image and Signal Processing and Analysis (ISPA)**. [S.l.: s.n.], 2005. <<https://doi.org/10.1109/ISPA.2005.195387>>.
- HOWARD, A. G. et al. Mobilenets: Efficient convolutional neural networks for mobile vision applications. **arXiv**, 2017.
- HU, S.; LI, K.; BAO, X. Wood species recognition based on sift keypoint histogram. In: **International Congress on Image and Signal Processing (CISP)**. [S.l.: s.n.], 2015. <<https://doi.org/10.1109/CISP.2015.7407968>>.
- HUANG, G. et al. Densely connected convolutional networks. In: **Conference on Computer Vision and Pattern Recognition (CVPR)**. [S.l.: s.n.], 2017. <<https://doi.org/10.1109/CVPR.2017.243>>.
- IANDOLA, F. N. et al. Squeezenet: Alexnet-level accuracy with 50x fewer parameters and <0.5mb model size. **arXiv**, 2016.
- IBRAHIM, I. et al. Statistical feature extraction method for wood species recognition system. **European Journal of Wood and Wood Products**, 2017. <<https://doi.org/10.1007/s00107-017-1163-1>>.

- \_\_\_\_\_. Tree species recognition system based on macroscopic image analysis. **Wood Science and Technology**, 2016. <<https://doi.org/10.1007/s00226-016-0859-4>>.
- ITO, E. et al. Virus particle detection by convolutional neural network in transmission electron microscopy images. **Food and Environmental Virology**, 2018. <<https://doi.org/10.1007/s12560-018-9335-7>>.
- JABID, T.; KABIR, M. H.; CHAE, O. Local directional pattern (ldp) – a robust image descriptor for object recognition. In: **IEEE International Conference on Advanced Video and Signal Based Surveillance**. [S.l.: s.n.], 2010. <<https://doi.org/10.1109/AVSS.2010.17>>.
- JAYAWICKRAMA, K. et al. Date of earlywood-latewood transition in provenances and families of loblolly pine, and its relationship to growth phenology and juvenile wood specific gravity. **Canadian Journal of Forest Research**, 1997. <<https://doi.org/10.1139/x97-091>>.
- JORDAN, R. et al. Classification of wood species by neural network analysis of ultrasonic signals. **Ultrasonics**, 1998. <[https://doi.org/10.1016/S0041-624X\(97\)00148-0](https://doi.org/10.1016/S0041-624X(97)00148-0)>.
- JUNIOR, J. J. de M. S.; BACKES, A. R. Elm based signature for texture classification. **Pattern Recognition**, 2016. <<https://doi.org/10.1016/j.patcog.2015.09.014>>.
- KENNEDY, J.; EBERHART, R. Particle swarm optimization. In: **International Conference on Neural Networks (ICNN)**. [S.l.: s.n.], 1995. <<https://doi.org/10.1109/ICNN.1995.488968>>.
- KHALID, M.; YUSOF, R.; KHAIRUDDIN, A. S. M. Tropical wood species recognition system based on multi-feature extractors and classifiers. In: **International Conference on Instrumentation Control and Automation**. [S.l.: s.n.], 2011. <<https://doi.org/10.1109/ICA.2011.6130117>>.
- KITANO, B. T. et al. Corn plant counting using deep learning and uav images. **IEEE Geoscience and Remote Sensing Letters**, 2019. <<https://doi.org/10.1109/LGRS.2019.2930549>>.
- KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. In: **Conference on Neural Information Processing Systems (NIPS)**. [S.l.: s.n.], 2012.
- KWON, O. et al. Automatic wood species identification of korean softwood based on convolutional neural networks. **Journal of the Korean Wood Science and Technology**, 2017. <<http://dx.doi.org/10.5658/WOOD.2017.45.6.797>>.
- KYLBERG, G.; UPPSTRÖM, M.; SINTORN, I.-M. Virus texture analysis using local binary patterns and radial density profiles. In: **Iberoamerican Congress on Pattern Recognition (CIARP)**. [S.l.: s.n.], 2011. <[https://doi.org/10.1007/978-3-642-25085-9\\_68](https://doi.org/10.1007/978-3-642-25085-9_68)>.
- KYLBERG, G. et al. Segmentation of virus particle candidates in transmission electron microscopy images. **Journal of microscopy**, 2011. <<https://doi.org/10.1111/j.1365-2818.2011.03556.x>>.

- LAKSHMANAPRABU, S. K. et al. Optimal deep learning model for classification of lung cancer on ct images. **Future Generation Computer Systems**, 2019. <<https://doi.org/10.1016/j.future.2018.10.009>>.
- LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. **Nature**, 2015. <<https://doi.org/10.1038/nature14539>>.
- LEE, C.-J.; WANG, S.-D. Fingerprint feature reduction by principal gabor basis function. **Pattern Recognition**, 2001. <[https://doi.org/10.1016/S0031-3203\(01\)00029-2](https://doi.org/10.1016/S0031-3203(01)00029-2)>.
- LI, P.; FLENLEY, J. R. Pollen texture identification using neural networks. **Grana**, Taylor & Francis, 1999. <<https://doi.org/10.1080/001731300750044717>>.
- LI, P. et al. Towards automation of palynology 2: the use of texture measures and neural network analysis for automated identification of optical images of pollen grains. **Journal of Quaternary Science**, 2004. <<https://doi.org/10.1002/jqs.874>>.
- LLORCA, D. F.; ARROYO, R.; SOTELO, M. A. Vehicle logo recognition in traffic images using hog features and svm. In: **IEEE International Conference on Intelligent Transportation Systems (ITSC)**. [S.l.: s.n.], 2013. <<https://doi.org/10.1109/ITSC.2013.6728559>>.
- LONG, J.; SHELHAMER, E.; DARRELL, T. Fully convolutional networks for semantic segmentation. In: **Conference on Computer Vision and Pattern Recognition (CVPR)**. [S.l.: s.n.], 2015. <<https://doi.org/10.1109/CVPR.2015.7298965>>.
- LUO, S. et al. Real-time action recognition in surveillance videos using convnets. In: **International Conference on Neural Information Processing (ICONIP)**. [S.l.: s.n.], 2016. <[https://doi.org/10.1007/978-3-319-46675-0\\_58](https://doi.org/10.1007/978-3-319-46675-0_58)>.
- MARCOS, J. V. et al. Automated pollen identification using microscopic imaging and texture analysis. **Micron**, 2015. <<https://doi.org/10.1016/j.micron.2014.09.002>>.
- NISGOSKI, S.; OLIVEIRA, A. A.; NIZ, G. I. B. M. Artificial neural network and simca classification in some wood discrimination based on near-infrared spectra. **Wood Science and Technology**, 2017. <<https://doi.org/10.1007/s00226-017-0915-8>>.
- OJALA, T.; PIETIKÄINEN, M.; HARWOOD, D. A comparative study of texture measures with classification based on featured distributions. **Pattern Recognition**, 1996. <[https://doi.org/10.1016/0031-3203\(95\)00067-4](https://doi.org/10.1016/0031-3203(95)00067-4)>.
- OJANSIVU, V.; HEIKKILÄ, J. Blur insensitive texture classification using local phase quantization. In: **International Conference on Image and Signal Processing (ICISP)**. [S.l.: s.n.], 2008. <[https://doi.org/10.1007/978-3-540-69905-7\\_27](https://doi.org/10.1007/978-3-540-69905-7_27)>.
- PENG, Z. Robust wood species recognition using variable color information. **Optik - International Journal for Light and Electron Optics**, 2013. <<https://doi.org/10.1016/j.ijleo.2012.08.058>>.
- POZO-BAÑOS, M. del et al. Features extraction techniques for pollen grain classification. **Neurocomputing**, 2015. <<https://doi.org/10.1016/j.neucom.2014.05.085>>.

- PUNYASENA, S. W. et al. Classifying black and white spruce pollen using layered machine learning. **New Phytologist**, 2012. <<https://doi.org/10.1111/j.1469-8137.2012.04291.x>>.
- RAHTU, E. et al. Local phase quantization for blur-insensitive image analysis. **Image and Vision Computing**, 2012. <<https://doi.org/10.1016/j.imavis.2012.04.001>>.
- REDONDO, R. et al. Pollen segmentation and feature evaluation for automatic classification in bright-field microscopy. **Computers and Electronics in Agriculture**, 2015. <<https://doi.org/10.1016/j.compag.2014.09.020>>.
- RONNEBERGER, O.; BURKHARDT, H.; SCHULTZ, E. General-purpose object recognition in 3d volume data sets using gray-scale invariants - classification of airborne pollen-grains recorded with a confocal laser scanning microscope. In: **International Conference on Pattern Recognition (ICPR)**. [S.l.: s.n.], 2002. <<https://doi.org/10.1109/ICPR.2002.1048297>>.
- RUSSAKOVSKY, O. et al. ImageNet Large Scale Visual Recognition Challenge. **International Journal of Computer Vision**, 2015. <<https://doi.org/10.1007/s11263-015-0816-y>>.
- SAJJADI, M.; JAVANMARDI, M.; TASDIZEN, T. Regularization with stochastic transformations and perturbations for deep semi-supervised learning. In: **Conference on Neural Information Processing Systems (NIPS)**. [S.l.: s.n.], 2016.
- SANTOS, F. L. C. dos et al. Computer vision for virus image classification. **Biosystems Engineering**, 2015. <<https://doi.org/10.1016/j.biosystemseng.2015.01.005>>.
- SARWAR, A. K. M. G.; HOSHINO, Y.; ARAKI, H. Pollen morphology and its taxonomic significance in the genus *bomarea mirb.* (alstroemeriaceae) - i. subgenera *baccata*, *sphaerine*, and *wichuraea*. **Acta Botanica Brasilica**, 2015. <<https://doi.org/10.1590/0102-33062015abb0077>>.
- SEVILLANO, V.; AZNARTE, J. L. Improving classification of pollen grain images of the polen23e dataset through three different applications of deep learning convolutional neural networks. **PLOS ONE**, 2018. <<https://doi.org/10.1371/journal.pone.0201807>>.
- SILVA, N. R. da et al. Automated classification of wood transverse cross-section micro-imagery from 77 commercial central-african timber species. **Annals of Forest Science**, 2017. <<https://doi.org/10.1007/s13595-017-0619-0>>.
- SRIVASTAVA, N. et al. Dropout: A simple way to prevent neural networks from overfitting. **Journal of Machine Learning Research**, 2014.
- SUNDARAM, M. et al. Wood species classification based on local edge distributions. **Optik - International Journal for Light and Electron Optics**, 2015. <<https://doi.org/10.1016/j.ijleo.2015.07.044>>.
- SZEGEDY, C. et al. Rethinking the inception architecture for computer vision. **Conference on Computer Vision and Pattern Recognition (CVPR)**, 2016. <<https://doi.org/10.1109/CVPR.2016.308>>.
- SZIBOR, R. et al. Pollen analysis reveals murder season. **Nature**, 1998. <<https://doi.org/10.1038/26646>>.

- TELLO-MIJARES, S.; FLORES, F. A novel method for the separation of overlapping pollen species for automated detection and classification. **Computational and Mathematical Methods in Medicine**, 2016. <<https://doi.org/10.1155/2016%2F5689346>>.
- TOU, J. Y.; TAY, Y. H.; LAU, P. Y. A comparative study for texture classification techniques on wood species recognition problem. In: **International Conference on Natural Computation**. [S.l.: s.n.], 2009. <<https://doi.org/10.1109/ICNC.2009.594>>.
- TRELOAR, W. J.; TAYLOR, G. E.; FLENLEY, J. R. Towards automation of palynology 1: analysis of pollen shape and ornamentation using simple geometric measures, derived from scanning electron microscope images. **Journal of Quaternary Science**, 2004. <<https://doi.org/10.1002/jqs.871>>.
- VALE, F. et al. Applications of transmission electron microscopy to virus detection and identification. **Microscopy: Science, Technology, Applications and Education**, 2010.
- VU, Q. D. et al. Methods for segmentation and classification of digital microscopy tissue images. **Frontiers in Bioengineering and Biotechnology**, 2019. <<https://doi.org/10.3389/fbioe.2019.00053>>.
- WEN, Z. et al. Virus image classification using multi-scale completed local binary pattern features extracted from filtered images by multi-scale principal component analysis. **Pattern Recognition Letters**, 2016. <<https://doi.org/10.1016/j.patrec.2016.04.022>>.
- WEN, Z.-J. et al. Latent local feature extraction for low-resolution virus image classification. **Journal of the Operations Research Society of China**, 2018. <<https://doi.org/10.1007/s40305-018-0212-8>>.
- WHEELER, E. A. Inside wood – a web resource for hardwood anatomy. **IAWA Journal**, 2011. <<https://doi.org/10.1163/22941932-90000051>>.
- XIAO, X.-Y. et al. Hog-based approach for leaf classification. In: **International Conference on Intelligent Computing (ICIC)**. [S.l.: s.n.], 2010. <[https://doi.org/10.1007/978-3-642-14932-0\\_19](https://doi.org/10.1007/978-3-642-14932-0_19)>.
- YADAV, A. R. et al. Classification of hardwood species using ann classifier. In: **National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG)**. [S.l.: s.n.], 2013. <<https://doi.org/10.1109/NCVPRIPG.2013.6776231>>.
- YOSINSKI, J. et al. How transferable are features in deep neural networks? In: **Conference on Neural Information Processing Systems (NIPS)**. [S.l.: s.n.], 2014.
- ZAMRI, M. I. P. et al. Tree species classification based on image analysis using improved-basic gray level aura matrix. **Computers and Electronics in Agriculture**, 2016. <<https://doi.org/10.1016/j.compag.2016.04.004>>.
- ZHANG, Y. et al. Towards automation of palynology 3: pollen pattern recognition using gabor transforms and digital moments. **Journal of Quaternary Science**, 2004. <<https://doi.org/10.1002/jqs.875>>.

ZHAO, P.; CAO, J. Wood species identification using spectral reflectance feature and optimal illumination radian design. **Journal of Forestry Research**, 2016. <<https://doi.org/10.1007/s11676-015-0171-4>>.