

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ  
PROGRAMA DE PÓS-GRADUAÇÃO EM TECNOLOGIAS COMPUTACIONAIS  
PARA O AGRONEGÓCIO

WELLINGTON DE OLIVEIRA

**SOFTWARE PARA RECONHECIMENTO DE ESPÉCIES FLORESTAIS  
A PARTIR DE IMAGENS DIGITAIS DE MADEIRAS UTILIZANDO  
DEEP LEARNING**

DISSERTAÇÃO

MEDIANEIRA-PR

2018

WELLINGTON DE OLIVEIRA

**SOFTWARE PARA RECONHECIMENTO DE ESPÉCIES FLORESTAIS  
A PARTIR DE IMAGENS DIGITAIS DE MADEIRAS UTILIZANDO  
DEEP LEARNING**

Dissertação apresentada ao Programa de Pós-Graduação em Tecnologias Computacionais para o Agronegócio – PPGTCA – da Universidade Tecnológica Federal do Paraná – UTFPR – Campus Medianeira, como requisito parcial para obtenção do título de Mestre em Tecnologias Computacionais para o Agronegócio / Área de Concentração: Tecnologias Computacionais Aplicadas à Produção Agrícola e Agroindústria.

Orientador: Prof. Dr. Pedro Luiz de Paula Filho

Coorientador: Prof. Dr. Jefferson Gustavo Martins

MEDIANEIRA-PR

2018

#### Dados Internacionais de Catalogação na Publicação

O48s

OLIVEIRA, Wellington de

Software para reconhecimento de espécies florestais a partir de imagens digitais de madeiras utilizando deep learning / Wellington de Oliveira - 2018

103 f. : il. ; 30 cm.

Orientador: Pedro Luiz de Paula Filho.

Coorientador: Jefferson Gustavo Martins.

Dissertação (Mestrado) – Universidade Tecnológica Federal do Paraná. Programa de Pós-Graduação em Tecnologias Computacionais para o Agronegócio. Medianeira, 2018.

Inclui bibliografias.

1. Processamento de imagens – Técnicas digitais. 2. Redes neurais (Computação). 3. Tecnologias Computacionais- Dissertações. I. Paula Filho, Pedro Luiz de. orient. II. Martins, Jefferson Gustavo, coorient. III. Universidade Tecnológica Federal do Paraná. Programa de Pós-Graduação em Tecnologias Computacionais para o Agronegócio. IV. Título.

CDD: 004



## TERMO DE APROVAÇÃO

### SOFTWARE PARA RECONHECIMENTO DE ESPÉCIES FLORESTAIS A PARTIR DE IMAGENS DIGITAIS DE MADEIRAS UTILIZANDO DEEP LEARNING

Por

**WELLINGTON DE OLIVEIRA**

Essa Dissertação foi apresentada às catorze horas, do dia dezesseis de março de dois mil e dezoito, como requisito parcial para a obtenção do título de Mestre em Tecnologias Computacionais para o Agronegócio, Linha de Pesquisa Tecnologias Computacionais Aplicadas à Produção Agrícola, no Programa de Pós-Graduação em Tecnologias Computacionais para o Agronegócio – PPGTCA, da Universidade Tecnológica Federal do Paraná. O candidato foi arguido pela Banca Examinadora composta pelos professores abaixo. Após deliberação, a Banca Examinadora considerou o trabalho APROVADO.

---

Prof. Dr. Pedro Luiz de Paula Filho (Orientador – PPGTCA)

---

Prof. Dr. Arnaldo Candido Junior (Membro Interno – PPGTCA)

---

Prof. Dr. Paulo Lopes de Menezes (Membro Externo – UTFPR)

---

Prof. Dr. Luiz Eduardo Soares de Oliveira (Membro Externo – UFPR, com participação a distância, sendo a assinatura no presente Termo dispensada conforme Regulamento da Pós-Graduação *Stricto Sensu* da UTFPR)

**A via original com as assinaturas encontra-se na secretaria do programa.**

Aos meus familiares e amigos que  
sempre me apoiaram.

## **AGRADECIMENTOS**

Acima de tudo agradeço a Deus que iluminou meu caminho desde o início de minha vida acadêmica.

Agradeço aos meus pais Joel Barros de Oliveira e Elizabete Maria de Oliveira que nunca mediram esforços para me apoiar e incentivar meus estudos, mesmo nos momentos difíceis. Com certeza essa conquista não é só minha, mas sim de cada um deles.

Agradeço também ao meu orientador Prof. Dr. Pedro Luiz de Paula Filho e coorientador Prof. Dr. Jefferson Gustavo Martins que, além de disponibilizarem as bases de imagens para desenvolver meu trabalho, sempre me atenderam e me deram o direcionamento para que esse projeto pudesse ser concluído.

Agradeço a equipe de professores do programa de mestrado da UTFPR, em especial ao Prof. Dr. Claudio Leones Bazzi e Prof. Me. Hamilton Pereira da Silva, os quais me atenderam em período de férias para que eu pudesse ter acesso à infraestrutura de TI e realizar os experimentos.

Agradeço a todos os professores e colegas de mestrado que hoje posso chamar de amigos.

"O valor das coisas não está no tempo em que elas duram, mas na intensidade com que acontecem. Por isso existem momentos inesquecíveis, coisas inexplicáveis e pessoas incomparáveis."  
(Fernando Pessoa)

## RESUMO

OLIVEIRA, Wellington; PAULA FILHO, Pedro Luiz; MARTINS, Jefferson Gustavo. Software para Reconhecimento de Espécies Florestais a partir de Imagens Digitais de Madeiras Utilizando Deep Learning. 2018. 103 f. Dissertação (Mestrado em Tecnologias Computacionais para o Agronegócio) – Programa de Pós-Graduação em Tecnologias Computacionais para o Agronegócio. Medianeira, 2018.

Classificar espécies florestais é um processo essencial para o correto manejo da madeira e controle florestal. Após o corte, perde-se muitas das características da espécie e a identificação torna-se uma tarefa bem mais difícil. Neste contexto, torna-se necessário a análise anatômica da madeira, a qual é realizada por especialistas que conhecem muito bem as estruturas celulares presentes em cada espécie. Porém, tal metodologia aborda técnicas pouco automatizadas, tornando a atividade demorada e passível de erros. Estes fatores prejudicam o controle e tomada de decisões por parte dos órgãos ambientais. O uso de visão computacional é uma alternativa para o reconhecimento automatizado, pois permite a construção de modelos inteligentes que, a partir de imagens, são capazes de detectar características e realizar a classificação final. Existem algumas técnicas que são o estado da arte em Processamento Digital de Imagens e Inteligência Artificial, como por exemplo, as redes neurais convolucionais. Tais redes são técnicas de aprendizado profundo (Deep Learning) que consiste na construção de modelos a partir de dados brutos (imagens, áudios, etc.). Em outras palavras, utiliza-se o mínimo de pré-processamento nas imagens e, no processo de treinamento, estes modelos são capazes de aprender os filtros a serem aplicados para a extração das características. Desta forma, tal técnica abstrai muito da complexidade existente na fase de pré-processamento em abordagens tradicionais. Esta pesquisa aborda a realização de alguns experimentos utilizando técnicas tradicionais e experimentos utilizando redes neurais convolucionais (foco da pesquisa). Foram utilizadas duas bases, sendo uma composta por imagens macroscópicas e outra por imagens microscópicas, para as quais foram gerados três modelos: reconhecimento de escala, reconhecimento de espécies a partir imagens macroscópicas e reconhecimento a partir de imagens microscópicas. Os melhores modelos proporcionaram taxas de reconhecimento de 100% para a base de escala, 98,73% para a base macroscópica e 99,11% para a base microscópica, os quais superaram os resultados obtidos para as respectivas bases em pesquisas correlatas. Por fim, foi desenvolvido um software web como produto final, para o qual foram acoplados os três melhores modelos obtidos.

**Palavras-chave:** Processamento Digital de Imagens. Redes Neurais Convolucionais. Anatomia da Madeira. Visão Computacional.



## ABSTRACT

OLIVEIRA, Wellington; PAULA FILHO, Pedro Luiz; MARTINS, Jefferson Gustavo. Software for Forest Species Recognition Based on Digital Images of Wood and Deep Learning Techniques. 2018. 103 f. Dissertação (Mestrado em Tecnologias Computacionais para o Agronegócio) – Programa de Pós-Graduação em Tecnologias Computacionais para o Agronegócio. Medianeira, 2018.

Classifying forest species is an essential process for the correct management of wood and forest control. After cutting off the trunk of the tree, many of the characteristics of the species are lost and identifying them becomes a much more difficult task. In this context, an anatomical analysis of the wood becomes necessary by specialists who know very well the cellular structures in each species. However, such methodology approaches few automated techniques, making it a delayed and error-prone activity. These factors undermine environmental control and decision-making. The use of computer vision is an alternative to automatic recognition, since it allows the development of intelligent models which, from images, are able to detect features and perform a final classification. There are some techniques that are state of the art in Digital Image Processing and Artificial Intelligence, such as, for example, convolutional neural networks. These networks are deep learning techniques that consists of creating models from raw data (images, audios, etc.). In this technique, the minimum preprocessing is used in the images and, in the training process, these models are able to learn the convolutional filters to be applied in feature extraction process. In this way, this technique abstracts a lot from the pre-processing complexity existing in traditional approaches. This research addresses some experiments using traditional techniques and experiments using convolutional neural networks (the focus of the research). Two original datasets were used, one including macroscopic images and the other including microscopic images, for which three models were created: scale recognition, species recognition from macroscopic images and species recognition from microscopic images. The best models provide 100% recognition rates for the scale dataset, 98.73% for the macroscopic dataset and 99.11% for the microscopic dataset, which they were better results than the ones obtained for the respective datasets in related researches. Finally, a web software was developed as a final product, using the three best models.

**Keywords:** Digital Image Processing. Convolutional Neural Networks. Wood Anatomy. Computer Vision.

## LISTA DE FIGURAS

Figura 1 – Exemplos de Imagens de Madeira .....	18
Figura 2 – Metodologia Clássica de Identificação de Espécies.....	23
Figura 3 - Estrutura Macroscópica da Madeira.....	24
Figura 4 – Estrutura Microscópica da Madeira .....	25
Figura 5 – Passos Fundamentais em PDI.....	28
Figura 6 – Matriz de Coocorrência (GLCM).....	30
Figura 7 – Modelo Matemático de um Neurônio Artificial .....	35
Figura 8 – Funções de Ativação.....	36
Figura 9 – Topologia MLP de RNA.....	37
Figura 10 – Treinamento da RNA e Algoritmo de <i>BP</i> .....	38
Figura 11 – Validação de Treinamento .....	39
Figura 12 – Conceito de Reconhecimento de Imagens com <i>Deep Learning</i> .....	41
Figura 13 – Convolução em Imagens.....	42
Figura 14 – Variações de Características .....	43
Figura 15 – Representações Invariantes em CNN .....	44
Figura 16 – Função Max-Pooling .....	44
Figura 17 – Rede Neural Convolutacional.....	45
Figura 18 – Arquitetura de CNN: AlexNet .....	55
Figura 19 – Arquitetura de CNN: VGG16 .....	56
Figura 20 – Arquitetura de CNN: ResNet50 .....	57
Figura 21 – Fluxograma de Atividades.....	61
Figura 22 – Número de Bits x Níveis de Cinza.....	63
Figura 23 – Transformações para Data Augmentation.....	65
Figura 24 – GLCM: Confusão de Classificação na Base Microscópica.....	71
Figura 25 – GLCM: Confusão de Classificação na Base Macroscópica .....	72
Figura 26 – Treinamentos com Imagens Redimensionadas .....	73
Figura 27 – Treinamentos com <i>Data Augmentation</i> – Base Macro - Fragmentos.....	73
Figura 28 – Treinamentos com <i>Data Augmentation</i> – Base Macro - Fusões .....	74
Figura 29 – Treinamentos com <i>Data Augmentation</i> – Base Micro - Fragmentos.....	74
Figura 30 – Treinamentos com <i>Data Augmentation</i> – Base Micro - Fusões .....	75
Figura 31 – Análise das Fusões por Votação e Média das Probabilidades.....	76
Figura 32 – Treinamento da Base de Escala – Análise dos Fragmentos.....	77

Figura 33 – Treinamento da Base de Escala – Análise das Imagens Inteiras .....	77
Figura 34 – Arquitetura Final – Modelo de Escala.....	78
Figura 35 – Treinamento da Base Macro – Análise dos Fragmentos.....	79
Figura 36 – Treinamento da Base Macro – Análise das Imagens Inteiras .....	79
Figura 37 – Arquitetura Final – Modelo Macro .....	80
Figura 38 – Treinamento da Base Micro – Análise dos Fragmentos.....	80
Figura 39 – Treinamento da Base Micro – Análise das Imagens Inteiras .....	81
Figura 40 – Arquitetura Final – Modelo Micro.....	81
Figura 41 – Interface Gráfica do Produto/Software .....	89
Figura 42 – Tela de Detalhes do Produto/Software.....	89
Figura 43 – Alerta de Possível Espécie Não Conhecida .....	90

## LISTA DE TABELAS

Tabela 1 – Equações dos 14 Descritores Texturais .....	31
Tabela 2 – Demonstração da Composição da Base Macro .....	60
Tabela 3 – Demonstração da Composição da Base Micro .....	60
Tabela 4 – Composição da Base de Imagens Macroscópicas .....	100
Tabela 5 – Composição da Base de Imagens Microscópicas .....	101

## LISTA DE QUADROS

Quadro 1 – Síntese das Pesquisas Correlatas (Bases da Presente Pesquisa) .....	52
Quadro 2 – Síntese das Pesquisas Correlatas (Outras Bases).....	52
Quadro 3 – Quantidade de Imagens nas Bases Após Data Augmentation.....	66
Quadro 4 – Exemplo de Fusão de Classificações por Votação.....	67
Quadro 5 – Exemplo de Fusão de Classificações por Média das Probabilidades.....	67
Quadro 6 – Descritores de GLCM e Taxas de Reconhecimento com RNA .....	69
Quadro 7 – Taxas de Reconhecimento dos Melhores Modelos .....	82
Quadro 8 – Erros de Classificação do Modelo Macroscópico .....	83
Quadro 9 – Erros de Classificação do Modelo Microscópico .....	83
Quadro 10 – Visualização de Erro de Classificação do Modelo Macroscópico .....	84
Quadro 11 – Visualização de Erro de Classificação do Modelo Microscópico .....	85
Quadro 12 – Ranking de Reconhecimento – Cinco Maiores Probabilidades .....	86
Quadro 13 – Taxas de Reconhecimento – GLCM vs. Deep Learning.....	87
Quadro 14 – Taxas de Reconhecimento – Pesquisas Correlatas .....	87

## LISTA DE SIGLAS E ABREVIATURAS

<b>AM</b>	Aprendizado de Máquina
<b>BP</b>	Back-Propagation
<b>CCD</b>	Charge Coupled Device
<b>cm</b>	Centímetros
<b>CNN</b>	Convolutional Neural Networks
<b>CPU</b>	Central Processing Unit
<b>CUDA</b>	Compute Unified Device Architecture
<b>DL</b>	Deep Learning
<b>GLCM</b>	Gray Level Co-occurrence Matrices
<b>GPU</b>	Graphics Processing Unit
<b>IBAMA</b>	Instituto Brasileiro do Meio Ambiente
<b>IA</b>	Inteligência Artificial
<b>ID</b>	Identificador
<b>ILSVRC</b>	ImageNet Large Scale Visual Recognition Challenge
<b>IAWA</b>	International Association of Wood Anatomists
<b>IDE</b>	Integrated Development Environment
<b>LBP</b>	Local Binary Pattern
<b>LPQ</b>	Local Phase Quantization
<b>MLP</b>	Multi-Layer Perceptron (Perceptron Multicamadas)
<b>mm</b>	Milímetros
<b>MSE</b>	Mean Squared Error
<b>PDI</b>	Processamento Digital de Imagens
<b>PF</b>	Polícia Federal
<b>RGB</b>	Red, Green, Blue
<b>ReLU</b>	Rectified Linear Unit
<b>RNA</b>	Redes Neurais Artificiais
<b>RNN</b>	Recurrent Neural Networks
<b>SIFT</b>	Scale Invariant Feature Transform
<b>SGBD</b>	Sistema Gerenciador de Banco de Dados
<b>SSE</b>	Sum of Squared Errors
<b>SURF</b>	Speed-Up Robust Feature
<b>SVM</b>	Support Vector Machines

**TI** Tecnologia da Informação  
**VC** Visão Computacional  
**UTFPR** Universidade Tecnológica Federal do Paraná

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>17</b>
1.1	OBJETIVOS	19
1.2	JUSTIFICATIVA E CONTRIBUIÇÕES	20
1.3	ORGANIZAÇÃO DO TRABALHO	21
<b>2</b>	<b>FUNDAMENTAÇÃO TEÓRICA</b>	<b>22</b>
2.1	ANATOMIA DA MADEIRA	22
2.1.1	Estrutura Macroscópica da Madeira	23
2.1.2	Estrutura Microscópica da Madeira	25
2.1.3	Metodologia Tradicional de Identificação de Espécies Florestais	26
2.2	VISÃO COMPUTACIONAL	27
2.2.1	Processamento Digital de Imagens	27
2.2.1.1	Passos fundamentais em PDI	28
2.2.1.2	Extração de Descritores a partir de Matriz de Coocorrência	30
2.2.2	Inteligência Artificial e Aprendizado de Máquina	32
2.2.2.1	Redes neurais artificiais Perceptron Multicamadas	35
2.2.3	Deep Learning	39
2.2.3.1	Redes neurais convolucionais	41
2.2.3.2	Treinamento de uma CNN	46
2.2.3.3	DropOut, Transfer Learning e Data Augmentation	47
2.3	ESTADO DA ARTE	49
2.3.1	Reconhecimento de Madeiras	49
2.3.1.1	Pesquisas correlatas	50
2.3.2	Arquiteturas de CNN	53
2.3.2.1	Arquitetura AlexNet	54
2.3.2.2	Arquitetura VGG16	55
2.3.2.3	Arquitetura ResNet50	56
<b>3</b>	<b>METODOLOGIA</b>	<b>58</b>
3.1	AMBIENTE DE EXPERIMENTOS E DESENVOLVIMENTO	58
3.1.1	Bases de Imagens	59
3.2	MÉTODO EXPERIMENTAL PROPOSTO	61
3.2.1	Abordagem Tradicional: GLCM e RNA MLP	62
3.2.2	Abordagem com Deep Learning: CNN	63
3.2.2.1	Data augmentation e descrição das novas bases	64
3.2.2.2	Técnicas de fusão das subimagens	66
3.2.2.3	Construção dos modelos de CNN e produto final	68
<b>4</b>	<b>RESULTADOS EXPERIMENTAIS</b>	<b>69</b>
4.1	ABORDAGEM TRADICIONAL: GLCM E RNA MLP	69
4.1.1	Erros de Classificação	71
4.2	ABORDAGEM COM DEEP LEARNING: CNN	72
4.2.1	Seleção de Arquitetura e Técnica de Fusão	73
4.2.2	Modelo para Reconhecimento de Escala	76
4.2.3	Modelo para Reconhecimento Macroscópico	78
4.2.4	Modelo para Reconhecimento Microscópico	80
<b>5</b>	<b>ANÁLISE E DISCUSSÃO DE RESULTADOS</b>	<b>82</b>
5.1	TAXAS DE RECONHECIMENTO E ANÁLISE DE ERROS: CNN	82
5.1.1	Ranking de Reconhecimento com as Cinco Maiores Probabilidades	85
5.2	COMPARAÇÃO DOS RESULTADOS COM O ESTADO DA ARTE	86

5.3 O PRODUTO FINAL: SOFTWARE WEB.....	88
<b>6 CONSIDERAÇÕES FINAIS.....</b>	<b>91</b>
6.1 TRABALHOS FUTUROS.....	92
<b>REFERÊNCIAS.....</b>	<b>94</b>
<b>ANEXO A - Composição da Base de Imagens Macroscópicas .....</b>	<b>100</b>
<b>ANEXO B - Composição da Base de Imagens Microscópicas .....</b>	<b>101</b>



## 1 INTRODUÇÃO

A classificação das diferentes espécies florestais é fundamental para o correto manejo de madeiras na construção civil, madeireiras, industrialização, exportação, entre outros. Tal classificação possibilita a identificação da comercialização de espécies ameaçadas de extinção e fraudes nas transações envolvendo madeiras como, por exemplo, a venda de uma espécie de menor valor como sendo uma espécie de maior valor comercial (MOHAN; VENKATACHALAPATHY; RAI, 2014).

Embora exista uma série de características presentes nestas espécies, como folhas, frutos, sementes e odores, após o corte e a retirada do tronco das florestas, perdem-se muitas destas características e a identificação torna-se uma tarefa bem mais difícil. Neste contexto, torna-se necessário uma análise anatômica detalhada da madeira por especialistas que conhecem muito bem as estruturas celulares presentes em cada espécie (MARTINS *et al.*, 2015 e PAULA *et al.*, 2014).

Segundo Correia (2009), o estudo anatômico da madeira contempla a análise das características morfológicas e de tonalidades das camadas macro e microscópicas de um tronco. Neste contexto, o reconhecimento de padrões nas diferentes camadas da madeira possibilita a identificação da respectiva espécie.

De acordo com Lesk (2008), para problemas como este, o uso de sistemas computacionais inteligentes pode automatizar processos e proporcionar resultados até melhores do que o esperado para o ser humano (LESK, 2008), já que muitas atividades podem ser demoradas e estressantes, aumentando a probabilidade de erros.

Os sistemas que utilizam técnicas de Visão Computacional, são exemplos de sistemas inteligentes que buscam assemelhar-se à visão humana para solução de problemas complexos como o reconhecimento de imagens. As técnicas de Visão Computacional têm sido utilizadas na indústria madeireira desde a década de 1980, auxiliando em problemas como identificação de defeitos em troncos, otimização de cortes, entre outros (PEDRINI; SCHWARTZ, 2008). Entretanto, existem outros problemas a serem explorados como, por exemplo, a construção de um sistema capaz de realizar o reconhecimento automático de espécies florestais a partir de imagens da madeira.

Ao visualizar imagens macro e microscópicas de madeiras, observam-se diferentes padrões morfológicos e de coloração presentes em cada espécie florestal. Como exemplo, a Figura 1 apresenta a estrutura anatômica macro e microscópica da espécie “Ocotea Porosa” (esquerda) e também da espécie “Araucaria Angustifolia” (direita).



**Figura 1 – Exemplos de Imagens de Madeira**  
**Fonte: Adaptado de Paula *et al.* (2014).**

Algumas abordagens se tornaram o estado da arte na área de Visão Computacional, como o Deep Learning, um conjunto de técnicas de aprendizado de máquina com abordagens profundas, as quais permitem a construção de modelos capazes de realizar o reconhecimento e classificação de sinais a partir de dados brutos (sem tratamento ou com o mínimo de pré-processamento).

Em um modelo baseado em Deep Learning como, por exemplo, uma rede neural convolucional, o treinamento é realizado de forma a automatizar as definições das técnicas para extração das características dos dados. Em outras palavras, o que antes (em modelos tradicionais) era realizado por intervenção e definição do

pesquisador, agora tornou-se uma etapa dinâmica e aprendida pelo próprio modelo na fase de treinamento (GOODFELLOW; BENGIO; COURVILLE, 2016).

Mediante o contexto apresentado, levanta-se a hipótese de que é possível classificar espécies florestais a partir de imagens de madeiras utilizando técnicas de Deep Learning. Logo, esta pesquisa aborda a realização de uma série de experimentos tradicionais e com Deep Learning a fim de avaliar tal hipótese com a finalidade de obter melhores resultados e menor complexidade em relação às pesquisas já existentes.

## 1.1 OBJETIVOS

O objetivo geral da presente pesquisa é desenvolver um software capaz de classificar espécies florestais a partir de imagens macro e microscópicas de madeiras utilizando técnicas de Deep Learning.

Para atingir o objetivo geral, foram definidos os seguintes objetivos específicos:

- a) Adequar duas bases de imagens já existentes: uma de imagens macroscópicas e outra de imagens microscópicas de madeiras de diferentes espécies florestais;
- b) Construir três modelos de reconhecimento baseados em técnicas tradicionais para posteriores comparações: modelo de escala, macroscópico e microscópico;
- c) Construir três modelos de reconhecimento baseados em Deep Learning: modelo de escala, macroscópico e microscópico;
- d) Desenvolver um software web como produto final utilizando os três melhores modelos baseados em Deep Learning.

## 1.2 JUSTIFICATIVA E CONTRIBUIÇÕES

A necessidade da identificação de espécies florestais tem aumentado no decorrer dos anos devido a sua aplicação em projetos de construção civil, produção de carvão, madeireiras, industrialização, exportação e afins. O desmatamento, as fraudes e a comercialização de espécies ameaçadas de extinção têm se tornado cada vez mais frequentes e a complexidade de identificação destas irregularidades dificulta a aplicação das medidas legais (PAULA *et al.*, 2014 e MOHAN; VENKATACHALAPATHY; RAI, 2014).

O cenário descrito é agravado pelo fato de que os órgãos como o Instituto Brasileiro do Meio Ambiente (IBAMA) e a Polícia Federal (PF) não possuem recursos suficientes para realizar a fiscalização de forma satisfatória. Isto tem chamado a atenção de pesquisadores de Ciência da Computação, Engenharia Florestal, Tecnologia da Madeira e áreas correlatas.

Algumas pesquisas que utilizaram a computação aplicada à identificação de espécies a partir de imagens de madeiras proporcionaram resultados muito relevantes. Em contrapartida, não são muitos os projetos publicados, os quais compreendem ou análise macroscópica ou microscópica (vide Seção 2.3.1).

Devido à essa problemática, surgiu a motivação por desenvolver um projeto que associe os dois níveis de análise e contemple também a construção de uma ferramenta específica de apoio aos laboratórios de manejo florestal, além da contribuição com futuras pesquisas por meio da publicação dos experimentos.

As contribuições científicas deste projeto compreendem a geração de um protocolo de reconhecimento de imagens utilizando técnicas de Deep Learning (estado da arte em Visão Computacional), que possibilita a construção de um modelo inteligente, capaz de extrair padrões relevantes para a classificação de uma imagem digital de madeira, com o mínimo de intervenção especialista e melhores resultados com menor complexidade técnica em comparação aos modelos tradicionais.

As contribuições sociais compreendem o desenvolvimento de um software web como produto final para auxílio no reconhecimento de espécies em laboratórios de manejo florestal, de forma precisa e ágil, além de auxiliar os órgãos de fiscalização ambiental na detecção de fraudes e comercialização ilegal.

### 1.3 ORGANIZAÇÃO DO TRABALHO

A Seção 2 é composta pela fundamentação teórica, a qual está subdividida em três grandes seções: Anatomia da Madeira, Visão Computacional e Estado da Arte. Estas três seções descrevem as técnicas e servem de embasamento para a fase experimental proposta.

Na Seção 3 são apresentadas as bases de imagens utilizadas, bem como os materiais, métodos e o planejamento experimental. Já a Seção 4 aborda os resultados parciais obtidos a partir dos experimentos propostos.

A Seção 5 apresenta a análise dos resultados obtidos com Deep Learning em comparação aos obtidos em pesquisas correlatas e aos obtidos com os experimentos realizados com abordagens tradicionais. Por fim, aborda também a descrição do produto final.

Finalmente, a Seção 6 aborda as considerações finais, análise da hipótese levantada e proposta de trabalhos futuros.

## 2 FUNDAMENTAÇÃO TEÓRICA

Esta seção aborda os conceitos bibliográficos essenciais para possibilitar a elaboração da pesquisa, experimentos e construção do produto final. Além da revisão técnica, também são mencionados os autores que trabalharam em projetos correlatos, bem como os resultados obtidos em suas pesquisas. A seção está subdividida em: Anatomia da Madeira, Visão Computacional e Estado da Arte.

### 2.1 ANATOMIA DA MADEIRA

O estudo anatômico da madeira se tornou um assunto tão importante, que em 1931 foi fundada a International Association of Wood Anatomists (IAWA), uma associação que tem como principais objetivos facilitar a coleta, armazenamento e troca de materiais de pesquisa relacionados à anatomia da madeira ao redor do mundo, criar a consciência da importância deste assunto para a ciência, tecnologia e conservação dos recursos naturais. O escritório sede da IAWA fica localizado na Holanda e é regido por um conselho internacional que padroniza a descrição de testes físicos e mecânicos já realizados para a identificação de madeiras.

Segundo Martins *et al.* (2015) e Mohan, Venkatachalapathy e Rai (2014), quando uma espécie não se encontra na floresta com suas características externas, a classificação da mesma consiste na análise anatômica da madeira. Esta análise é tradicionalmente realizada por especialistas que observam aspectos macro e microscópicos presentes na madeira a fim de encontrar padrões para reconhecimento da respectiva espécie.

A Figura 2 representa uma ideia do processo tradicional de identificação de espécies a partir de madeiras. Neste contexto, os troncos são retirados da natureza, serrados e, quando necessário, analisados por profissionais especialistas.



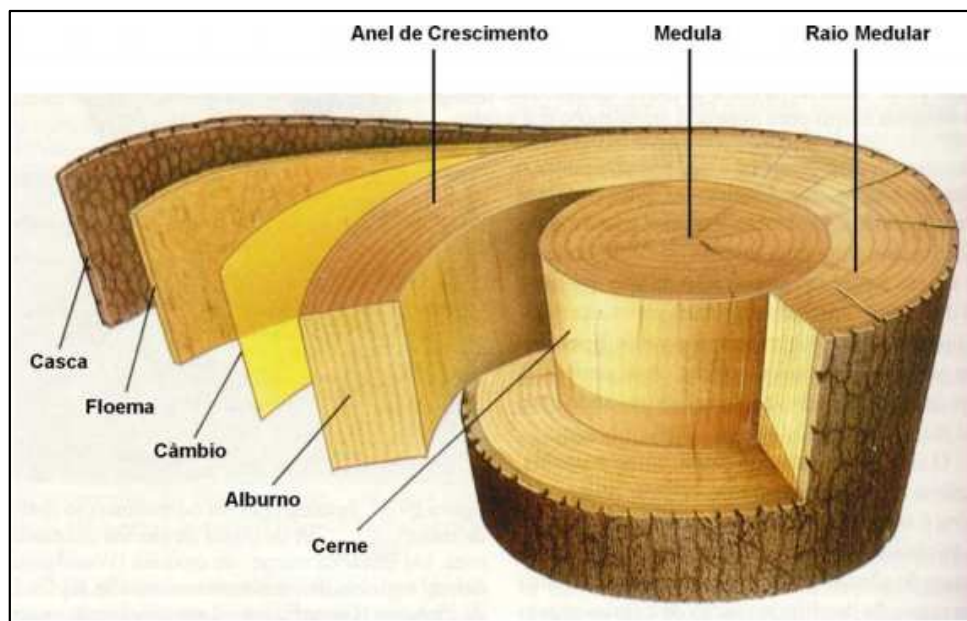
**Figura 2 – Metodologia Clássica de Identificação de Espécies**  
 Fonte: Adaptado de Coradin (2008).

Existem muitas características associadas à estrutura e anatomia da madeira, podendo abordar assuntos relacionados aos aspectos químicos, botânicos, substâncias moleculares, detalhes anatômicos, entre outros. Entretanto, as subseções a seguir estão delimitadas à abordagem anatômica, a qual é relevante o suficiente os experimentos computacionais propostos.

### 2.1.1 Estrutura Macroscópica da Madeira

De acordo com Zenid e Ceccantini (2007), as características anatômicas macroscópicas da madeira são aquelas observadas à olho nu ou com aproximações em até 10 vezes utilizando lupas ou microscópios. Segundo Amabis e Martho (1995) e Correia (2009), a estrutura macroscópica da madeira é composta por algumas características morfológicas e diferentes tonalidades de cor que podem ser observadas na Figura 3 e são descritas a seguir.





**Figura 3 - Estrutura Macroscópica da Madeira**  
**Fonte: Amabis e Martho (1995).**

Medula: parte central do tronco, constituída por tecido com pouca espessura e pouca resistência, o que a torna suscetível a apodrecimentos causados por fungos. Tem por função armazenar nutrientes e realizar o transporte ascendente de líquidos (MARTINS *et al.*, 2015).

Lenho (Cerne e Alburno): camada interior que possui uma coloração mais escura (Cerne) na região mais próxima à Medula e uma coloração um pouco mais clara na região mais próxima ao Câmbio (Alburno). O Cerne é a camada com maior densidade e resistência perante fatores de degradação biológica. O Alburno é uma camada mais úmida e menos resistente que corresponde à parte mais jovem da madeira com células vivas, grande quantidade de água e nutrientes que garantem a condução da seiva bruta, desde a raiz até a copa. É nesta seção que estão localizados os anéis de crescimento (AMABIS; MARTHO, 1995 e CORREIA, 2009).

Raios Medulares: trata-se de raios que se estendem de forma perpendicular ao eixo do tronco, da medula à casca e são muito importantes na classificação da madeira. Transportam nutrientes para as demais partes do tronco e contribuem para a rigidez da estrutura (AMABIS; MARTHO, 1995 e CORREIA, 2009).

Anéis de Crescimento: as atividades do Câmbio variam de acordo com as mudanças climáticas em função das estações do ano, gerando anéis que podem ser identificados facilmente por meio de um corte transversal. Estes anéis representam os ciclos anuais que caracterizam a idade da espécie, os quais representam a



transição entre os lenhos inicial e tardio formados anualmente, em que as atividades fisiológicas das plantas aumentam durante a primavera e diminuem durante o outono (ZENID; CECCANTINI, 2007).

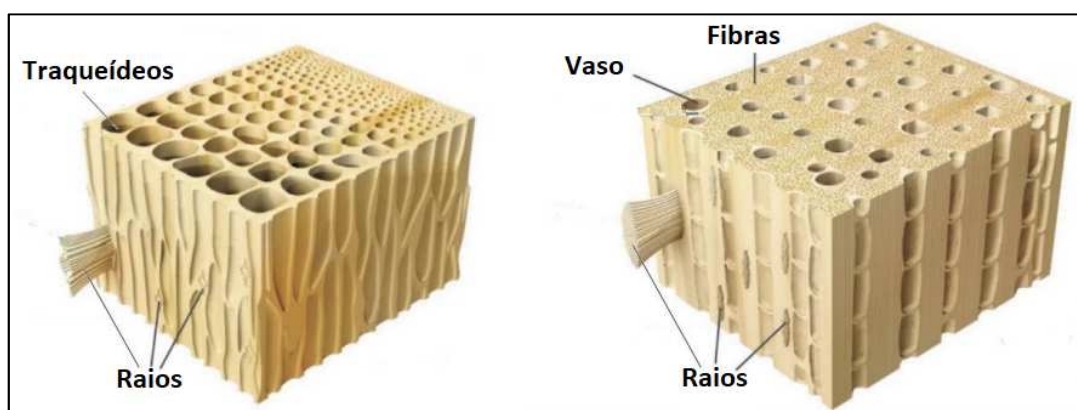
Câmbio: tem por função gerar novas células, pois é constituído por um tecido em constante transformação celular que são de suma importância para as demais camadas (AMABIS; MARTHO, 1995 e CORREIA, 2009).

Floema: trata-se de uma subcamada da casca que é responsável pela proteção do tronco e por conduzir a seiva das folhas para o crescimento da árvore (AMABIS; MARTHO, 1995 e CORREIA, 2009).

Casca: é a camada externa, com espessura irregular que contém duas partes: a Camada epidérmica, parte externa que é formada por células mortas que protegem o lenho e o Floema, formado por células vivas, conduzindo a seiva elaborada das folhas (AMABIS; MARTHO, 1995 e CORREIA, 2009).

### 2.1.2 Estrutura Microscópica da Madeira

De acordo com Zenid e Ceccantini (2007), as características anatômicas microscópicas da madeira são aquelas observadas com aproximações acima de 10 vezes, utilizando lupas ou microscópios. Segundo Koch *et al.* (2015), o tronco de uma árvore é composto por milhões de células microscópicas de diferentes tamanhos e formas. Estas células são divididas nos seguintes grupos: Traqueóides, Vasos, Fibras e Raios Medulares e podem ser observadas na Figura 4.



**Figura 4 – Estrutura Microscópica da Madeira**  
Fonte: Santos (2015).

Traqueídeos: células alongadas e orientadas em direção ao eixo longitudinal do tronco. Possuem dimensões aproximadas à 5mm de comprimento e 0,6mm de diâmetro, podendo compor até 95% da madeira. Tem por função transportar a seiva bruta de uma célula para outra, armazenar substâncias polimerizadas e atribuir resistência mecânica ao tronco (SANTOS, 2015).

Vasos: células longas com dimensões aproximadas à 1mm de comprimento e 0,3mm de diâmetro. Podem compor até 50% da madeira. Possui uma camada transversal arredondada e vazada e tem por função transportar a seiva bruta e armazenar substâncias polimerizadas (SANTOS, 2015).

Fibras: células longas com comprimento aproximado à 1,5mm. Possuem uma camada transversal arredondada e vazada, paredes grossas e pode compor até 50% da madeira. Tem por função atribuir resistência mecânica e rigidez ao tronco (SANTOS, 2015).

Raios Medulares (Células Radiais): células longas e achatadas localizadas horizontalmente da Casca à Medula. Tem por função transportar a seiva elaborada ao longo do lenho (SANTOS, 2015).

### 2.1.3 Metodologia Tradicional de Identificação de Espécies Florestais

De acordo com Mohan, Venkatachalapathy e Rai (2014), o processo de identificação da madeira é realizado por especialistas que realizam observação de aspectos anatômicos macro e microscópicos. Estes aspectos estão ligados à forma, tamanho e distribuição dos elementos na madeira que podem ser comparados a ferramentas de auxílio, como manuais ilustrados, chaves dicotômicas, bancos de dados eletrônicos, entre outros.

Na análise macroscópica, os especialistas observam dois grupos de características: organolépticas e anatômicas. As características organolépticas englobam gosto, densidade, odor, cor, grã, textura, brilho, dureza e desenhos na madeira. As características anatômicas podem ser observadas em três planos: transversal (perpendicular ao eixo do tronco), radial (paralelo ao eixo da haste) e tangencial (paralelo ao eixo do tronco). Nestes planos observam-se as camadas de crescimento, tipos de parênquima, vasos e raios (KOCH *et al.*, 2015).

Na análise microscópica, os especialistas observam as características dos tecidos e as estruturas celulares, tais como ornamentações da parede celular, tipos de pontuações, composição das células nos raios, dimensões das células, presença de cristais, entre outros (ZENID; CECCANTINI, 2007 e KOCH *et al.*, 2015).

Esta abordagem tradicional de identificação é uma tarefa difícil, não completamente automatizada e exige extrema concentração e conhecimento específico.

## 2.2 VISÃO COMPUTACIONAL

A Visão Computacional (VC) é uma área de estudo que engloba a construção de sistemas computacionais inteligentes que buscam simular a visão humana e podem ser aplicados em múltiplas áreas do conhecimento, como Agronomia, Biologia, Medicina, entre outras (NEVES; VIEIRA; GONZAGA, 2012).

De acordo com Bebis, Egbert e Shah (2003) e Neves, Vieira e Gonzaga (2012), a VC consiste na integração de técnicas de Processamento Digital de Imagens (PDI) e Inteligência Artificial (IA), possibilitando a análise, manipulação e interpretação de imagens por meio das características digitais nelas contidas

### 2.2.1 Processamento Digital de Imagens

O Processamento Digital de Imagens (PDI) é um campo da VC que trata da extração de características a partir de imagens digitais. Em outras palavras, o PDI consiste na entrada de uma imagem e a saída de um conjunto de valores que podem compor outra imagem ou serem interpretados por técnicas de IA (NEVES; VIEIRA; GONZAGA, 2012).

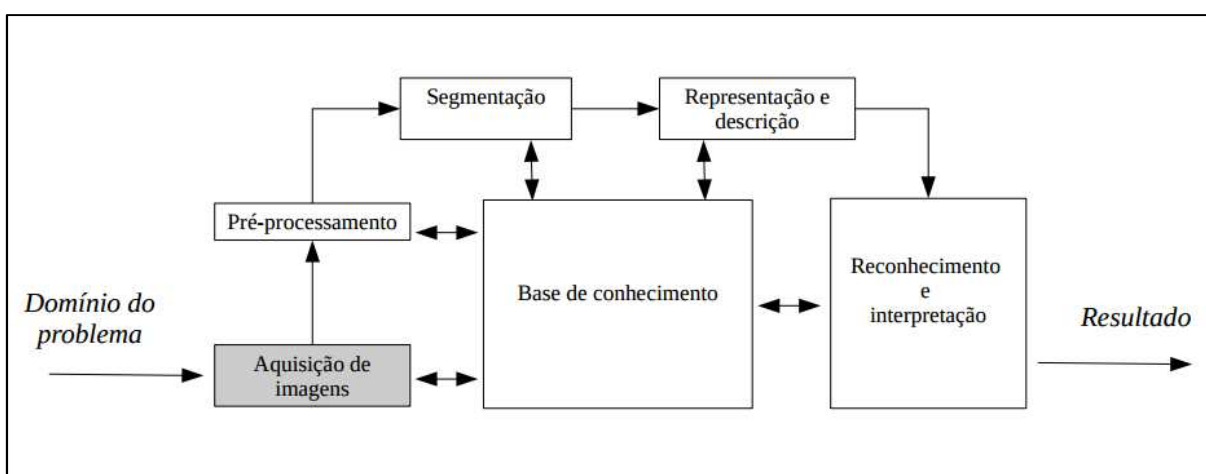
Uma imagem digital pode ser definida como uma função bidimensional  $f(x,y)$ , sendo  $x$  e  $y$  coordenadas do plano e a amplitude de  $f()$  é chamada de intensidade, para qualquer par de coordenadas. Esta é composta por números finitos e discretos em suas coordenadas e nos valores de intensidade, além de possuir um número

finito de elementos denominados *pixels*. Em outras palavras, trata-se da discretização de uma cena ou de uma imagem em formato analógico, resultando em uma matriz bidimensional (GONZALEZ; WOODS, 2010).

Segundo Akram, Dar e Quayoum (2010), as técnicas de PDI podem ser utilizadas para melhorar a informação visual por meio de transformações estruturais de uma imagem, a fim de facilitar posteriores processamentos ou simplesmente auxiliar na interpretação humana. Estas transformações podem ser realizadas, por exemplo, a partir da manipulação de dados nos canais de cores e no histograma de uma imagem digital.

### 2.2.1.1 Passos fundamentais em PDI

Segundo Gonzalez e Woods (2010) e Akram, Dar e Quayoum (2010), existem alguns passos fundamentais necessários para a realização de uma tarefa de PDI, são eles: Aquisição; Pré-processamento; Segmentação; Representação / Descrição e Interpretação. A Figura 5 apresenta o fluxo de execução destes passos.



**Figura 5 – Passos Fundamentais em PDI**  
Fonte: Gonzalez e Woods (2010).

De acordo com Gonzalez e Woods (2010) e Akram, Dar e Quayoum (2010), estes processos podem ser descritos como:

Aquisição: trata-se da obtenção de imagens digitais por meio de sensores. A aquisição pode ser por meio de câmeras digitais com sensores do tipo CCD (Charge

Coupled Device) com capacidade de digitalização. A função  $f(x,y)$ , mencionada na Seção 2.2.1, é o resultado do produto entre a luminância  $i(x,y)$  e a refletância do objeto  $r(x,y)$ , com a respectiva resposta da luz do ambiente (dia, noite, etc.).

Pré-processamento: fase utilizada para a aplicação de filtros específicos como realce de contrastes, remoção de ruídos, entre outros, para melhorar os aspectos da imagem obtida e facilitar as etapas seguintes, aumentando as chances de sucesso na atividade.

Segmentação: trata-se da divisão da imagem de entrada em objetos distintos, em que apenas os elementos de interesse na imagem devem ser segmentados para a etapa posterior. É uma das etapas mais complexas em PDI, pois uma segmentação mal definida pode comprometer o resultado final do processamento.

Representação / Descrição: fase também conhecida como Extração de Características, em que os objetos obtidos na segmentação devem ser convertidos a uma forma adequada de processamento. Com isso, é necessário tomar algumas decisões em relação a representação dos dados, de forma que estes sejam representados como fronteiras (cantos, pontos de inflexão, etc.) e/ou regiões completas (esqueleto, textura, etc.). Além disso, é necessário escolher métodos de descrição de dados para extrair as informações relevantes para serem utilizadas na fase seguinte.

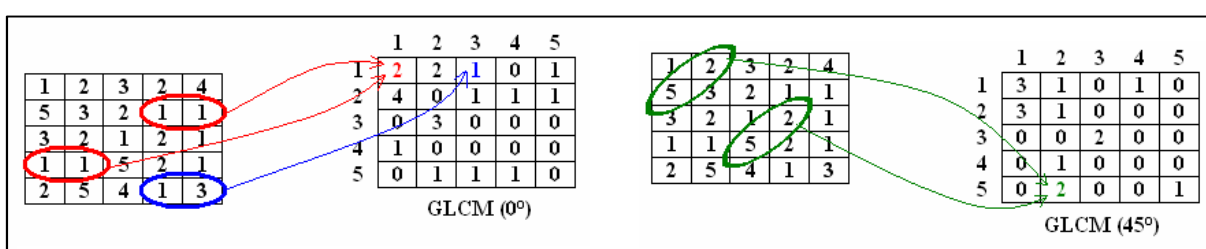
Reconhecimento / Interpretação: a finalidade desta fase é atribuir rótulos e dar significado aos objetos. A atribuição de rótulos denomina-se Reconhecimento e a atribuição de significado trata-se da Interpretação dos objetos reconhecidos.

É válido ressaltar que este é um modelo clássico de execução de uma tarefa em PDI, entretanto, estas etapas podem ser adaptadas ou eliminadas de acordo com as demais técnicas a serem utilizadas na solução do problema. A fase de Reconhecimento em PDI é fortemente associada às técnicas de IA, uma vez que os valores extraídos da imagem precisam passar por um processo de interpretação para transformar efetivamente os dados brutos (sem tratamentos prévios) em informações.

### 2.2.1.2 Extração de Descritores a partir de Matriz de Coocorrência

De acordo com Gonzalez e Woods (2010) e Haralick, Shanmugam e Dinstein (1973), existem diferentes técnicas de extração de descritores/características, seja por análise de cor, forma, textura, regiões, entre outras. Em se tratando de imagens com aspectos texturais, como a imagem de um corte de madeira por exemplo, a obtenção dos descritores pode ser realizada por meio de técnicas estatísticas, uma forma de obter tais descritores é com a realização de cálculos sobre uma matriz denominada Matriz de Coocorrência (Gray Level Co-occurrence Matrix - GLCM).

A GLCM trata-se de uma matriz quadrada em que seu tamanho é definido pela quantidade de níveis de cinza possíveis em uma imagem. Para imagens com pixels representados por 8 bits, existem 256 possíveis níveis de cinza ( $2^8 = 256$ ), portanto a GLCM terá dimensões de  $256 \times 256$ . Para cada combinação de  $x$  e  $y$ , calcula-se a quantidade de ocorrências simultâneas destes valores em toda a imagem e armazena-se na matriz. A Figura 6 (esquerda) apresenta um exemplo do cálculo de uma matriz de coocorrência, considerando o ângulo  $0^\circ$ . Neste caso, observam-se apenas 5 níveis de cinza e uma imagem de  $5 \times 5$  (OLIVEIRA *et al.*, 2012).



**Figura 6 – Matriz de Coocorrência (GLCM)**  
 Fonte: Adaptado de Martins *et al.* (2015).

Para calcular uma matriz de coocorrência, é necessário definir um ângulo  $\Theta$  ( $0^\circ$ ,  $45^\circ$ ,  $90^\circ$ ,  $135^\circ$ , ...) e uma distância  $D$  entre os pixels (1, 2, 3, ...). Quando um pixel vizinho é inexistente, isto é, quando a distância  $D$  em relação ao pixel atual extrapola as dimensões da imagem, a solução mais comum é ignorar a respectiva correlação (ITO, KIM e SALCEDO, 2009).

Ainda segundo Ito, Kim e Salcedo (2009), após a contagem das coocorrências, realiza-se a normalização dos valores da matriz, de forma a gerar uma matriz de probabilidades, na qual a soma das probabilidades resulta em 1. Tal normalização é dada pela equação (1), para a qual  $p(i,j)$  é o novo valor normalizado,  $i$  e  $j$  são os índices de linha e coluna do valor na matriz de coocorrência,  $N_g$  é a quantidade de níveis de cinza e  $P(i,j)$  é o valor antes de normalizar.

$$p(i,j) = \frac{P(i,j)}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} P(i,j)} \quad (1)$$

A normalização da matriz não é suficiente para representar e descrever uma imagem. Para isso, Haralick, Shanmugam e Dinstein (1973) propõem 14 descritores que consistem em cálculos sobre os valores da GLCM normalizada. Estes descritores são dados pelas equações apresentadas na Tabela 1.

**Tabela 1 – Equações dos 14 Descritores Texturais**

ID	Nome / Sigla	Fórmula
1	Energia ou Segundo Momento Angular (SMA)	$\sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} (P(i,j))^2$
2	Contraste (CON)	$\sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} (i-j)^2 P(i,j)$
3	Correlação (COR)	$\sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} P(i,j) \frac{(i-\mu)(j-\mu)}{\sigma^2}$
4	Soma dos Quadrados ou Variância (VAR)	$\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i-\mu)^2 p(i,j) \quad \mu = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} i \cdot P(i,j)$
5	Homogeneidade ou Momento da Diferença Inversa (MDI)	$\sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} \frac{P(i,j)}{1+(i-j)^2}$
6	Média da Soma (MDS)	$\sum_{i=2}^{2N_g} i p_{x+y}(i) \quad p_{x+y}(k) = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \delta_{i+j,k} p(i,j)$
7	Variância da Soma (VSM)	$\sum_{i=2}^{2N_g} (i - \text{ESM})^2 p_{x+y}(i)$

ID	Nome / Sigla	Fórmula
8	Entropia da Soma (ESM)	$\sum_{i=2}^{2N_g} p_{x+y}(i) \log(p_{x+y}(i))$
9	Entropia (ENT)	$\sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} P(i, j) \log P(i, j)$
10	Variância da Diferença (VDF)	Variância de $p_{x-y}(k) = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \delta_{ i-j ,k} p(i, j)$
11	Entropia da Diferença (EDF)	$\sum_{i=0}^{N_g-1} p_{x-y}(i) \log(p_{x-y}(i))$
12	Medida de Informação de Correlação 1 (MC1)	$\frac{ENT - HXY1}{\max(HX, HY)}$ $HX = - \sum_i p_x(i) \log(p_x(i)) = \text{entropy of } p_x$ $HY = - \sum_j p_y(j) \log(p_y(j)) = \text{entropy of } p_y$ $HXY1 = - \sum_i \sum_j p(i, j) \log(p_x(i)p_y(j))$
13	Medida de Informação de Correlação 2 (MC2)	$[1 - \exp(-2(HXY2 - ENT))]^{1/2}$ $HXY2 = - \sum_i \sum_j p_x(i)p_y(j) \log(p_x(i)p_y(j))$
14	Máximo Coeficiente de Correlação (MCC)	Raiz quadrada do segundo maior valor de Q, EM QUE $Q(i, j) = \sum_k \frac{p(i, k)p(j, k)}{p_x(i)p_y(k)}$

**Fonte: Adaptado de Haralick, Shanmugam e Dinstein (1973).**

Segundo Baraldi e Parmiggiani (1995), apenas 6 destes 14 descritores são os mais relevantes: SMA, CON, VAR, COR, ENT e MDI. Os demais não fazem diferença significativa na classificação.

Após a obtenção dos descritores, realiza-se o treinamento de um classificador (modelo de Inteligência Artificial), no qual parte dos conjuntos de descritores de GLCM são utilizados. Posteriormente, o classificador torna-se capaz de generalizar e reconhecer as classes de imagens desconhecidas.

## 2.2.2 Inteligência Artificial e Aprendizado de Máquina

A Inteligência Artificial (IA) é uma área da computação que busca compreender o pensamento humano e desenvolver sistemas artificiais inteligentes para a tomada de decisão e resolução de problemas complexos. É amplamente



utilizada em casos de mineração de dados, diagnósticos médicos, robótica, reconhecimento de voz, reconhecimento de imagem, processamento de linguagem natural, entre outros (NIU *et al.*, 2016).

De acordo com Goldschmidt (2010), os sistemas computacionais que utilizam IA podem ser subdivididos em: baseados em raciocínio (sistemas que simulam a forma humana de pensar) e baseados em comportamento (sistemas que simulam o comportamento humano e atuam como seres humanos). Estes são construídos por meio de processos de aprendizagem de máquina.

Segundo Jordan e Mitchell (2015), o Aprendizado de Máquina (AM) é uma área da IA que compreende o desenvolvimento de algoritmos, técnicas e programas que possibilitam que os computadores sejam capazes de adquirir conhecimento e aperfeiçoar seu desempenho com atividades complexas para a execução humana. É um dos campos mais crescentes, juntando conceitos de ciência da computação, matemática, estatística, física e ciência dos dados para a solução de problemas em diversas áreas, como saúde, biologia, fabricação, educação, finanças, segurança, marketing, entre outros.

Um sistema inteligente contempla a realização de tarefas simulando o raciocínio e comportamento humano, não se tratando apenas de dados decorados e comparações. O AM pode ser realizado por duas abordagens distintas:

- a) Aprendizado Dedutivo: trata-se da aplicação de uma regra geral para casos particulares, baseando-se no raciocínio dedutivo;
- b) Aprendizado Indutivo: trata-se de inferências lógicas para a obtenção de conclusões genéricas a partir de um conjunto particular de exemplos (GOLLIN, 1998).

Segundo Talwar e Kumar (2013) e Goldschmidt (2010), o aprendizado indutivo é base para os principais algoritmos de aprendizado existentes. Neste tipo, é comum que a obtenção do modelo de conhecimento seja abstraída de uma ou mais iterações sobre um conjunto de exemplos históricos. O Aprendizado Indutivo pode ser realizado de duas formas:

- a) Indutivo Supervisionado: nesta forma de aprendizado, os exemplos devem estar rotulados, isto é, devem conter qual o resultado esperado a ser produzido pelo modelo artificial. Desta forma, realiza-se a análise do resultado esperado com o resultado produzido pelo modelo;

- b) Indutivo Não Supervisionado: neste caso, os algoritmos realizam o agrupamento dos exemplos buscando alguma similaridade entre eles. Neste tipo de aprendizado a saída é previamente desconhecida.

De acordo com Rezende (2003) e Goldschmidt (2010), a IA possui cinco paradigmas clássicos de AM, os quais podem ser utilizados na construção de sistemas de forma individual ou híbrida. Os cinco paradigmas são:

Simbólico: o aprendizado acontece por meio de representações simbólicas, utilizando exemplos e contraexemplos. Geralmente acontece na forma de expressões lógicas, regras de produção, árvore de decisão ou rede semântica.

Estatístico: trata-se da utilização de modelos estatísticos para a aproximação do conceito induzido, podendo ser métodos paramétricos que assumem alguma forma de modelo a fim de encontrar os valores apropriados para estes parâmetros.

Baseado em exemplos: sistemas que classificam novos exemplos com base em exemplos conhecidos. Estes não utilizam os exemplos para indução de modelos, precisam estar armazenados previamente para que a classificação do novo exemplo aconteça.

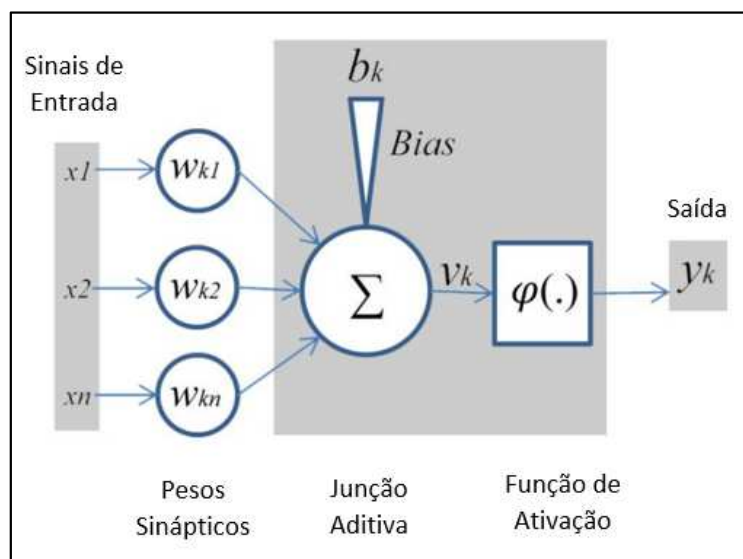
Evolucionário: baseia-se na evolução natural e na reprodução genética, simulando modelos biológicos para evoluir soluções e abstrair novos modelos que podem resolver problemas complexos.

Conexionista: trata-se da dinâmica das conexões entre pequenos nós chamados neurônios. Paradigma capaz de representar o conhecimento, simulando funções dos neurônios humanos. Para isso, foi desenvolvido o conceito de Redes Neurais Artificiais (RNA) capaz de simular características do cérebro humano como aprendizado, associação, generalização e abstração.

A seção a seguir aborda os conceitos de Redes Neurais Artificiais (RNA) de topologia Multi-Layer Perceptron (MLP). De acordo com Rezende (2003) e Goldschmidt (2010), tais redes são caracterizadas como estruturas de Aprendizado Indutivo Supervisionado Conexionista.

### 2.2.2.1 Redes neurais artificiais Perceptron Multicamadas

As Redes Neurais Artificiais (RNAs) são modelos matemáticos inspirados nas Redes Neurais Biológicas, as quais são pesquisadas desde 1943 e são utilizadas para a construção de sistemas inteligentes que podem solucionar problemas como reconhecimento de imagens, caracteres, voz, previsão de séries temporais, diagnósticos médicos, entre outros. Estes modelos são compostos por unidades simples de processamento conhecidas como Neurônios, os quais são organizados em camadas e interligados com outros neurônios através de sinapses (HAYKIN, 2001). A Figura 7 apresenta a estrutura de um neurônio artificial.



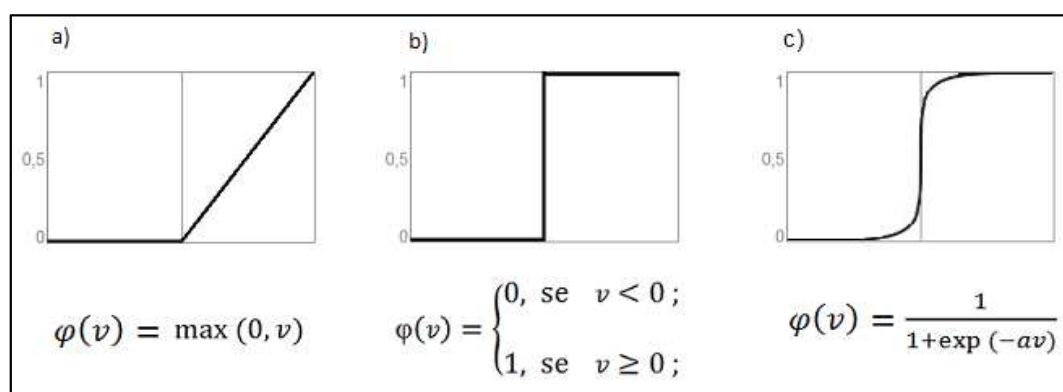
**Figura 7 – Modelo Matemático de um Neurônio Artificial**  
**Fonte: Taravat et al. (2015).**

Para melhor conceituar, Haykin (2001) descreve o neurônio artificial da Figura 7 por meio da Equação (2), em que  $k$  é o neurônio;  $x_1, x_2, \dots, x_m$  são os sinais de entrada;  $w_{k1}, w_{k2}, \dots, w_{km}$  são os pesos sinápticos para cada sinal de entrada;  $b_k$  é o bias (unidade extra na camada de entrada que serve para ajustar o valor da soma);  $v_k$  é o resultado da soma do  $b_k$  com os sinais de entrada ponderados pelos pesos sinápticos ( $w_k$ ) e  $y_k$  é o sinal de saída do neurônio  $k$ , após a aplicação da função de ativação sobre  $v_k$ .

$$v_k = \sum_{j=0}^m w_{kj} \cdot x_j + b_k \quad \text{e,} \quad y_k = \varphi(v_k) \quad (2)$$

De acordo com Haykin (2001), a função de ativação, representada por  $\varphi(v_k)$ , é aplicada para restringir o sinal de saída de um neurônio a um valor finito, em geral normalizado para o intervalo  $[0,1]$  ou  $[-1,1]$ . Existem algumas funções de ativação básicas mais utilizadas (Figura 8), são elas:

- Função Rectified Linear Unit (ReLU): para Goodfellow, Bengio e Courville (2016), é a função mais adequada na construção de Redes Neurais Convolucionais. Ela executa uma operação de limiar, em que qualquer valor de entrada inferior a zero é definido como zero;
- Função Limiar ou Degrau: consiste na normalização do sinal de saída em duas possibilidades de valores: 0 ou 1;
- Função Sigmoide Logística: é definida como uma função estritamente crescente.



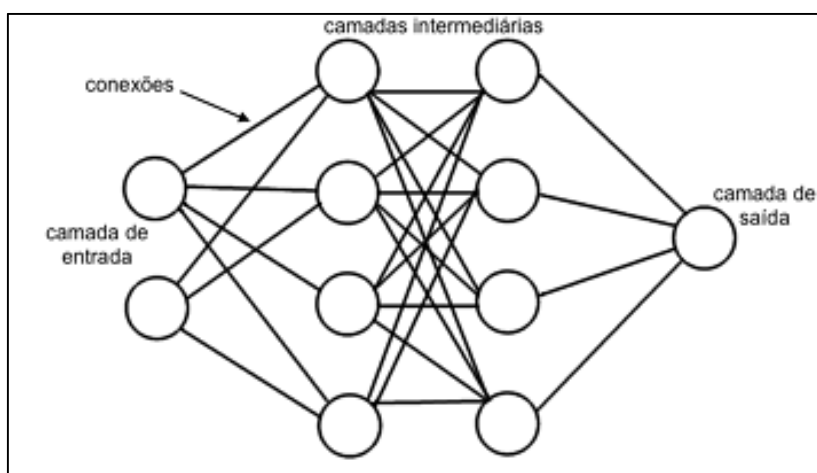
**Figura 8 – Funções de Ativação**  
Fonte: Adaptado de Haykin (2001).

Segundo Haykin (2001), para a construção de uma RNA completa é necessário estabelecer uma topologia que representará a estrutura da mesma, isto é, a distribuição dos neurônios, definição de camadas e suas conexões. Existem alguns modelos tradicionais, como o Perceptron, Adaline, *Multi-Layers*, entre outros.

A Perceptron Multicamadas (MultiLayer Perceptron - MLP) é uma topologia de RNA que serve de base para a técnica de AM com DL (GOODFELLOW; BENGIO; COURVILLE, 2016). Esta topologia consiste em um conjunto de unidades sensoriais que constituem a camada de entrada, uma ou mais camadas ocultas de

unidades computacionais e uma camada de saída de unidades computacionais. Os sinais inseridos na camada de entrada, são as características obtidas de um determinado problema, podendo ser desde dados geométricos de um determinado objeto até um conjunto de características extraídas a partir de uma tarefa de PDI (HAYKIN, 2001).

A representação da topologia MLP de RNA pode ser visualizada na Figura 9, na qual a arquitetura pode ser considerada como completamente conectada, em que cada neurônio se conecta com todos os neurônios da camada anterior.

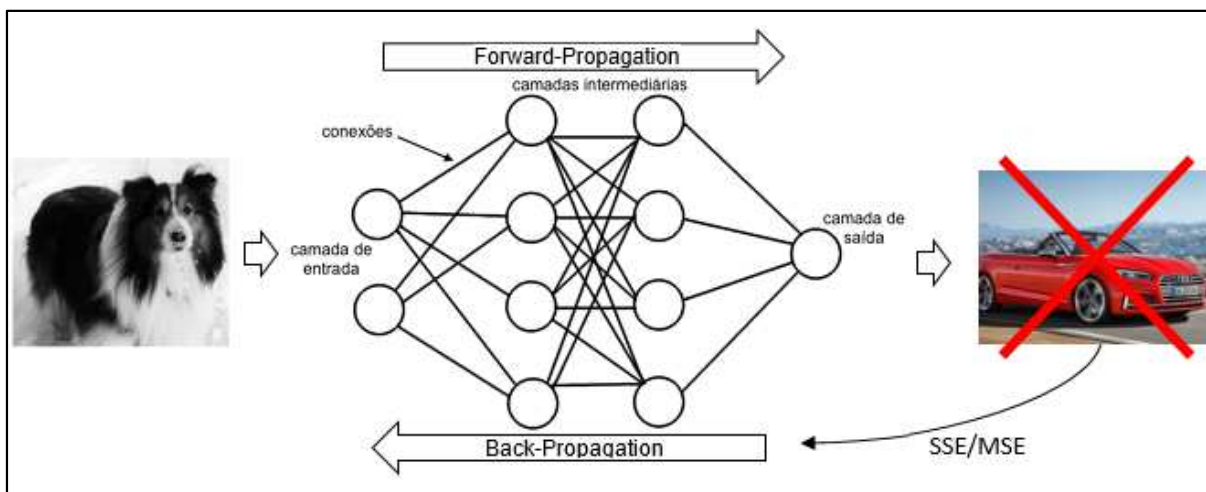


**Figura 9 – Topologia MLP de RNA**  
**Fonte: Adaptado de Haykin (2001).**

De acordo com Haykin (2001), para que uma RNA MLP funcione adequadamente, é necessário realizar o treinamento supervisionado por meio de um algoritmo de aprendizagem. Desta forma, a rede neural aprende com um conjunto de dados já conhecidos e torna-se capaz de generalizar respostas para entradas desconhecidas. Um dos algoritmos de aprendizagem que esteve presente em diversos casos de sucesso é o Back-Propagation, capaz de realizar o treinamento de uma RNA e resolver diversos problemas complexos.

O Back-Propagation (BP) é um algoritmo que foi criado em 1975, baseado na regra de aprendizado por correção de erro, no qual o treinamento é feito com um conjunto de dados de entrada e saída esperada. Os pesos sinápticos da RNA são iniciados aleatoriamente no intervalo de  $[-1, 1]$  e os sinais percorrem toda a rede até chegar na camada de saída, em que a resposta obtida é comparada com a resposta esperada. Para cada iteração (chamada também de época), é calculada a soma dos erros quadrados (SSE - Sum of Squared Errors) ou a média dos erros quadrados

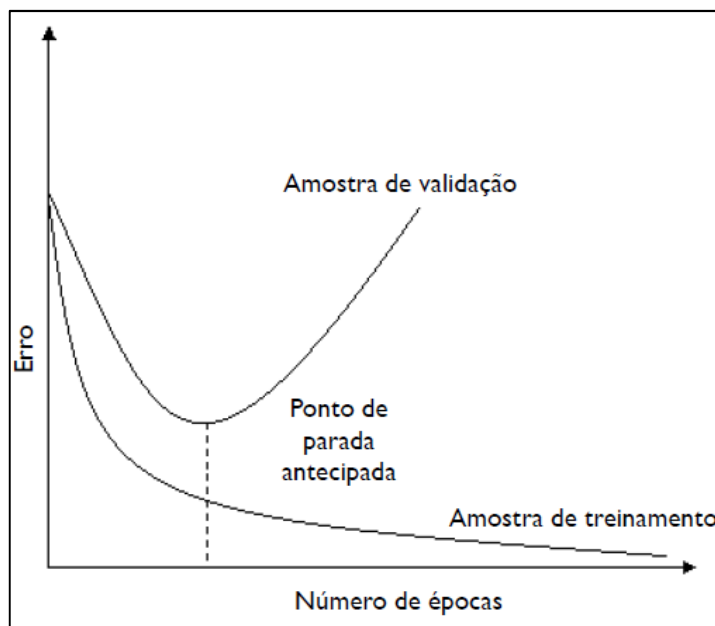
(MSE - Mean Squared Error) e é realizado o reajuste dos pesos sinápticos desde a camada de saída até a camada de entrada, utilizando o erro calculado como parâmetro de correção, como pode ser visualizado na Figura 10 (HAYKIN, 2001).



**Figura 10 – Treinamento da RNA e Algoritmo de BP**  
**Fonte: Adaptado de Haykin (2001).**

Uma das dificuldades em RNA é identificar o melhor ponto de parada de treinamento, isto é, estabelecer o melhor número de épocas para aprendizado. Quando o treinamento é iniciado, a taxa de erro é alta e diminui à medida em que novas épocas/iterações são realizadas. Entretanto, se houver treinamento em excesso da rede (*overfitting*), as entradas passarão a ser decoradas e não mais aprendidas efetivamente. Em outras palavras, a rede ficará tão ajustada ao conjunto de dados de treinamento que não será mais capaz de generalizar, tornando-se ineficiente para novos conjuntos de dados (HAYKIN, 2001).

Segundo Haykin (2001), para buscar a melhor generalização da rede, uma alternativa é realizar a validação do modelo durante o treinamento. Tal validação consiste em um conjunto de dados diferente do utilizado para treinamento para estabelecer um ponto de parada antecipada com base na comparação dos erros de validação e a taxa de aprendizado. Com isso, é possível estabelecer o ponto de parada (número de épocas) ideal, no qual a rede tem uma boa taxa de treinamento e baixa taxa de erros de validação, como pode ser visualizado na Figura 11.



**Figura 11 – Validação de Treinamento**  
Fonte: Adaptado de Haykin (2001).

A partir dos conceitos clássicos desta topologia de RNA, foram construídas arquiteturas cada vez mais complexas de aprendizado profundo e agrupadas em um conjunto de técnicas denominado Deep Learning. Tais arquiteturas e conceitos são abordados na seção a seguir.

### 2.2.3 Deep Learning

Segundo Goodfellow, Bengio e Courville (2016), as abordagens tradicionais de AM permitem a construção de sistemas inteligentes que simulam o conhecimento humano. Desta forma, as tarefas que são intelectualmente difíceis para os humanos podem não ser tão difíceis para o computador. Entretanto, para que um sistema adquira o conhecimento e execute tarefas, os problemas precisam ser descritos formalmente, por meio de regras matemáticas, treinamentos por representações de características específicas, árvores de decisão, regras de produção, etc.

O grande desafio na IA é lidar com problemas que são fáceis para pessoas executarem, porém difíceis de serem descritos. Normalmente, são problemas que as pessoas resolvem intuitivamente, como reconhecer a letra de determinada pessoa, determinar rostos em imagens, etc., mas não conseguem representar uma lógica

completamente formal. Como a representação e descrição de características é fundamental para abordagens tradicionais de AM, se o desenvolvedor não conseguir representar muito bem quais características de um determinado conjunto de dados são importantes para a aprendizagem, a precisão dos resultados finais poderá estar comprometida (SAFARI, 2013 e GOODFELLOW; BENGIO; COURVILLE, 2016).

Deep Learning (DL) é um conjunto de técnicas para solução de problemas humanamente intuitivos e subjetivos, isto é, quando o problema não é tão fácil de se descrever. Com isso, a técnica manipula os dados brutos (sem tratamentos prévios) em abordagens profundas, procurando compreender o problema através de hierarquia de características e suas inter-relações. Assim, evita-se a dependência de definição de técnicas por parte dos especialistas humanos. Essa hierarquia pode conter muitas camadas para compreensão do problema, por isso é chamada de aprendizagem profunda ou DL (LECUN; BENGIO; HINTON, 2015).

Para exemplificar um cenário de aprendizagem tradicional, observa-se o problema de um jogo de Xadrez:

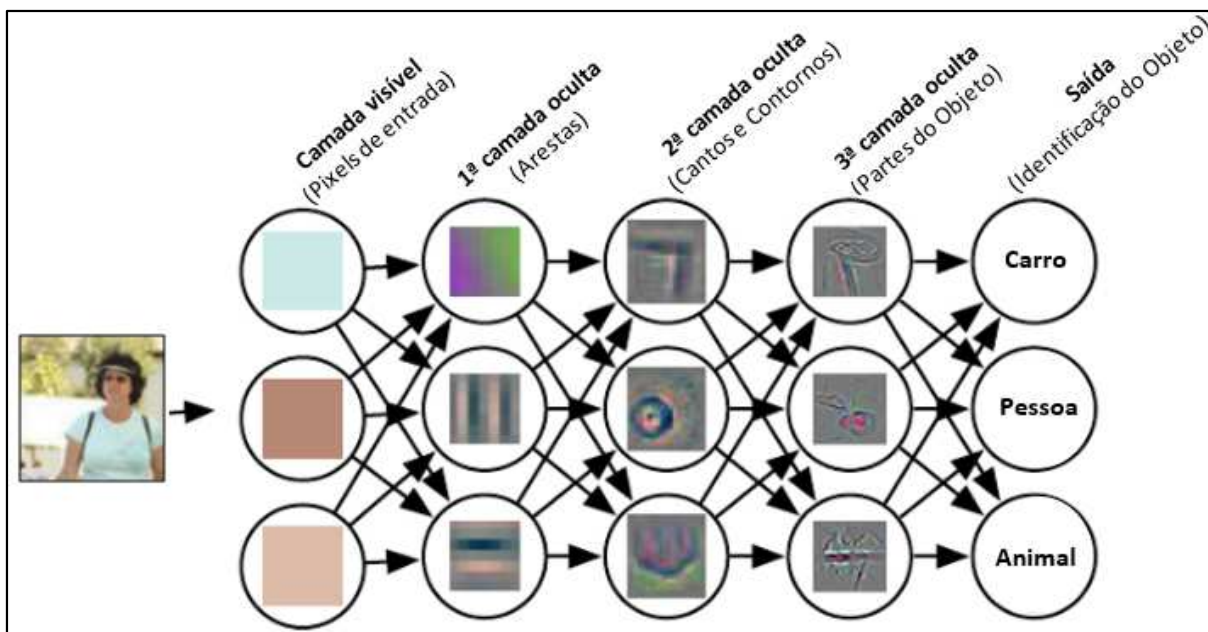
O Xadrez é, naturalmente, um cenário muito simples, contendo apenas sessenta e quatro locais e trinta e duas peças que podem transitar em apenas caminhos rígidos e circunscritos. Conceber uma estratégia de xadrez bem-sucedida é um enorme desafio, mas o desafio não se deve à dificuldade de descrever o conjunto de peças de xadrez e movimentos permitidos para o computador. O jogo pode ser completamente descrito por uma breve lista de regras formais fornecida com antecedência pelo programador (GOODFELLOW; BENGIO; COURVILLE, 2016, p. 2).

Para exemplificar um cenário de aprendizagem profunda, observa-se o problema de detecção de carros em fotografias:

Sabemos que os carros têm rodas, então gostaríamos de usar a presença de uma roda como um recurso. Infelizmente, é difícil descrever exatamente o que roda parece em termos de valores de pixel. Uma roda tem uma forma geométrica simples, mas sua imagem pode ser distorcida por sombras, o brilho do sol, as peças metálicas da roda, o para-choque do carro ou um objeto em primeiro plano obscurecendo parte da roda, e assim por diante (GOODFELLOW; BENGIO; COURVILLE, 2016, p. 3).

A Figura 12 mostra como um sistema baseado em DL pode representar o reconhecimento de imagem de uma pessoa, combinando características simples, como cantos e contornos processados em uma estrutura de camadas distribuídas de forma hierárquica, em que os elementos de cada camada se conectam.





**Figura 12 – Conceito de Reconhecimento de Imagens com *Deep Learning***  
**Fonte: Adaptado de Goodfellow, Bengio e Courville (2016).**

Segundo Arel, Rose e Kamowski (2010), existe uma técnica de DL específica para reconhecimento de imagens, as Redes Neurais Convolucionais. Uma Rede Neural Convolutiva é uma estrutura especializada em processamento de sinais brutos (imagens, áudios, etc.), para os quais são extraídas características por meio de uma série de convoluções aplicadas a partir de filtros com pesos convolucionais previamente aprendidos no processo de treinamento do modelo.

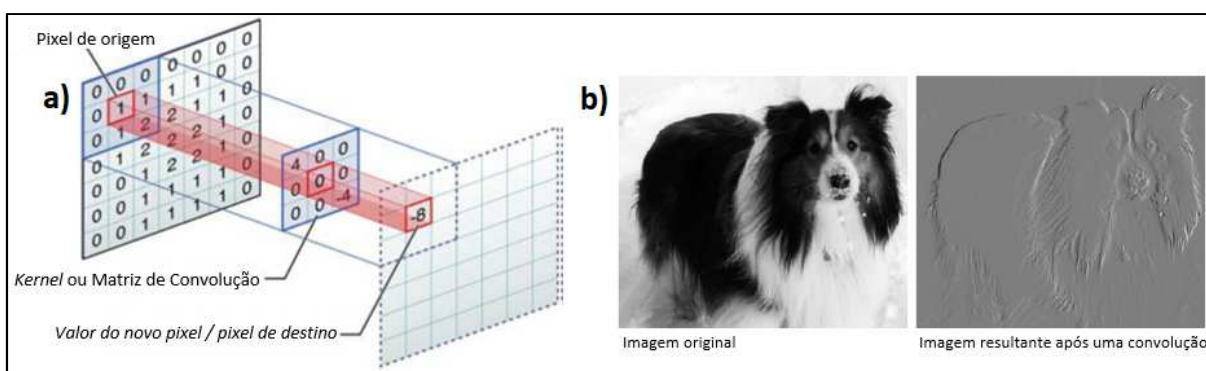
### 2.2.3.1 Redes neurais convolucionais

Uma Rede Neural Convolutiva (*Convolutional Neural Network – CNN*) é um tipo de RNA baseada na topologia MLP (vide Seção 2.2.2.1) e especializada em processamento de dados brutos (sem tratamentos prévios). Uma CNN consiste na realização de convoluções em um determinado conjunto de dados, aplicando filtros previamente aprendidos para a abstração de características. Este tipo de rede tem sido utilizado amplamente em aplicações de classificação, detecção e reconhecimento em imagens e vídeos (VARGAS; CARVALHO; VASCONCELOS, 2016 e AREL; ROSE; KAMOWSKI, 2010).

Segundo Goodfellow, Bengio e Courville (2016), a convolução ( $\otimes$ ) é uma operação que realiza a transformação linear entre uma matriz primária e outra denominada *kernel* ou matriz de convolução. Realiza-se a superposição do *kernel* sobre a matriz primária em função do deslocamento entre elas, gerando uma terceira matriz (vide Figura 13). A matriz primária pode ser uma imagem, uma vez que esta refere-se a uma coleção bidimensional de pixels.

Um *kernel* deve conter um tamanho  $n \times m$ , seus valores em  $f(n,m)$ , o valor  $x$  do salto/passo de deslocamento e o elemento central que deve ser posicionado sobre o primeiro pixel da imagem. Assim, realiza-se a soma dos produtos obtidos entre os elementos que se sobrepõem, gerando o novo valor para a respectiva posição do pixel original na matriz de saída. Em seguida, o *kernel* é deslocado  $x$  pixels e a mesma operação é realizada até que toda a imagem seja percorrida e as novas saídas sejam geradas (GOODFELLOW; BENGIO; COURVILLE, 2016).

A Figura 13 (a) representa uma convolução utilizando um *kernel* de tamanho  $3 \times 3$ , no qual os valores dos elementos são, respectivamente, 4; 0; 0; 0; 0; 0; 0; 0; -4. A Figura 13 (b) representa uma possível imagem de saída após a realização de uma convolução utilizando um filtro aprendido. Estas imagens de saída são denominadas mapas de características.



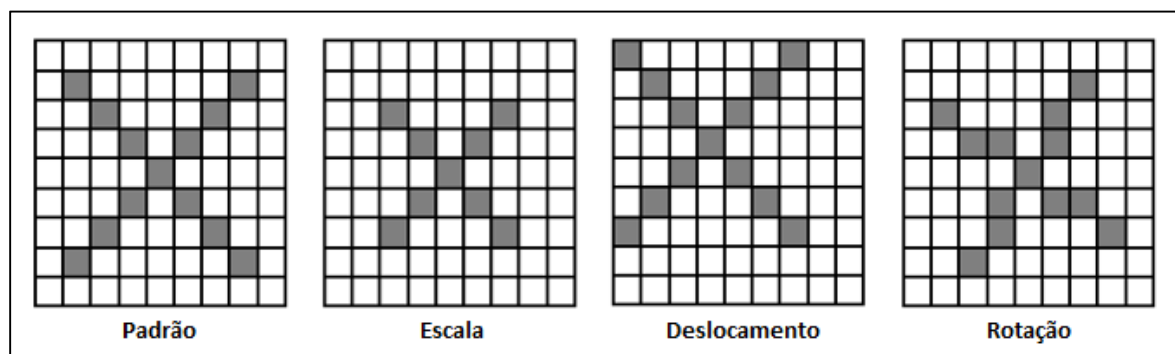
**Figura 13 – Convolução em Imagens**

Fonte: Adaptado de Neves, Vieira e Gonzaga (2012) e Goodfellow, Bengio e Courville (2016).

A fórmula da convolução pode ser representada pela Equação (3), na qual um pixel  $f(x,y)$  de uma imagem é multiplicado pelo respectivo elemento  $W_i$  de um *kernel* de tamanho  $k$  (sendo  $k = n \times m$ ) em que a soma destes produtos gera um novo pixel  $g(x,y)$ .

$$g(x, y) = \sum_{i=1}^k w_i \cdot f(x, y) \quad (3)$$

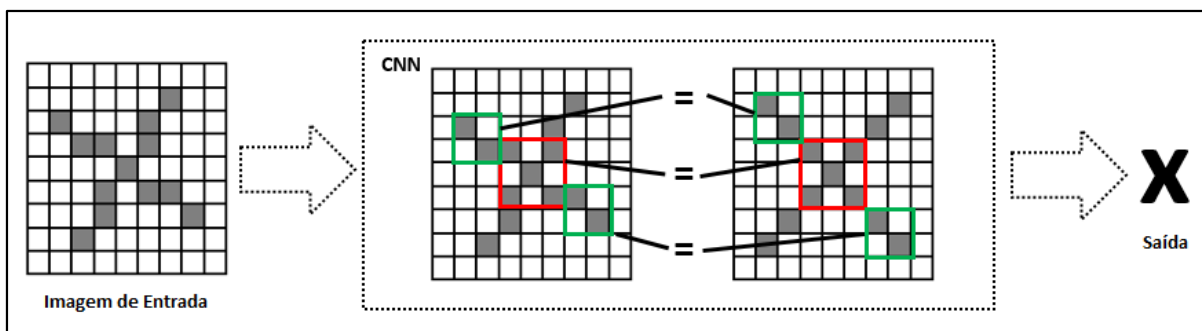
O grande desafio em tarefas que envolvem PDI e IA é lidar com as variações de características existentes nas imagens, como deslocamento, escala, rotação, oclusão, entre outros (vide Figura 14). Para uma RNA tradicional, estas variações precisam ser muito bem tratadas no PDI por meio da aplicação de filtros e utilização de descritores de forma, para a extração de características mais representativas. Desta forma, na fase de processamento existe uma dependência de predefinição destas técnicas para que as características extraídas sejam reconhecidas e classificadas corretamente pela RNA (MANSANO; MATSUOKA; AFONSO, 2012).



**Figura 14 – Variações de Características**  
**Fonte: Autoria Própria.**

De acordo com Goodfellow, Bengio e Courville (2016), em uma abordagem com DL as variações de características são tratadas de uma maneira diferente dos métodos tradicionais. A própria CNN realiza a análise de partes da imagem de entrada e resume as informações em regiões, de forma que a maioria das saídas não mudem. Este tratamento consiste muito mais em determinar se uma característica está presente na imagem do que onde exatamente ela está.

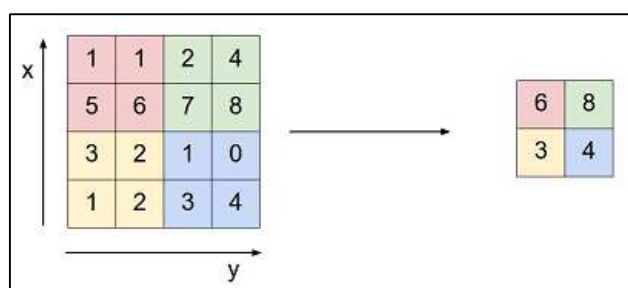
É claro que este conceito funciona para pequenas variações na imagem. Sendo assim, a realização do treinamento com variações maiores é essencial para que o modelo apresente menos erros. A Figura 15 apresenta um modelo conceitual de reconhecimento de imagens com CNN utilizando representações invariantes.



**Figura 15 – Representações Invariantes em CNN**  
**Fonte: Autoria Própria.**

Para que haja esse tratamento de representação invariante, utiliza-se uma técnica denominada *Pooling* ou Agrupamento. Trata-se de uma maneira de resumir uma imagem preservando suas informações mais importantes. Em outras palavras, uma função de *Pooling* consiste em substituir uma saída de uma parte filtrada da imagem por uma saída resumida em relação aos valores vizinhos, tornando a rede mais robusta e de menor custo computacional.

De acordo com Vargas, Carvalho e Vasconcelos (2016), uma das funções de agrupamento mais utilizadas e indicadas é o *Max-Pooling* (Figura 16, esquerda). Após a obtenção de mapas de características a partir das convoluções, cada mapa é dividido em sub-regiões ( $m \times n$ ) e, para cada sub-região, escolhe-se o valor máximo do conjunto para compor uma nova saída, na qual cada elemento é o máximo de cada sub-região (Figura 16 - direita). Com isso, mesmo que uma imagem esteja um pouco deslocada, a CNN é capaz de capturar as características essenciais da imagem para realizar uma classificação.



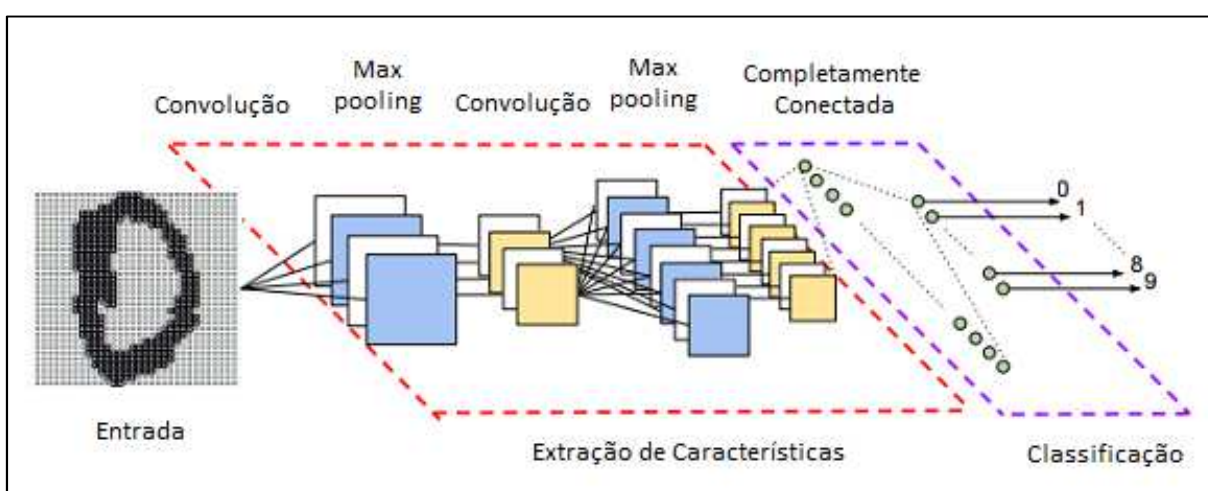
**Figura 16 – Função Max-Pooling**  
**Fonte: Adaptado de Bezerra, 2016.**

Assim como a MLP, uma CNN é composta por uma camada de entradas, várias camadas intermediárias e uma camada de saída. Para normalizar os valores de saída das camadas intermediárias, Nair e Hinton (2010) mencionam o uso da

função de ativação ReLU (vide Seção 2.2.2.1) como alternativa mais adequada para CNN do que o uso de funções sigmóides muito utilizadas em RNAs tradicionais.

As camadas intermediárias deste tipo de rede podem ser conceitualmente divididas em dois grupos (Figura 17): extração de características e classificação. No grupo de extração de características estão localizadas as camadas de convolução, *pooling* e funções de ativação (ReLU). No grupo de classificação encontra-se uma ou mais camadas completamente conectadas que atuam como um classificador da rede, tomando a decisão final sobre o processo de reconhecimento. Para a representação das probabilidades das saídas na última camada, normalmente utiliza-se *SoftMax* como função de ativação (BEZERRA, 2016 e VARGAS; CARVALHO; VASCONCELOS, 2016).

Conforme ilustrado na Figura 17, após a camada de entrada (representada pela imagem), observa-se uma camada de convolução onde ficam os mapas de características gerados a partir de filtros aprendidos. Cada filtro convolucional é composto de vários neurônios clonados, mas visualizam características diferentes da imagem de entrada. Em analogia a uma MLP, cada célula de um filtro equivale ao peso do respectivo neurônio. Em seguida, observa-se uma camada de *pooling*, outra de convolução e outra de *pooling*, completando o grupo de extração de características. No segundo grupo conceitual, encontram-se duas camadas completamente conectadas, as quais realizam a classificação final da rede.



**Figura 17 – Rede Neural Convolutional**  
**Fonte: Vargas, Carvalho e Vasconcelos (2016).**

A obtenção de um mapa de características por meio de uma convolução utilizando um filtro aprendido conceitua-se como Conectividade Local, na qual um neurônio em uma camada está conectado a um subconjunto de saídas da camada anterior. Em outras palavras, cada neurônio está ligado a uma sub-região da camada anterior e estas sub-regiões são conhecidas como Campos Respectivos Locais (GOODFELLOW; BENGIO; COURVILLE, 2016).

Assim como em uma RNA MLP, é necessário realizar o treinamento supervisionado da CNN para que ela seja capaz de generalizar as respostas para imagens desconhecidas. A seção a seguir aborda como este treinamento deve ser realizado e quais são as técnicas recomendadas para este tipo de rede neural.

### 2.2.3.2 Treinamento de uma CNN

Em um processo tradicional de VC, utiliza-se as etapas fundamentais de PDI para definição dos filtros adequados a serem aplicados sobre a imagem original para a extração de características relevantes. A definição dos filtros varia de acordo com o tipo de imagem a ser tratada, dependendo fortemente da interação do programador para essa tomada de decisão. Posteriormente, utiliza-se uma MLP que classifica este conjunto de características. Já em um processo de VC utilizando CNN, a definição dos filtros para extração das características é realizada dinamicamente pela arquitetura na fase de treinamento. É necessário realizar somente definições como a quantidade de filtros por camada, os passos e tamanhos (VARGAS; CARVALHO; VASCONCELOS, 2016).

Ao longo do treinamento, o processo de aprendizado altera os pesos (valores nos filtros) até encontrar os melhores valores para o conjunto de imagens utilizado. Geralmente, as primeiras camadas costumam aprender filtros de cores, arestas e bordas e as camadas mais profundas tornam-se especialistas em características complexas, detectando detalhes muito específicos que, muitas vezes, não são abordados pela metodologia clássica de VC (ZEILER; FERGUS, 2014).

Para a realização do treinamento, torna-se necessário a definição da arquitetura de CNN a ser utilizada, isto é, de que maneira as camadas estão dispostas, profundidade da rede, etc. De acordo com Vargas, Carvalho e

Vasconcelos (2016), existem algumas arquiteturas já utilizadas no meio científico, como a *AlexNet* e a *VGGNet*.

Segundo Lecun, Bengio e Hinton (2015), o processo de treinamento é similar ao de uma MLP, no qual normalmente utiliza-se o algoritmo de *BP*. Para uma CNN, o treinamento final pode ser realizado após um processo de pré-treinamento, o qual evita complicadores como dissipação dos gradientes (*vanishing gradients*).

A dissipação dos gradientes é um problema que acontece durante a aplicação do algoritmo de *BP* em redes muito profundas. Se a unidade de ativação de uma determinada camada está próxima da saturação, então a derivada parcial resultante será um valor próximo de zero. Isso significa que, à medida que a retro propagação vai chegando às camadas mais próximas da entrada, os valores dos gradientes do erro serão cada vez mais próximos de zero e as primeiras camadas serão corrigidas de forma extremamente lenta (BEZERRA, 2016).

Segundo Bezerra (2016), o objetivo do pré-treinamento é diminuir o erro de generalização dos modelos resultantes e também o tempo de treinamento a ser realizado, convergindo mais rapidamente o aprendizado do modelo. O processo consiste na iniciação dos valores dos pesos da rede de forma otimizada, tal que facilite a propagação de correções pelo algoritmo de *BP*. As camadas de extração de características são adicionadas à rede iterativamente e os pesos são definidos de forma parcial camada a camada (NAIR; HINTON, 2010).

Após o pré-treinamento, as camadas completamente conectadas (de classificação) são adicionadas à rede e, neste caso, os respectivos pesos são iniciados aleatoriamente. Desta forma, têm-se a CNN completa e pré inicializada com os valores não tão aleatórios, pronta para a realização do treinamento oficial e obtenção do modelo resultante.

### 2.2.3.3 Dropout, Transfer Learning e Data Augmentation

Para que uma CNN se torne mais robusta no treinamento, é possível parametrizar um valor de desligamento aleatório de alguns neurônios, de forma que outros não confiem demais na especialização de determinados neurônios. Esta técnica é chamada de *DropOut* e faz com que sempre que uma entrada é

apresentada à rede, a mesma se comporte de maneira diferente, o que ajuda a evitar o *overfitting*. Desta forma, mesmo que alguns neurônios não detectem determinadas características em uma classificação, outros serão capazes de fazê-lo (KRIZHEVSKY *et al.*, 2012).

De acordo com Huqqania *et al.* (2013) e Huang *et al.* (2017), o treinamento de uma CNN é um processo custoso, computacionalmente falando, e pode demorar muitos dias em uma unidade de processamento comum. Com isso, para agilizar a generalização de novos modelos, foi estabelecido um conceito denominado *Transfer Learning*. Tal conceito consiste na obtenção da arquitetura contendo pesos pré-inicializados, os quais extraem características suficientes para serem utilizadas em novos modelos. Algumas das arquiteturas mais conhecidas que possuem pré-inicializações disponibilizadas são apresentadas na Seção 2.3.2.

Neste contexto, tal conceito ainda aborda a possibilidade de desacoplar as últimas camadas completamente conectadas da arquitetura e acoplar uma ou mais camadas que representem a quantidade de saídas necessárias para o novo problema abordado. Com isso, o pesquisador pode congelar uma ou todas as camadas convolucionais e treinar (habilitar o *BP*) somente nas camadas finais, que realizam a classificação final. Desta forma, aproveita-se o conhecimento já existente na arquitetura e agiliza-se o tempo de treinamento (HUANG *et al.*, 2017).

Para que uma CNN consiga generalizar melhor, o processo de treinamento precisa ser realizado com uma grande quantidade de exemplos. Quando o pesquisador não dispõe de uma base de testes muito grande, torna-se necessário realizar transformações para aumentar o número de exemplos em sua base, este conceito é chamado de *Data Augmentation*. Em uma base de imagens, as transformações podem ser: fragmentação, rotação, translação, escala, manipulação de cores, entre outros (LAPTEV; BUHMANN, 2015). A grande vantagem é aumentar artificialmente o tamanho do conjunto de dados com variações na representação dos exemplos, aumentando as chances de sucesso ao realizar o treinamento.

De acordo com Huqqania *et al.* (2013), uma das alternativas para agilizar o treinamento de redes profundas é a utilização de uma ou mais Unidades de Processamento Gráfico (Graphics Processing Unit – GPU), utilizando conceitos de computação paralela para obter melhor desempenho. Além disso, os dados de treinamento normalmente são carregados e treinados em lotes (Batches ou Minibatches), de forma a não extrapolar o uso de memória.



A empresa NVIDIA disponibiliza uma plataforma denominada Compute Unified Device Architecture (CUDA), a qual permite a integração de linguagens de programação com as instruções executadas pela GPU. Segundo Izotov *et al.* (2011), esta estrutura permite o treinamento de uma CNN com tempo muito menor em relação à Unidade Central de Processamento (Central Processing Unit – CPU).

Muitas das arquiteturas que fazem parte do Estado da Arte em CNN foram treinadas em várias unidades de GPU para otimização de desempenho (KRIZHEVSKY *et al.*, 2012 e RUSSAKOVSKY *et al.*, 2015). Algumas destas arquiteturas são descritas na Seção 2.3.2.

## 2.3 ESTADO DA ARTE

Esta seção descreve alguns dos projetos que representam o Estado da Arte em reconhecimento de madeiras utilizando Visão Computacional, bem como as principais arquiteturas de CNN.

Dentre os autores apresentados a seguir, destacam-se os pesquisadores Paula *et al.* (2014) e Martins *et al.* (2015), os quais realizaram as aquisições das imagens e disponibilizaram as bases utilizadas na presente pesquisa.

### 2.3.1 Reconhecimento de Madeiras

Não são muitos os projetos de pesquisa que abordam o reconhecimento automático de espécies florestais. Paula *et al.* (2014) mencionam apenas três grupos de pesquisa: UTM – CAIRO (Malásia) - Análises em imagens macroscópicas; UTAR – CVIS (Malásia) - Análises em imagens macroscópicas e UFPR – VRI - Análises em imagens macro e microscópicas.

Além de algumas publicações com outras bases de imagens de espécies florestais, entre os anos de 2015 e 2016, foram identificados outros experimentos que utilizaram as mesmas bases de imagens utilizadas nesta pesquisa. Algumas destas publicações também abordam técnicas de DL, mais especificamente CNN, e

apresentam resultados bastante relevantes em relação a outros trabalhos que utilizaram técnicas tradicionais. A subseção a seguir apresenta uma breve descrição das principais pesquisas, as técnicas utilizadas e as taxas de reconhecimento obtidas pelas referidas publicações.

### 2.3.1.1 Pesquisas correlatas

Paula *et al.* (2014) desenvolveram uma pesquisa abordando a identificação de espécies florestais a partir de 2942 imagens macroscópicas de madeiras de 41 espécies diferentes adquiridas por equipamentos de baixo custo. Em tal pesquisa, a base de imagens macroscópicas foi construída por aquisição em campo a partir de espécies encontradas em território brasileiro.

A proposta destes pesquisadores não foi a construção de um software, mas sim um protocolo baseado em uma estratégia denominada “dividir para conquistar”. Nesta estratégia as imagens foram divididas em subpartes, de forma que alguns ruídos locais não comprometessem a classificação final. A partir das subimagens, foram extraídas características como cor e textura para posterior classificação.

Foram utilizadas técnicas de extração de características, como análises de cor, GLCM, histograma de borda, Fractais, LBP (Local Binary Pattern), LPQ (Local Phase Quantization) e Gabor, além de técnicas de classificação com SVM (Support Vector Machines) com tomada de decisão baseada em *pairwise* (para múltiplas classes) para a construção dos modelos de conhecimento, realização de testes e validações. Posteriormente, os resultados das classificações parciais passaram por fusões de baixo e alto nível para se chegar à classificação final da espécie, proporcionando taxa de reconhecimento de 97,77%.

Martins *et al.* (2015) desenvolveram uma pesquisa abordando a identificação de espécies florestais a partir de 2240 imagens de madeiras de 112 espécies diferentes em âmbito microscópico. A proposta não foi a construção de um produto final, mas sim um protocolo de classificação baseado em seleção e combinação de classificadores no espaço de dissimilaridade para a identificação de características texturais das imagens.

Foram utilizadas técnicas de extração de características, como SIFT (Scale Invariant Feature Transform), SURF (Speed-Up Robust Feature), LBP, LPQ, GLCM e Gabor, além de técnicas de classificação com SVM para a construção dos modelos de conhecimento, realização de testes e validações.

A melhor taxa de reconhecimento obtida com os diferentes classificadores foi de 93,03%. Em outras palavras, as taxas de reconhecimento em nível microscópico também foram muito bem-sucedidas e a base de imagens foi disponibilizada para a realização de trabalhos futuros.

A partir da base de imagens microscópicas elaborada por Martins *et al.* (2015), Cavalin, Kapp e Oliveira (2015) propuseram uma abordagem adaptativa com combinação de classificações utilizando diferentes grupos de características extraídas. A pesquisa teve por finalidade obter um modelo capaz de apresentar bons resultados com o menor custo e complexidade computacional.

Utilizando as bases de imagens macro e microscópicas elaboradas por Paula *et al.* (2014) e Martins *et al.* (2015), alguns pesquisadores realizaram também alguns experimentos utilizando CNN.

Andrearczyk e Whelan (2016) construíram uma arquitetura de CNN com duas camadas convolucionais e três camadas completamente conectadas e nomearam como Texture CNN (T-CNN). A ideia foi elaborar um modelo capaz de extrair características de textura automaticamente e comparar a eficiência das técnicas de extração aprendidas com as técnicas de extração pré-estabelecidas pelo pesquisador em abordagens mais tradicionais. Para tal pesquisa, foi realizado o treinamento completo da estrutura (sem *Transfer Learning*) e o melhor resultado para a base de imagens macroscópica foi de 95,77% de taxa de reconhecimento e 97% para a base microscópica.

Hafemann, Cavalin e Oliveira (2014) também construíram uma arquitetura de CNN com duas camadas convolucionais e três camadas completamente conectadas. Foi realizado o treinamento completo da estrutura (sem *Transfer Learning*) e o melhor resultado para a base de imagens macroscópica foi de 95,77% de taxa de reconhecimento e 97,32% para a base microscópica.

Hafemann, Cavalin e Oliveira (2015) também desenvolveram outro experimento com a base de imagens macroscópicas. Em tal experimento, um modelo foi inicialmente treinado para a classificação de uma base de texturas

denominada Brodatz-32<sup>1</sup> e, então, foi realizado o *Transfer Learning* para obter um novo modelo específico para a classificação de imagens de madeiras. Para o melhor modelo, a taxa de reconhecimento foi de 97,04%.

O Quadro 1 apresenta a relação geral dos trabalhos dos diferentes pesquisadores mencionados anteriormente, os quais utilizaram as mesmas bases de imagens em utilizadas na presente pesquisa.

**Quadro 1 – Síntese das Pesquisas Correlatas (Bases da Presente Pesquisa)**

PESQUISADOR(ES)	TÉCNICAS	RECONHECIMENTO
PAULA <i>et al.</i> (2014)	Análises de cor, GLCM, histograma de borda, Fractais, LBP, LPQ e Gabor / SVM	<b>97,77% (Macro)</b>
MARTINS <i>et al.</i> (2015)	SIFT, SURF, LBP, LPQ, GLCM e Gabor / SVM	93,03% (Micro)
CAVALIN, KAPP E OLIVEIRA (2015)	Abordagem adaptativa com combinação de classificações	% não disponível (Micro)
HAFEMANN, CAVALIN E OLIVEIRA (2014)	CNN	95,77% (Macro) <b>97,32% (Micro)</b>
HAFEMANN, CAVALIN E OLIVEIRA (2015)	CNN – Transfer Learning de Texturas	97,04% (Macro)
ANDREARCZYK E WHELAN (2016)	Texture CNN (T-CNN)	95,77% (Macro) 97,00% (Micro)

**Fonte: Autoria Própria.**

O Quadro 2 apresenta a relação de algumas pesquisas mencionadas por Paula *et al.* (2014) e Martins *et al.* (2015), as quais também abordam o reconhecimento automático de espécies, porém a partir de outras bases de imagens.

**Quadro 2 – Síntese das Pesquisas Correlatas (Outras Bases)**

PESQUISADOR(ES)	TÉCNICAS	RECONHECIMENTO
TOU <i>et al.</i> (2007)	GLCM / RNA	72,00% (Macro)
TOU <i>et al.</i> (2008)	GLCM e 1DGLCM / RNA e K-nn	72,80% (Macro)
TOU <i>et al.</i> (2009)	GLCM, filtros Gabor, comb. GLCM e Gabor, matriz de covariância e <i>varication-based recognition</i> / K-nn	85,00% (Macro)
KHALID <i>et al.</i> (2008)	GLCM / RNA	95,00% (Macro)

<sup>1</sup> Disponível em: <http://www.ux.uis.no/~tranden/brodatz.html>

PESQUISADOR(ES)	TÉCNICAS	RECONHECIMENTO
YUSOF <i>et al.</i> (2010)	Gabor e GLCM / RNA	90,33% (Macro)
NASIRZADEH <i>et al.</i> (2010)	LBP / RNA	96,60% (Macro)
DE PAULA <i>et al.</i> (2009)	Cor / RNA	65,20% (Macro)
DE PAULA <i>et al.</i> (2009)	Cor e GLCM	82,00% (Macro)
DE PAULA <i>et al.</i> (2010)	Cor e GLCM	80,80% (Macro)
YU <i>et al.</i> (2009)	GLCM	<i>Informação não disponível no artigo</i>
YU <i>et al.</i> (2009)	HSV, GLCM, LRE, Wavelets e Fractais	<i>Informação não disponível no artigo</i>

Fonte: Adaptado de Paula *et al.* (2014) e Martins *et al.* (2015).

### 2.3.2 Arquiteturas de CNN

Em 1989 um grupo de pesquisadores publicou uma pesquisa abordando o reconhecimento de dígitos manuscritos do código postal fornecidos pelo Serviço Postal dos EUA. Tal pesquisa resultou em uma rede de processamento de dados brutos (sem tratamentos prévios) com taxa de erro de 1%. Este projeto é considerado o pioneiro em CNN (LECUN *et al.*, 1989).

As CNN's foram popularizadas a partir do ano de 2010 por meio de uma competição científica denominada ILSVRC - ImageNet Large Scale Visual Recognition Challenge (Desafio de Reconhecimento Visual de Grande Escala da ImageNet) realizada anualmente pela ImageNet, uma base de imagens para pesquisadores e educadores.

O ILSVRC<sup>2</sup> é um desafio realizado entre equipes do mundo todo e consiste em solucionar problemas complexos de reconhecimento e classificação de imagens, englobando centenas de classes diversas e milhões de exemplos. O intuito do evento é contribuir com conteúdo científico por meio da criação ou melhoria de técnicas computacionais (RUSSAKOVSKY *et al.*, 2015). Algumas das arquiteturas mais populares originadas por este evento são descritas a seguir.

<sup>2</sup> Mais informações em: <http://www.image-net.org/challenges/LSVRC>

### 2.3.2.1 Arquitetura AlexNet

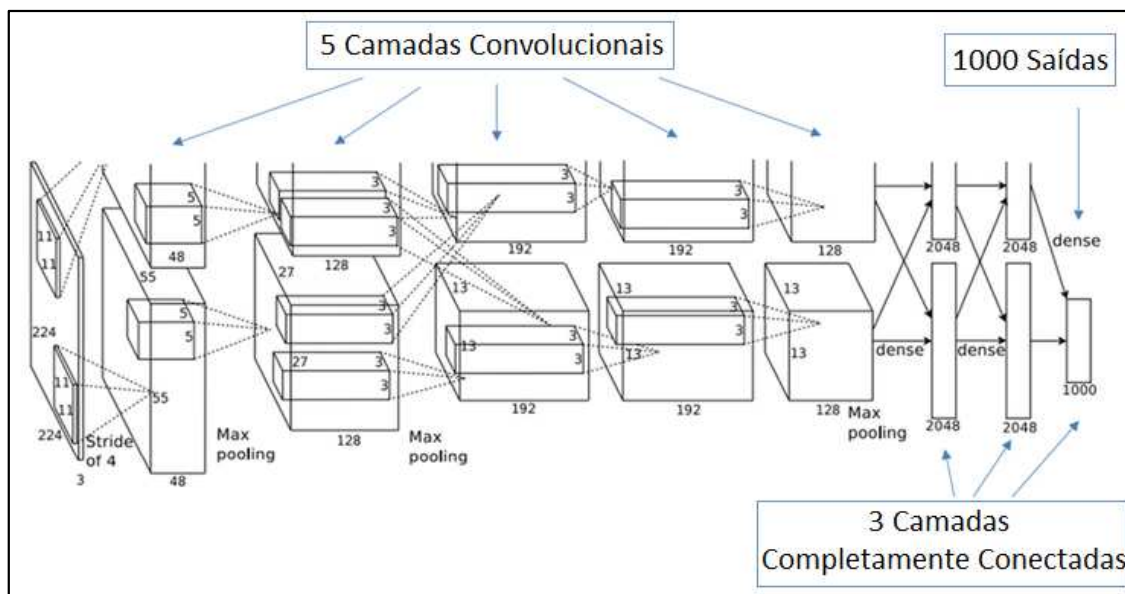
No ano de 2010, realizou-se o primeiro desafio ILSVRC, no qual o objetivo foi averiguar a melhor técnica para classificação de 1,2 milhões de imagens em 1000 classes diversas. Neste evento, uma equipe destacou-se ao desenvolver uma CNN bastante profunda e que recebeu o nome de um de seus integrantes: AlexNet. A rede mostrou-se muito eficiente, resultando em taxas de erro de 17%, sendo melhor que o estado da arte anterior. No desafio ILSVRC de 2012, a mesma equipe utilizou uma variação da rede original e venceu o campeonato, atingindo taxas de erro de 15,3%, enquanto o segundo lugar atingiu de 26,2% (KRIZHEVSKY *et al.*, 2012).

Ainda segundo Krizhevsky *et al.* (2012), a AlexNet é uma rede composta por 5 camadas convolucionais, sendo as camadas 1, 2 e 5 seguidas de *MaxPooling* (3 x 3), 3 camadas completamente conectadas, sendo a última com 1000 saídas. Utiliza-se a função ReLU para a ativação dos neurônios, os quais conectam-se apenas à camada anterior.

A primeira camada filtra a imagem de entrada com 96 *kernels* convolucionais de dimensões 11 x 11 x 3 e passo de 4 pixels. A estrutura original foi treinada e é otimizada para imagens de entrada com dimensões de 224px x 224px com 3 canais (*Red, Green e Blue* - RGB). Após resumir a resposta da primeira camada com *MaxPooling*, a segunda realiza a próxima convolução, utilizando 256 *kernels* de dimensões 5 x 5 x 48, seguida de *MaxPooling*. A terceira camada utiliza 384 *kernels* de dimensões 3 x 3 x 256, a quarta utiliza 384 *kernels* de dimensões 3 x 3 x 192 e a quinta utiliza 256 *kernels* de tamanho 3 x 3 x 192, seguida de *MaxPooling*. As duas primeiras camadas completamente conectadas são compostas por 4096 neurônios cada e a terceira é composta com 1000 neurônios de saída.

Em outras palavras, a convolução realizada com os *kernels* (filtros com vários neurônios clonados) gera mapas de características. Sendo assim, em uma camada, o número de *kernels* equivale a quantidade de neurônios.

De modo geral, a arquitetura engloba 60 milhões de parâmetros e 650.000 de neurônios e o processo de treinamento da rede original durou original durou 6 dias com melhor época = 90. Foram utilizadas duas GPU's de modelo NVIDIA GTX580 3GB. A arquitetura resumida da rede pode ser observada na Figura 18.



**Figura 18 – Arquitetura de CNN: AlexNet**  
 Fonte: Adaptado de Krizhevsky *et al.* (2012).

### 2.3.2.2 Arquitetura VGG16

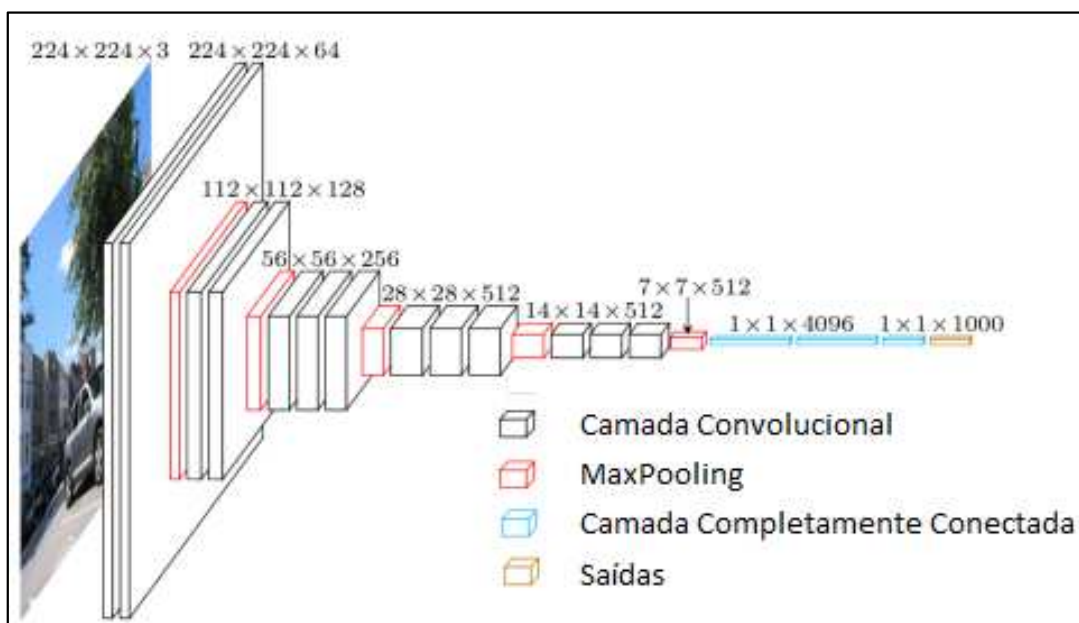
No desafio ILSVRC do ano de 2014, foi desenvolvida outra CNN utilizando uma abordagem um pouco diferente da AlexNet. O objetivo foi construir uma rede mais profunda, porém com filtros convolucionais menores. A arquitetura foi denominada VGGNet e possui 6 configurações, sendo a mais utilizada a VGG16, também conhecida como VGGNet-D. Tal arquitetura mostrou-se bastante robusta e atingiu taxa de erro de 7,3% no desafio de classificação de quase 1,5 milhões de imagens de 1000 classes de objetos diversos (SIMONYAN; ZISSERMAN, 2015).

De acordo com Simonyan e Zisserman (2015), a VGG16 é composta por 13 camadas convolucionais, sendo as camadas 2, 4, 7, 10 e 13 seguidas de *MaxPooling* (2 x 2), 3 camadas completamente conectadas e 1000 saídas. Para ambas, utiliza-se a função ReLU para a ativação dos neurônios, os quais conectam-se somente à camada anterior.

Os modelos originais foram treinados com imagens de entrada de dimensões 224px x 224px com 3 canais (RGB), as quais passam por um processo de pré-processamento, resultando em imagens bidimensionais (1 único canal) com os valores médios de RGB para cada pixel.

Todas as camadas convolucionais filtram a imagem de entrada com *kernels* bidimensionais de  $3 \times 3$  e passo de 1 pixel. A quantidade de *kernels* por camada convolucional é de: Camadas 1 e 2 – 64; Camadas 3 e 4 – 128; Camadas 5, 6 e 7 – 256; Camadas 8, 9, 10, 11, 12 e 13 – 512. Além disso, o modelo original possui as duas primeiras camadas completamente conectadas compostas por 4096 neurônios cada e a terceira composta por 1000 neurônios de saída.

No desafio ILSVRC foram utilizadas várias configurações e o treinamento foi realizado em quatro GPU's de modelo NVIDIA Titan Black. O melhor modelo foi obtido entre 2 e 3 semanas de treinamento, o qual foi realizado com *batches* de tamanho 256, 370.000 iterações e melhor época = 74. A arquitetura resumida da VGG16 pode ser observada na Figura 19.



**Figura 19 – Arquitetura de CNN: VGG16**  
 Fonte: Adaptado de Simonyan e Zisserman (2015).

### 2.3.2.3 Arquitetura ResNet50

No ILSVRC de 2015, pesquisadores da *Microsoft* desenvolveram uma arquitetura de CNN ainda mais profunda, porém com filtros convolucionais não tão complexos. A arquitetura ResNet possui 5 configurações de arquitetura: ResNet18, ResNet34, Resnet50, ResNet101 e ResNet152, nas quais o número no final da



nomenclatura representa a quantidade de camadas. A ResNet152 atingiu taxa de erro de 3,57% no desafio de classificação de quase 1,5 milhões de imagens, sendo 1,28 milhões de imagens para treinamento e 150 mil imagens para validação e testes, entre 1000 possíveis classes de objetos diversos (HE *et al.*, 2015).

De acordo com He *et al.* (2015), uma das suas configurações mais otimizadas, a ResNet50, é composta por 49 camadas convolucionais (com ativação ReLU), sendo que a camada 1 é seguida de *MaxPooling* (3x3 e passo 2) e a camada 49 é seguida de *AveragePooling*, e uma camada extra completamente conectada. Os modelos originais foram treinados com imagens de entrada de dimensões 224px × 224px com 3 canais (RGB) e são otimizadas para imagens de até 300px × 300px.

No desafio ILSVRC foram utilizadas várias configurações e o treinamento foi realizado em duas GPU's. A arquitetura resumida da ResNet50 pode ser observada na Figura 20.

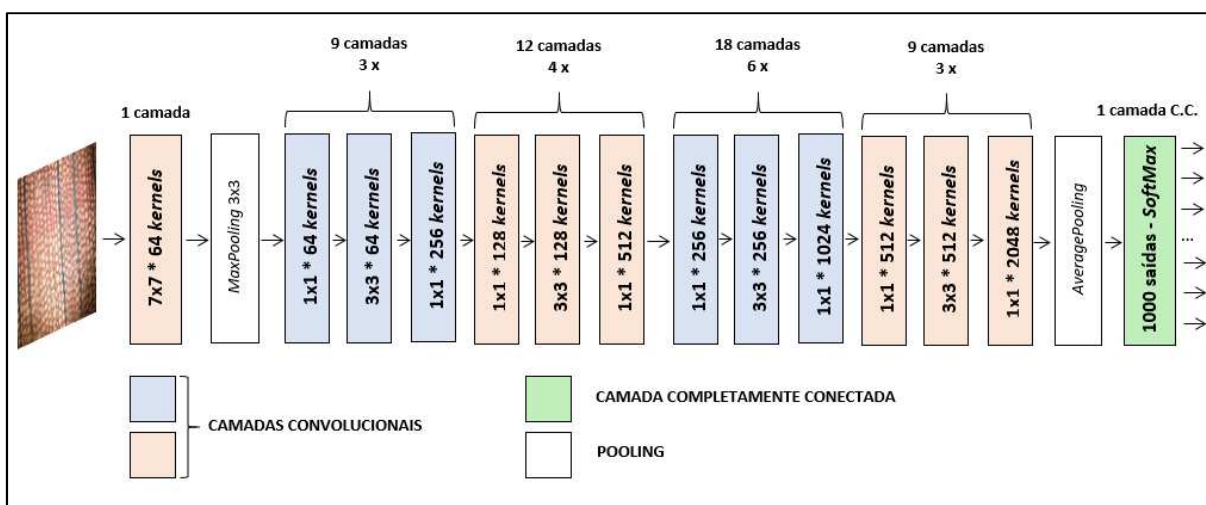


Figura 20 – Arquitetura de CNN: ResNet50

Fonte: Adaptado de He *et al.* (2016).

### 3 METODOLOGIA

De acordo com as definições de Gil (2008), Marconi e Lakatos (2003) e Gerhardt e Silveira (2009), a presente pesquisa pode ser caracterizada como quantitativa com método hipotético-dedutivo, na qual há proposta de experimentos e posterior avaliação de resultados a partir da hipótese: **é possível classificar espécies florestais a partir de imagens utilizando técnicas de Deep Learning**. Em relação à natureza, objetivos e procedimentos técnicos, também pode ser classificada como pesquisa aplicada, descritiva, experimental e operacional.

As seções a seguir apresentam as técnicas, métodos, materiais e planejamentos utilizados para a realização dos experimentos.

#### 3.1 AMBIENTE DE EXPERIMENTOS E DESENVOLVIMENTO

Para iniciar a fase experimental, tornou-se necessário definir algumas ferramentas e técnicas de auxílio para a realização de tais experimentos e construção do produto final:

Experimentos com abordagem tradicional:

- a) C++: linguagem de programação e plataforma CodeBlocks 16;
- b) OpenCV<sup>3</sup>: biblioteca de VC e PDI (versão 3.1);
- c) Weka: plataforma para experimentos e classificações utilizando RNA's tradicionais, entre outros (versão 3.8.1).

Experimentos com abordagem em DL:

- a) Java: linguagem de programação e IDE (Integrated Development Environment) Eclipse Oxygen;
- b) DeepLearning4j (DL4J)<sup>4</sup>: *framework* de código aberto para a linguagem Java, específica para a construção de modelos artificiais com abordagens em *Deep Learning* (versão 0.9.1). A licença de uso desta plataforma permite registros de patente;

---

<sup>3</sup> Disponível em: <https://opencv.org/>

<sup>4</sup> Disponível em: <https://deeplearning4j.org/>

c) Treinamento: CPU.

Desenvolvimento do produto:

- a) Java: Linguagem de programação e IDE (Integrated Development Environment) Eclipse Oxygen;
- b) TomCat: Servidor de aplicações *web* para Java (Versão 8.5);
- c) AngularJS e Bootstrap: *Frameworks* de desenvolvimento *front-end* para aplicações *web*;
- d) Design Responsivo: *Design* que adapta o *layout* de acordo com o tipo de dispositivo e sua resolução.

Foi utilizada a infraestrutura de TI (Tecnologia da Informação) da UTFPR (Universidade Tecnológica Federal do Paraná) – Câmpus Medianeira e a infraestrutura do Google Cloud<sup>5</sup> para a realização dos experimentos e armazenamento do produto final.

### 3.1.1 Bases de Imagens

Para os experimentos com imagens macroscópicas, foi utilizada a base elaborada e disponibilizada por Paula *et al.* (2012)<sup>6</sup>, composta por 2942 imagens macroscópicas de madeiras de 41 espécies com resoluções de 3264x2448 pixels.

Para a construção desta base, Paula *et al.* (2014) capturaram as imagens no plano transversal a partir de blocos de madeira de 2cm<sup>3</sup> previamente lixados para a uniformização das amostras. Foi construída uma estrutura com duas lâmpadas halógenas com filtros de papel sulfite para a padronização da iluminação e uma câmera fotográfica SONY modelo DSC T20. Para a aquisição das imagens, as amostras foram posicionadas à aproximadamente 1cm da lente.

A Tabela 2 apresenta uma parte da relação de espécies macroscópicas presentes na base e o número de exemplares por espécie. A tabela com a relação completa pode ser consultada no Anexo A desta pesquisa.

---

<sup>5</sup> Disponível em: <https://cloud.google.com>

<sup>6</sup> Disponível em: <http://web.inf.ufpr.br/vri/databases/forest-species-database-macroscopic/>

**Tabela 2 – Demonstração da Composição da Base Macro**

ID	Nome Popular	Nome Científico	Quantidade
01	Acrocarpus	<i>Acrocarpus fraxinifolius</i>	53
02	Amapá	<i>Brosimum paraensis</i>	63
03	Andiroba	<i>Carapa guianensis</i>	43
04	Angelim Pedra	<i>Hymenolobium petraeum</i>	99
05	Araucaria	<i>Araucaria angustifolia</i>	63
06	Assacu	<i>Hura crepitans</i>	41
07	Bracatinga	<i>Mimosa scabrella</i>	48
08	Cabriuva Vermelha	<i>Myroxylon balsamum</i>	75
09	Castanheira	<i>Bertolothia excelsa</i>	72
10	Cedrinho	<i>Erismia uncinatum</i>	98
...	...	...	...

Fonte: Adaptado de Paula *et al.* (2014).

Para os experimentos com imagens microscópicas, foi utilizada a base elaborada e disponibilizada por Martins *et al.* (2012)<sup>7</sup>. Esta base é composta por 2240 imagens de madeiras de 112 espécies florestais diferentes (20 amostras de cada espécie) com resoluções de 1024x768 pixels.

Para a construção desta base, Martins *et al.* (2015) confeccionaram blocos de madeira, inicialmente de 2cm<sup>3</sup>, os quais passaram por um processo de cozimento e cortes histológicos no plano transversal utilizando um equipamento chamado micrótomo de deslizamento para melhorar os aspectos anatômicos. Foi realizada a tripla coloração dos cortes e a desidratação em série alcoólica ascendente. Em seguida, foram coletadas as imagens por meio de um microscópio Olympus modelo CX40 utilizando aproximação ótica de 100x.

A Tabela 3 apresenta parte da relação de espécies microscópicas da base. A tabela com a relação completa pode ser consultada no Anexo B desta pesquisa.

**Tabela 3 – Demonstração da Composição da Base Micro**

ID	Ordem	Família	Gênero	Espécie	Quantidade
001	Ginkgoales	Ginkgoaceae	<i>Ginkgo</i>	<i>biloba</i>	20
002	Pinales	Araucariaceae	<i>Agathis</i>	<i>becarii</i>	20
003	Pinales	Araucariaceae	<i>Araucaria</i>	<i>angustifolia</i>	20
004	Pinales	Cephalotaxaceae	<i>Cephalotaxus</i>	<i>drupacea</i>	20
005	Pinales	Cephalotaxaceae	<i>Cephalotaxus</i>	<i>harringtonia</i>	20
006	Pinales	Cephalotaxaceae	<i>Torreya</i>	<i>nucifera</i>	20
007	Pinales	Cupressaceae	<i>Calocedrus</i>	<i>decurrens</i>	20
008	Pinales	Cupressaceae	<i>Chamaecyparis</i>	<i>formosensis</i>	20
009	Pinales	Cupressaceae	<i>Chamaecyparis</i>	<i>pisifera</i>	20
010	Pinales	Cupressaceae	<i>Cupressus</i>	<i>arizonica</i>	20
...	...	...	...	...	...

Fonte: Adaptado de Martins *et al.* (2015).

<sup>7</sup> Disponível em: <http://web.inf.ufpr.br/vri/databases/forest-species-database-microscopic/>

### 3.2 MÉTODO EXPERIMENTAL PROPOSTO

A fim de organizar o trabalho, foram estabelecidas algumas etapas e atividades experimentais para possibilitar o cumprimento dos objetivos específicos (vide fluxograma na Figura 21). O planejamento para tais etapas e atividades é descrito nas subseções a seguir e os respectivos resultados são descritos separadamente na Seção 4.

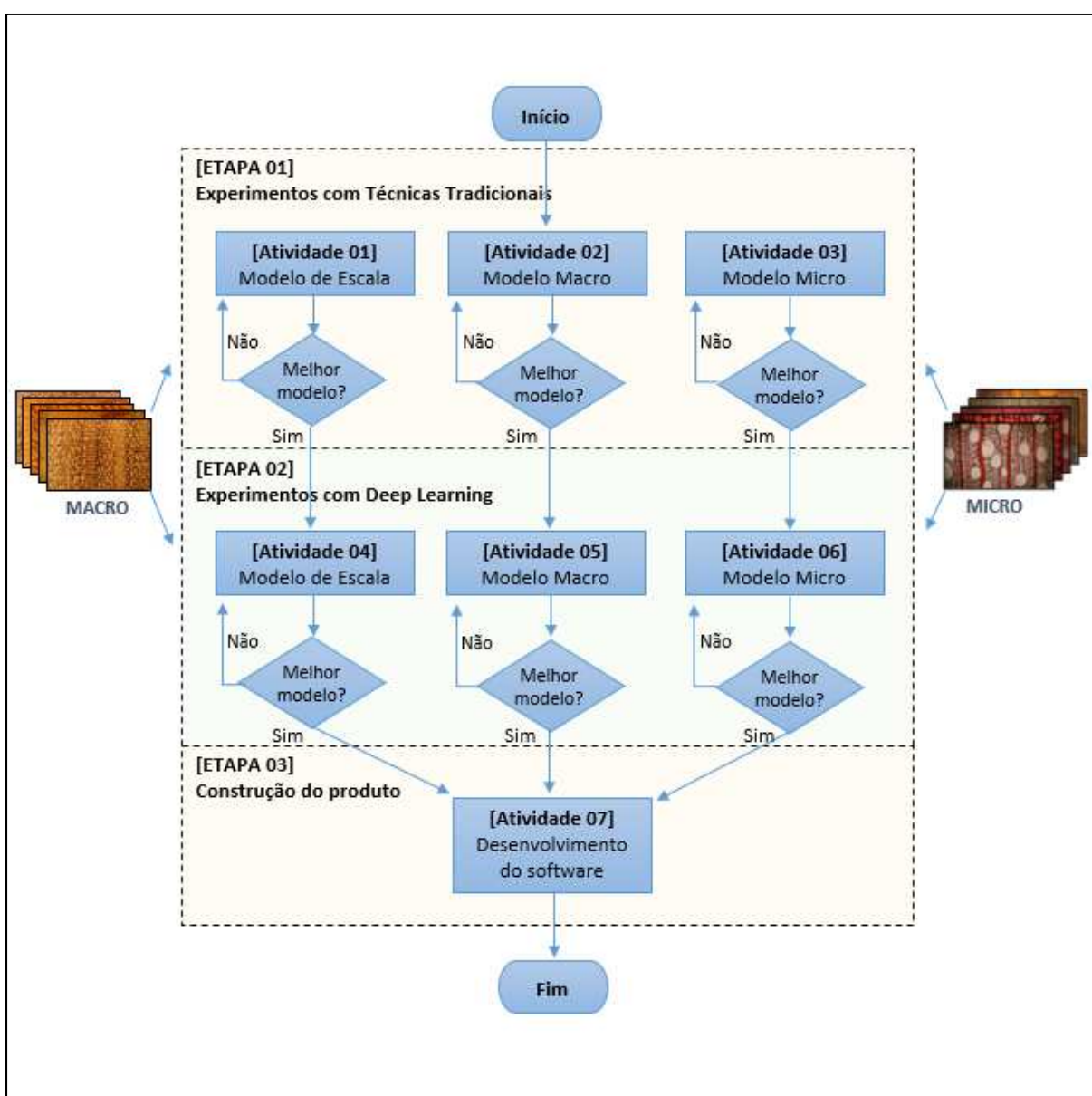


Figura 21 – Fluxograma de Atividades  
Fonte: Autoria Própria.

### 3.2.1 Abordagem Tradicional: GLCM e RNA MLP

Para esta etapa, realizou-se a divisão das imagens da seguinte maneira: 10% para testes finais e, dos 90% restantes, 70% para treinamento e 30% para validação do processo de treinamento. Esta divisão foi aplicada a cada base de imagens, resultando em:

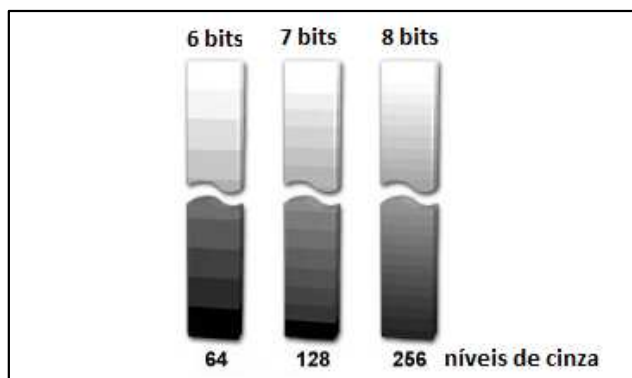
- a) Base macroscópica (41 classes): 1854 imagens para treinamento; 795 imagens para validação e 293 imagens para testes;
- b) Base microscópica (112 classes): 1411 imagens para treinamento; 605 imagens para validação e 224 imagens para testes;
- c) Base de escala (2 classes): 3265 imagens para treinamento; 1400 imagens para validação e 517 imagens para testes.

A base de escala trata-se de uma nova base criada a partir da junção das imagens das bases macro e micro. Lidar com todas as 153 classes das duas bases originais de uma só vez poderia aumentar muito a complexidade do modelo final, bem como o custo computacional. Sendo assim, o objetivo desta base é classificar inicialmente se uma imagem é macroscópica ou microscópica e, posteriormente, repassar a imagem para um segundo modelo (modelo de macro ou modelo micro) para que finalmente aconteça a classificação da espécie.

Para esta etapa, foram utilizados os seguintes parâmetros e técnicas:

- a) Extração dos 14 descritores de GLCM e análise dos 6 descritores teoricamente mais relevantes (vide Seção 2.2.1.2);
- b) Graus:  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$  e  $135^\circ$  e Distâncias: 1 e 5;
- c) Número de bits por pixel: 8 bits ( $2^8 = 256$  níveis de cinza), 7 bits ( $2^7 = 128$  níveis de cinza) e 6 bits ( $2^6 = 64$  níveis de cinza);
- d) Classificador: RNA.

Optou-se pela variação dos números de bits por pixel, já que a redução do número de bits implica na perda de detalhes e diminui as variações níveis de cinza. Desta forma, quanto menor a quantidade de níveis de cinza, mais genéricas ficam as coocorrências, tornando a análise menos sensível a pequenas variações das características da madeira (Vide Figura 22).



**Figura 22 – Número de Bits x Níveis de Cinza**  
**Fonte: Aatoria Própria.**

Para cada imagem, foram extraídas 336 características, considerando 14 descritores \* 2 distâncias \* 4 ângulos \* 3 variações de níveis de cinza. Para cada GLCM calculada, foi realizada a normalização para a obtenção da matriz de probabilidades (GLCM normalizada), possibilitando a extração dos 14 descritores. A partir de tais características, foi possível realizar a seleção e combinação das mesmas para gerar os resultados apresentados na Seção 4.1.

### 3.2.2 Abordagem com Deep Learning: CNN

Para esta etapa, realizou-se a divisão das imagens das três bases em 70% para treinamento e 30% para testes. Tal divisão resultou nas seguintes quantidades:

- a) Base de escala (2 classes): 3641 imagens para o conjunto de treinamento e 1541 para testes;
- b) Base macroscópica (41 classes): 2073 imagens para o conjunto de treinamento e 869 para testes;
- c) Base microscópica (112 classes): 1568 imagens para o conjunto de treinamento e 672 para testes.

A fim de estabelecer a arquitetura ideal para cada um dos três modelos e, posteriormente, treiná-los, foi realizada uma época de treinamento para a base macroscópica e uma para a base microscópica, utilizando cada uma das seguintes arquiteturas: AlexNet, VGG16 e ResNet50 (vide Seção 2.3.2).

Com base no conceito de *Transfer Learning*, as três arquiteturas foram obtidas por meio do *framework* DL4J com os pesos das camadas convolucionais

pré-treinados (vide Seção 2.3.2). Sendo assim, tais camadas já conhecem muito sobre o reconhecimento de padrões em imagens (arestas, curvas, bordas, etc.). Desta forma, foi possível desacoplar apenas a camada completamente conectada de saídas e acoplar uma nova camada com a quantidade de neurônios de saída correspondente ao número de possíveis classes para cada um dos modelos.

De forma a otimizar o processamento e aproveitar o conhecimento já existente nas arquiteturas pré-treinadas, optou-se por manter todas as camadas convolucionais congeladas, deixando apenas as camadas de classificação final com os pesos treináveis. Desta forma, o Back-Propagation ocorreu apenas nas camadas completamente conectadas. A utilização de modelos pré-treinados e o congelamento de camadas é uma prática de Transfer Learning que possibilita que novos modelos convirjam mais rápido, com menor poder computacional em menos tempo.

Como as arquiteturas são otimizadas para trabalhar com imagens de 3 canais com dimensões de até 300px x 300px, a primeira tentativa foi avaliar os resultados a partir do redimensionamento de todas as imagens das três bases para dimensões proporcionais a estas, ou seja, com largura ou altura máxima de 300px. Os resultados obtidos a partir deste experimento são apresentados na Seção 4.2.

### 3.2.2.1 Data augmentation e descrição das novas bases

Utilizando redimensionamento das imagens, todos os modelos apresentaram taxas de reconhecimento iniciais relativamente baixas (vide Seção 4.2) em se tratando de uma arquitetura pré-inicializada. Com isso, observou-se que o número de exemplos por classe e a perda de detalhes com o redimensionamento foram fatores que interferiram nos resultados.

Devido ao prazo de conclusão da pesquisa, realizar a aquisição de novas imagens de cada classe para o aumento das bases seria uma tarefa inviável. Então, foi necessário estabelecer uma estratégia para manter o máximo de características das imagens originais e aumentar a quantidade de exemplos por classe.

O conceito de *Data Augmentation* foi a estratégia mais adequada nesta fase. Todas as imagens da base macroscópica possuem dimensões de 3264x2448 pixels e as imagens da base microscópica 1024x768 pixels. Com isso, foi possível dividí-



las em subimagens de dimensões próximas ao limite recomendado pelas arquiteturas (*crop* de imagens). A técnica de divisão das imagens também foi abordada na pesquisa de Paula *et al.* (2014), a qual foi denominada “Dividir para Conquistar” e foi motivação para esta etapa da pesquisa.

A fim de padronizar a quantidade de fragmentos, optou-se por redimensionar as imagens originais macroscópicas para as mesmas dimensões das imagens originais microscópicas, isto é, 1024x768 pixels. Em seguida, cada imagem original foi dividida em partes iguais de 256x256 pixels, resultando em 12 subimagens (4 na horizontal e 3 na vertical). Além da divisão, para cada fragmento realizou-se 5 transformações: espelhamento, rotação 90°, rotação 180°, rotação 270° e inversão de cores (imagem negativa). Logo, cada imagem original gerou 72 subimagens, ou seja, 12 fragmentos \* (5 transformações + fragmento original). A representação destas operações pode ser observada na Figura 23.

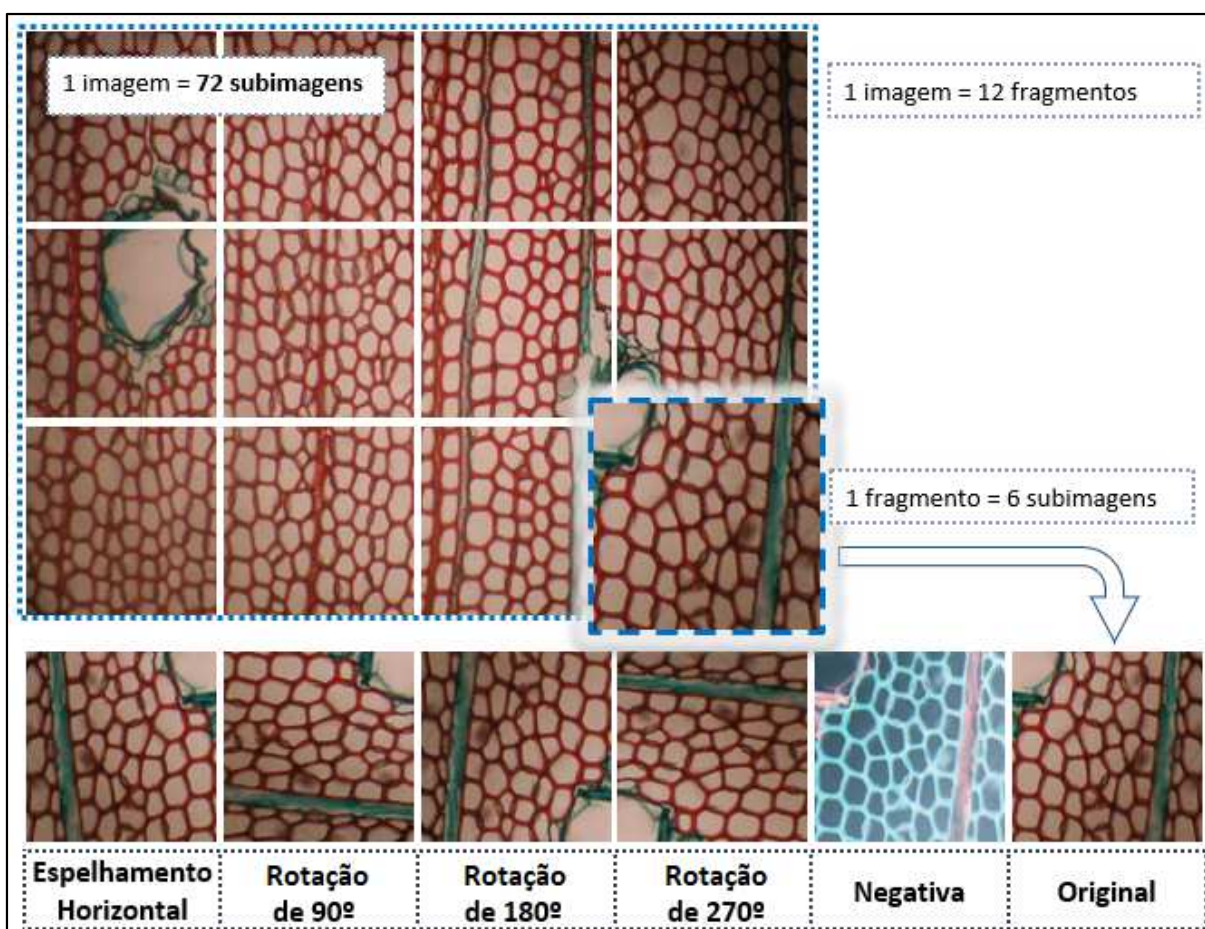


Figura 23 – Transformações para Data Augmentation  
Fonte: Autoria Própria.

Após a realização do *Data Augmentation*, as novas bases de imagens ficaram bem maiores. Na base macroscópica, a classe que possui menor número de exemplos originais é a 11 (Espécie Cedro), com 37 imagens (26 para treinamento e 11 para testes) e, após as transformações, esta mesma classe passou a possuir 2.664 imagens (1.872 para treinamento e 792 para testes). Na base microscópica, todas as classes possuem 20 exemplos originais e, após as transformações, passaram a possuir 1.440 imagens (1.008 para treinamento e 432 para testes).

A nova base de escala ficou com o total de 373.104 imagens (262.152 para treinamento e 110.952 para testes), a base macroscópica com o total de 211.824 imagens (149.256 para treinamento e 62.568 para testes) e a base microscópica com o total de 161.280 imagens (112.896 para treinamento e 48.384 para testes). A correlação entre imagens originais e novas pode ser visualizada no Quadro 3.

**Quadro 3 – Quantidade de Imagens nas Bases Após Data Augmentation**

BASE DE IMAGENS	TREINAMENTO (70%)		TESTES (30%)		TOTAL GERAL	
	IMAGENS ORIGINAIS	SUBIMAGENS	IMAGENS ORIGINAIS	SUBIMAGENS	IMAGENS ORIGINAIS	SUBIMAGENS
Escala	3.641	262.152	1.541	110.952	5.182	373.104
Macro	2.073	149.256	869	62.568	2.942	211.824
Micro	1.568	112.896	672	48.384	2.240	161.280






Fonte: Autoria Própria.

### 3.2.2.2 Técnicas de fusão das subimagens

Com o incremento das bases, foi possível treinar e testar novamente os modelos, porém com quantidade muito maior de exemplos. Entretanto, as acurácias analisadas não correspondiam a imagens inteiras, mas sim aos fragmentos das imagens originais. Logo, tornou-se necessário definir uma técnica de fusão das classificações dos fragmentos para obter a classificação final. Foram testadas as seguintes técnicas: Soma e Votação Majoritária (KITTER *et al.*, 1998 e JAIN, DUIN, MAO, 2000).

Na votação, para cada fragmento, analisa-se a classe prevista e, posteriormente, realiza-se a contagem de votos para cada classe. A classe vencedora é a que possui maior número de votos (vide o exemplo do Quadro 4).

**Quadro 4 – Exemplo de Fusão de Classificações por Votação**






	MODELO ESCALA		MODELO MACRO				
	LABEL: MACRO	LABEL: MICRO	LABEL: 01	LABEL: 02	LABEL: 04	...	LABEL: 41
	SIM	NÃO	NÃO	NÃO	SIM	...	NÃO
	SIM	NÃO	NÃO	NÃO	SIM	...	NÃO
...	Fragmentos da Imagem Original		...	...	...	...	...
	NÃO	SIM	SIM	NÃO	NÃO	...	NÃO
	SIM	NÃO	NÃO	NÃO	SIM	...	NÃO
	SIM	NÃO	NÃO	NÃO	SIM	...	NÃO
<b>VOTOS</b>	<b>04</b>	<b>01</b>	<b>01</b>	<b>00</b>	<b>04</b>	...	--

Annotations: Red dashed boxes highlight the 'SIM' and 'NÃO' counts for Label 04 in the Macro Model, and the '04' and '01' vote counts in the 'VOTOS' row. Red arrows point from the 'SIM' counts to the '04' vote count and from the 'NÃO' counts to the '01' vote count.

Fonte: Autoria Própria.

Para a fusão por média das probabilidades, considera-se as probabilidades de cada classe para cada fragmento. Em seguida, realiza-se a média destas probabilidades para cada classe, considerando todos os fragmentos. Logo, a classe que apresentar maior probabilidade média representa a classificação final (vide o exemplo do Quadro 5).

**Quadro 5 – Exemplo de Fusão de Classificações por Média das Probabilidades**

	MODELO ESCALA		MODELO MACRO				
	LABEL: MACRO	LABEL: MICRO	LABEL: 01	LABEL: 02	LABEL: 04	...	LABEL: 41
	0,999999999	0,000000001	0,000000000	0,000000000	0,999994040	...	0,000000000
	1,000000000	0,000000000	0,000000000	0,000000000	0,999999881	...	0,000000000
...	Fragmentos da Imagem Original		...	...	...	...	...
	0,999977589	0,000000003	0,000000000	0,000000004	0,999921083	...	0,000000000
	0,999999965	0,000000035	0,000000000	0,000000358	0,999847651	...	0,000000000
	1,000000000	0,000000000	0,000000000	0,000000000	0,999999999	...	0,000000000
<b>Média</b>	<b>0,999998132</b>	<b>0,000001871</b>	<b>0,000000000</b>	<b>0,000000030</b>	<b>0,999972950</b>	...	<b>0,000000000</b>

Annotations: Red dashed boxes highlight the 'Média' row for Label 04 and Label MICRO. Red arrows point from the 'Média' values to the corresponding 'Média' row.

Fonte: Autoria Própria.

Após a definição das técnicas de fusão a serem testadas, iniciaram-se novamente os treinamentos com as três arquiteturas e as novas bases de imagens. Foram observadas as acurácias em relação aos fragmentos, imagens inteiras com fusão por votação e imagens inteiras com fusão por média das probabilidades. Desta forma, tornou-se possível identificar a técnica de fusão adequada para a continuação dos experimentos (vide resultados experimentais na Seção 4.2.1).

### 3.2.2.3 Construção dos modelos de CNN e produto final

O primeiro modelo criado foi o Modelo de Escala, responsável por identificar se a imagem de entrada está em escala macro ou microscópica (classes “Macro” e “Micro”). Em seguida, foi construído o Modelo Macro, responsável pela classificação da espécie a partir de uma imagem macroscópica, de acordo com as 41 possíveis espécies que foram apresentadas na Tabela 2. Por fim, foi construído o Modelo Micro, responsável pela classificação da espécie a partir de uma imagem microscópica, de acordo com as 112 espécies que foram apresentadas na Tabela 3.

A fim de cumprir o último objetivo específico desta pesquisa, realizou-se o desenvolvimento do software web como produto final, utilizando os três modelos de CNN construídos. Foi utilizada a linguagem Java no *back-end*, os *frameworks* AngularJS e *Bootstrap* no *front-end* e o *framework* DL4J para carregar e executar os classificadores de CNN. A interface gráfica foi construída com o cuidado de ser a mais simples possível para o usuário final, além de possuir *design* responsivo (*layout* que se adapta ao dispositivo e sua resolução).

Para disponibilizar o acesso ao software, foi criado um subdomínio do site da UTFPR associado à um servidor no qual o software encontra-se alocado. Os detalhes e algumas imagens deste produto final encontram-se na Seção 5.3.

## 4 RESULTADOS EXPERIMENTAIS

Esta seção apresenta os experimentos realizados e os resultados parciais obtidos a partir do planejamento descrito na Seção 3.2. A análise final e discussão dos resultados são descritas separadamente na Seção 5.

### 4.1 ABORDAGEM TRADICIONAL: GLCM E RNA MLP

De acordo com o planejamento apresentado na Seção 3.2.1, foram treinados modelos de MLP no software Weka utilizando uma série de combinações experimentais a fim de obter as melhores taxas de reconhecimento em relação aos conjuntos de validação. O ponto de parada para todos os experimentos foi a época 2000, após este ponto, identificou-se início de *overfitting*. Em todos os experimentos foi considerada a concatenação dos descritores obtidos para cada um dos quatro ângulos, conforme descrito previamente.

Para os dois melhores modelos (macro e microscópico) obtidos a partir das diferentes combinações, foram analisadas também as taxas de reconhecimento com os conjuntos de testes. No Quadro 6, considera-se (V) como a taxa de reconhecimento para os conjuntos de validação e (T) a taxa de reconhecimento para os conjuntos de testes.

**Quadro 6 – Descritores de GLCM e Taxas de Reconhecimento com RNA**

COMBINAÇÃO	BASE DE ESCALA		BASE MACROSCÓPICA		BASE MICROSCÓPICA	
	Distância 1	Distância 5	Distância 1	Distância 5	Distância 1	Distância 5
<b>Níveis de Cinza:</b> 256 <b>Descritores (14):</b> SMA; CON; COR; VAR; MDI; MDS; VSM; ESM; ENT; VDF; EDF; MC1; MC2; MCC.	100% (V)	100% (V)	55,85% (V)	59,60% (V)	57,85% (V)	63,80% (V)
<b>Níveis de Cinza:</b> 128 <b>Descritores (14):</b> SMA; CON; COR; VAR; MDI; MDS; VSM; ESM; ENT; VDF; EDF; MC1; MC2; MCC.	100% (V)	100% (V)	55,22% (V)	60,75% (V)	57,68% (V)	61,00% (V)
<b>Níveis de Cinza:</b> 64 <b>Descritores (14):</b> SMA; CON; COR; VAR; MDI; MDS; VSM; ESM; ENT; VDF; EDF; MC1; MC2; MCC.	100% (V)	100% (V) 100% (T)	54,09% (V)	62,65% (V) 66,90% (T)	53,88% (V)	63,15% (V) 66,52% (T)

COMBINAÇÃO	BASE DE ESCALA		BASE MACROSCÓPICA		BASE MICROSCÓPICA	
	Distância 1	Distância 5	Distância 1	Distância 5	Distância 1	Distância 5
Níveis de Cinza: 256 Descritores (6): SMA; CON; VAR; COR; ENT; MDI.	100% (V)	100% (V)	51,24% (V)	55,72% (V)	51,02% (V)	55,04% (V)
Níveis de Cinza: 128 Descritores (6): SMA; CON; VAR; COR; ENT; MDI.	100% (V)	100% (V)	50,01% (V)	54,46% (V)	55,36% (V)	59,50% (V)
Níveis de Cinza: 64 Descritores (6): SMA; CON; VAR; COR; ENT; MDI.	100% (V)	100% (V)	53,81% (V)	57,87% (V)	55,99% (V)	60,00% (V)

Fonte: Autoria Própria.

Observa-se que ao se utilizar os 14 descritores de Haralick *et al.* (1973), as taxas de reconhecimento apresentaram um ganho médio de 4 pontos percentuais em relação aos resultados obtidos ao se utilizar apenas os 6 descritores mais relevantes, segundo Baraldi e Parmiggiani (1995). Ao se realizar uma análise individual das taxas de reconhecimento, observa-se ganhos de até 8,76 pontos percentuais, como por exemplo, as taxas de 63,80% (com 14 descritores) e 55,04% (com 6 descritores), taxas obtidas para a base microscópica, utilizando 256 níveis de cinza e distância 5.

Em análise ao custo computacional, o tempo médio de treinamento, utilizando as imagens macroscópicas e os 14 descritores, foi de 400 segundos (aproximadamente 7 minutos). Para as mesmas imagens, porém utilizando apenas 6 descritores, o tempo médio de treinamento foi de 250 segundos (aproximadamente 4 minutos).

Para as imagens microscópicas, o tempo médio de treinamento, utilizando os 14 descritores, foi de 1500 segundos (aproximadamente 25 minutos). Para as mesmas imagens, porém utilizando apenas 6 descritores, o tempo médio de treinamento foi de 1900 segundos (aproximadamente 32 minutos).

Para este problema de pesquisa, a teoria dos 6 descritores mais relevantes é questionável, uma vez que os demais descritores se complementam e melhoraram os índices, com custo computacional não muito impactante. Sendo assim, para outros problemas de pesquisa, a decisão do uso de 14 ou 6 descritores dependerá do nível de importância do ganho em relação ao custo computacional.

As melhores taxas de reconhecimento foram 66,90% para a base macroscópica e 66,52% para a base microscópica em relação aos conjuntos de testes. Já com relação aos conjuntos de validação, obteve-se 62,65% para a base

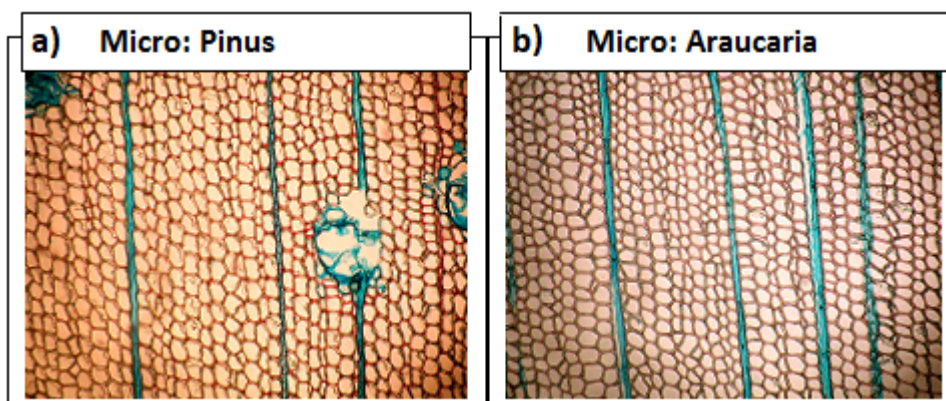


macroscópica e 63,80% para a base microscópica em relação aos conjuntos de validação. Observa-se que a utilização de menos níveis de cinza apresentou melhores resultados, além de ser uma técnica de menor custo computacional, uma vez que se trabalha com representações mais compactas das imagens (menor número de bits).

Neste contexto, observou-se também que a utilização de distância 5 tornou o processo menos sensível a pequenas coocorrências não relevantes para a classificação, bem como a utilização de 6 bits para cada pixel da imagem implicou em resultados bastante significativos, além de diminuir o custo computacional.

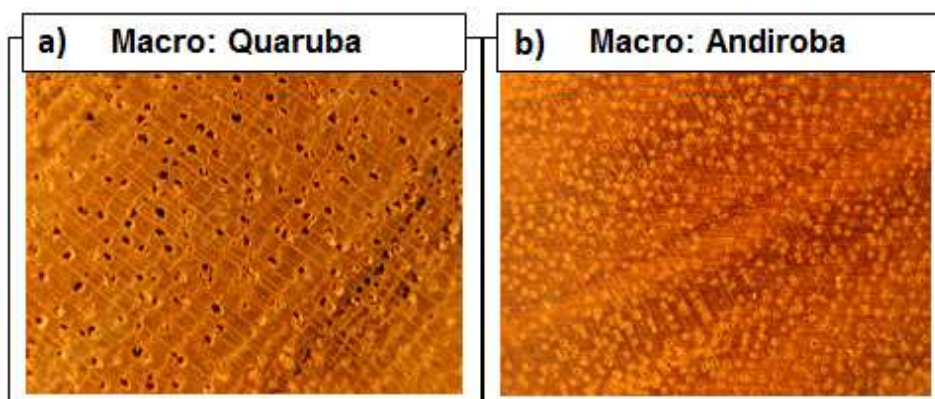
#### 4.1.1 Erros de Classificação

Ao analisar as matrizes de confusão, observou-se que alguns dos erros cometidos pelos classificadores ocorreram em amostras de diferentes classes com características texturais muito semelhantes. A Figura 24 apresenta uma das confusões entre as espécies microscópicas “*Pinus Elliottii*” (Figura 24 - a) e “*Araucaria Angustifolia*” (Figura 24 - b).



**Figura 24 – GLCM: Confusão de Classificação na Base Microscópica**  
**Fonte: Autoria Própria.**

A Figura 25 apresenta uma das confusões entre as espécies macroscópicas “*Quaruba Vochysia sp*” (Figura 25 - a) e “*Andiroba Carapa guianensis*” (Figura 25 - b), também muito semelhantes.



**Figura 25 – GLCM: Confusão de Classificação na Base Macroscópica**  
**Fonte: Autoria Própria.**

De modo geral, a GLCM mostrou-se uma técnica que pode trazer bons resultados com a classificação de imagens por análise de textura, principalmente quando explorada com diferentes combinações de parâmetros. Entretanto, exige forte intervenção do pesquisador para a extração dos descritores.

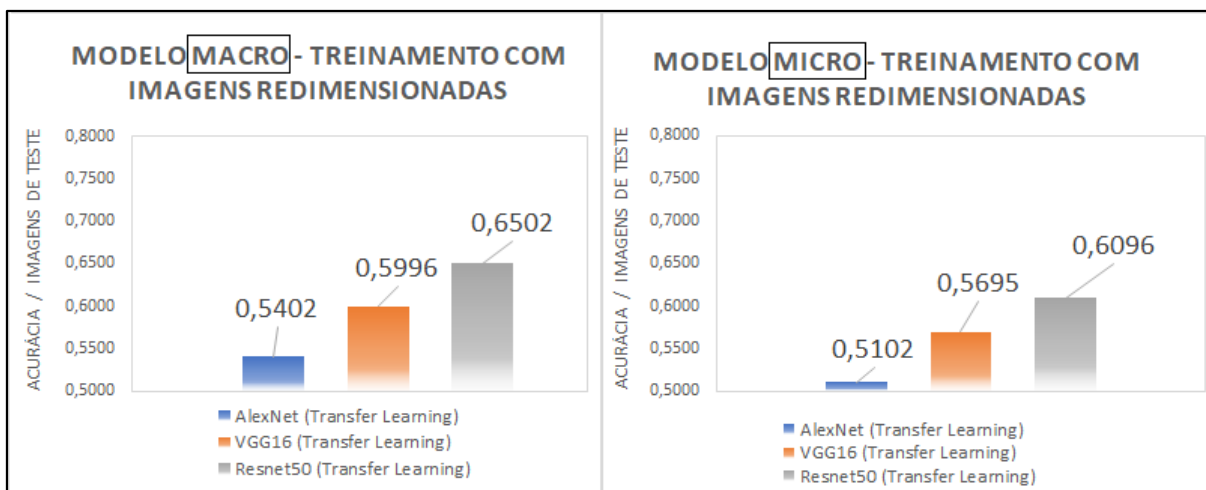
Em se tratando da técnica descritores de GLCM, estes resultados já superam as taxas de reconhecimento obtidas com GLCM pelos autores das bases. Para Paula *et al.* (2014), o melhor resultado com as imagens macroscópicas foi de 55,97% e para Martins *et al.* (2015), o melhor resultado com as imagens microscópicas foi de 50%.

#### 4.2 ABORDAGEM COM DEEP LEARNING: CNN

De acordo com o planejamento apresentado na Seção 3.2.2, o primeiro passo foi analisar as taxas de reconhecimento a partir do redimensionamento das imagens das bases. Com as imagens redimensionadas, camadas convolucionais congeladas e uma nova camada de neurônios de saída da arquitetura acoplada, foram realizados os primeiros treinamentos a fim de identificar a arquitetura que produzisse o modelo com maior acurácia em relação ao conjunto de testes.

Inicialmente os modelos com maior acurácia foram os treinados com a arquitetura ResNet50. Observa-se na Figura 26 que a taxa de reconhecimento para o modelo macroscópico foi de 65,02% e para o modelo microscópico foi de 60,96%.



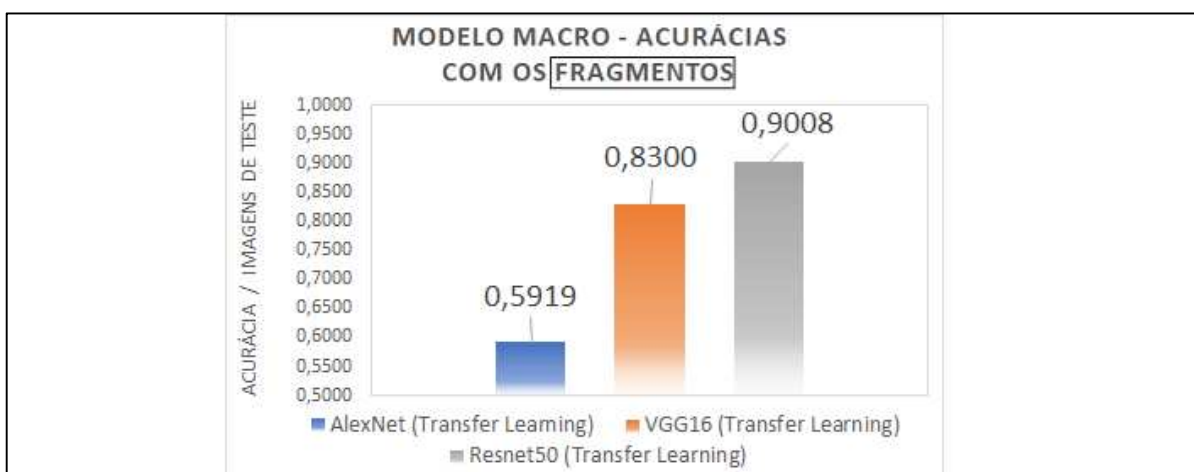


**Figura 26 – Treinamentos com Imagens Redimensionadas**  
**Fonte: Autoria Própria.**

As taxas iniciais obtidas não foram satisfatórias, então foram realizadas as técnicas de Data Augmentation para aumentar as bases, bem como foram definidas as técnicas de fusão de subimagens a serem testadas nos experimentos (vide Seções 3.2.2.1 e 3.2.2.2).

#### 4.2.1 Seleção de Arquitetura e Técnica de Fusão

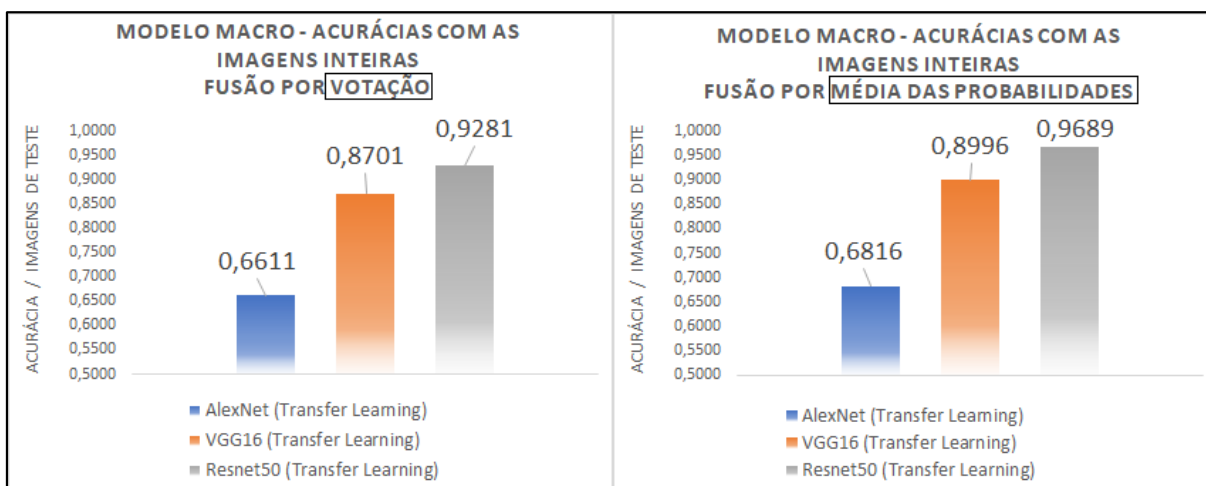
A Figura 27 apresenta as acurácias em relação aos fragmentos de testes da base macroscópica para cada uma das arquiteturas utilizadas. Observa-se que o aumento da base melhorou expressivamente as taxas de reconhecimento.



**Figura 27 – Treinamentos com Data Augmentation – Base Macro - Fragmentos**  
**Fonte: Autoria Própria.**

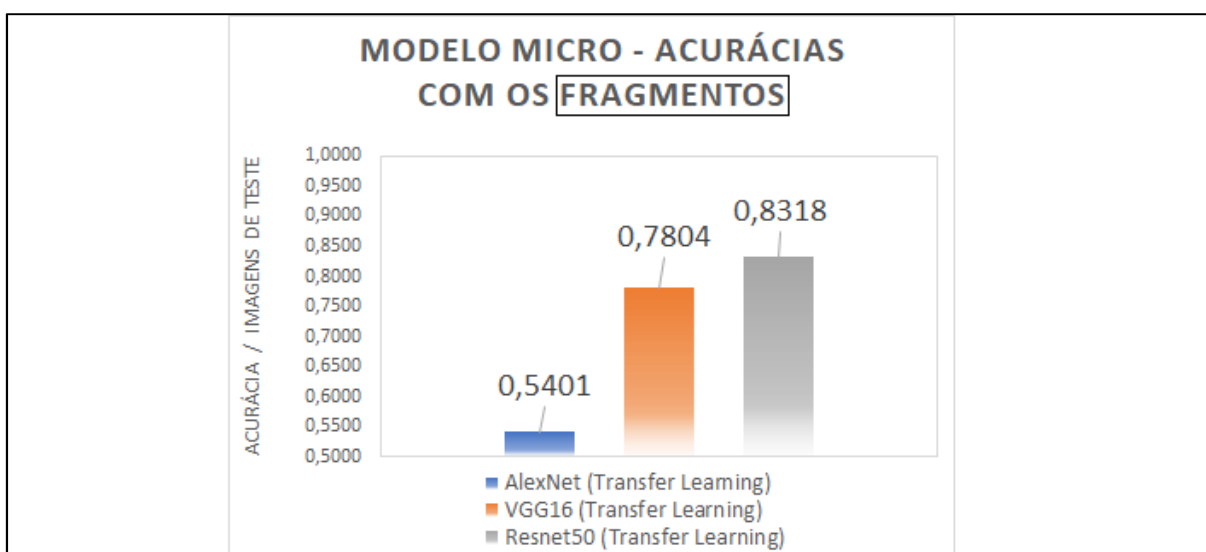
A Figura 28 (esquerda) apresenta as acurácias em relação as imagens inteiras de testes, utilizando fusão por votação. A Figura 28 (direita) apresenta as acurácias utilizando fusão por média das probabilidades.

Nota-se que a fusão por média das probabilidades proporcionou ganhos entre 2 e 4 pontos percentuais nas taxas de reconhecimento.



**Figura 28 – Treinamentos com *Data Augmentation* – Base Macro - Fusões**  
Fonte: Autoria Própria.

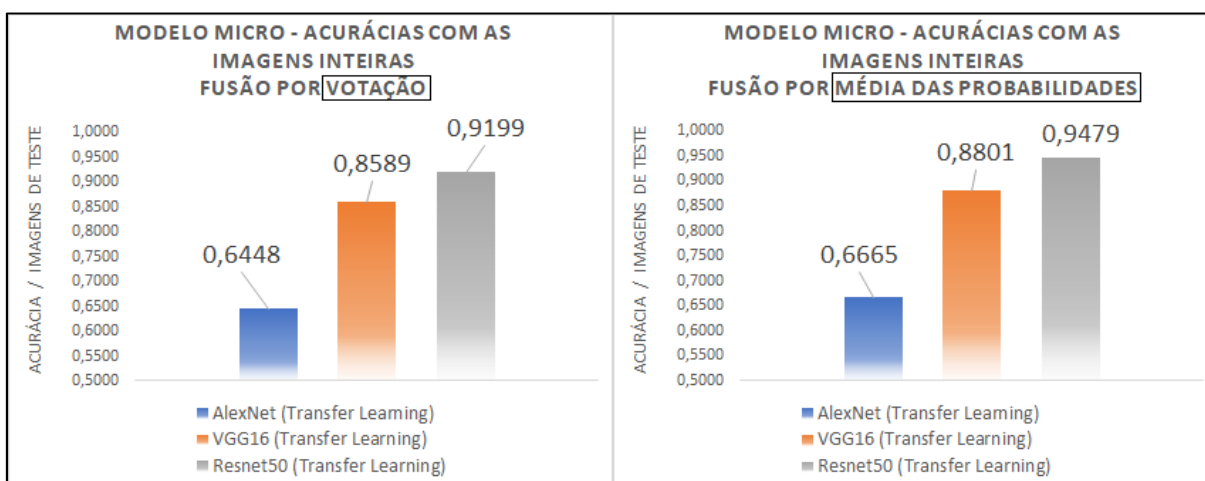
A Figura 29 apresenta as acurácias em relação aos fragmentos de testes da base microscópica para cada arquitetura utilizada. Observa-se que o aumento da base microscópica também melhorou expressivamente as taxas de reconhecimento.



**Figura 29 – Treinamentos com *Data Augmentation* – Base Micro - Fragmentos**  
Fonte: Autoria Própria.

A Figura 30 (esquerda) apresenta as acurácias em relação as imagens inteiras de testes, utilizando fusão por votação. A Figura 30 (direita) apresenta as acurácias utilizando fusão por média das probabilidades.

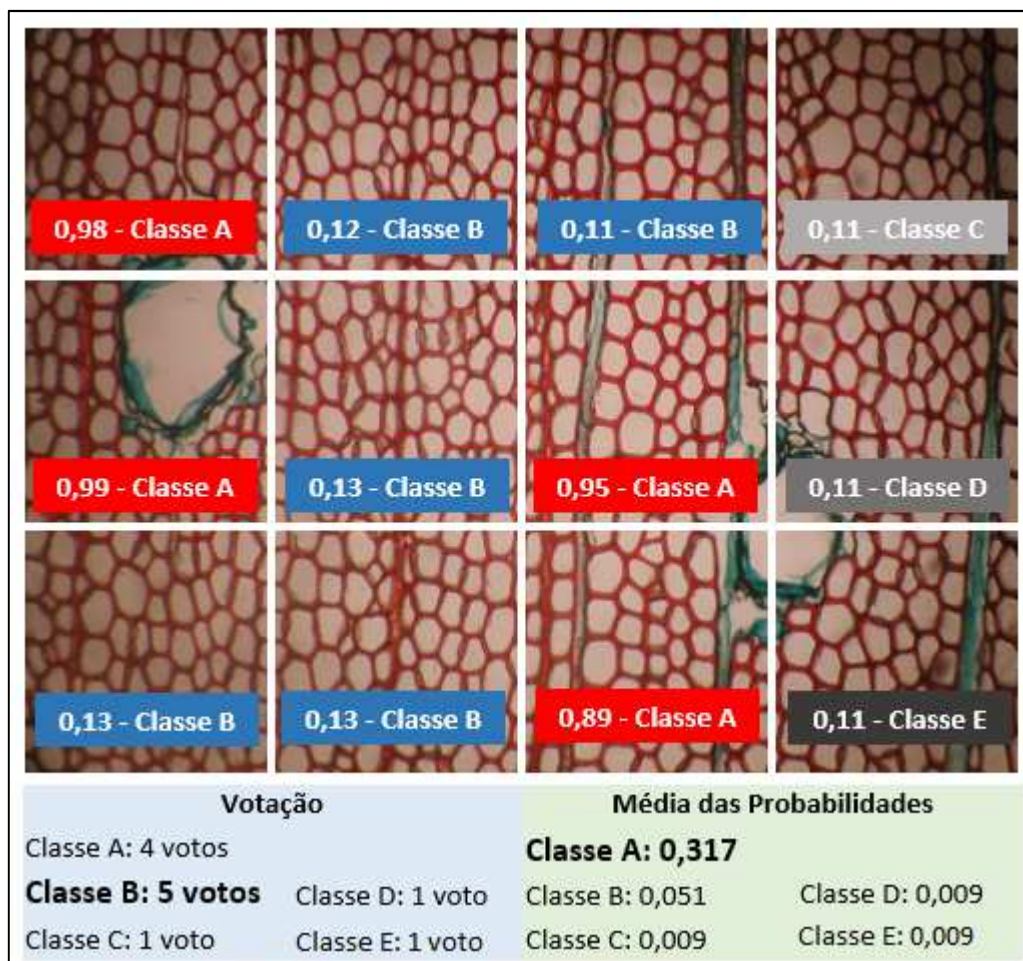
Assim como nos gráficos da base macroscópica, nota-se que a fusão por média das probabilidades também proporcionou ganhos entre 2 e 3 pontos percentuais nas taxas de reconhecimento da base microscópica.



**Figura 30 – Treinamentos com *Data Augmentation* – Base Micro - Fusões**  
**Fonte: Autoria Própria.**

Ao observar alguns exemplos, os quais foram classificados com rótulos diferentes por cada um dos métodos de fusão, observou-se que alguns poucos fragmentos apresentavam probabilidades muito próximas a 1, enquanto vários outros possuíam as possibilidades mais distribuídas ao longo das classes. Neste contexto, a votação elegeu a classe com mais previsões, porém probabilidades não tão altas. Já a média das probabilidades, elegeu a classe para a qual houve maior detecção de características (maior probabilidade), tornando a média maior do que a classe mais votada.

Para melhor contextualizar, observa-se o exemplo apresentado na Figura 31, para a qual espera-se que a previsão seja Classe A. Tal figura demonstra a eleição da Classe B utilizando fusão por votação e a eleição da Classe A utilizando fusão por média de probabilidades. Observa-se que mais fragmentos apostaram na Classe B, porém alguns fragmentos encontraram mais características (maior detecção) referentes à Classe A.

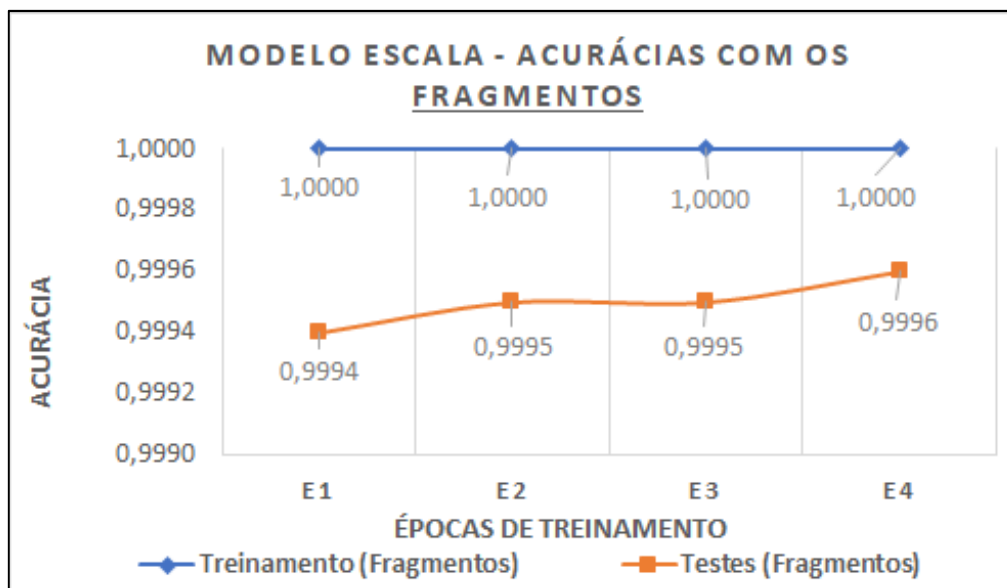


**Figura 31 – Análise das Fusões por Votação e Média das Probabilidades**  
 Fonte: Autoria Própria.

Após a análise dos ganhos nas acurácias, elegeu-se a arquitetura ResNet50 e a técnica de fusão por média das probabilidades para serem utilizadas nas próximas fases.

#### 4.2.2 Modelo para Reconhecimento de Escala

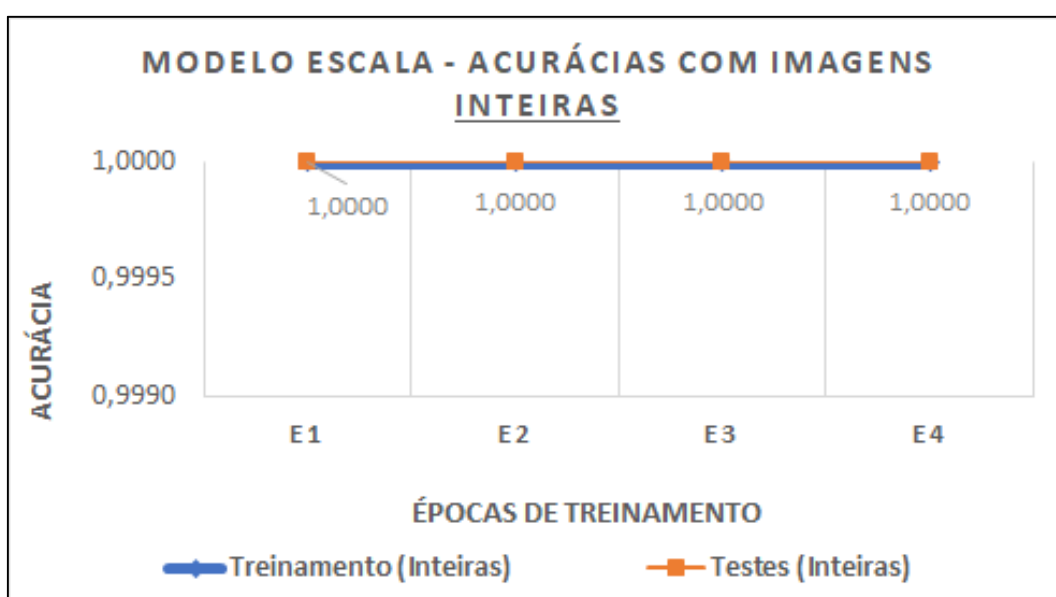
Utilizando a arquitetura ResNet50 pré-inicializada, já na primeira época de treinamento, as taxas de reconhecimento atingiram 100% para o conjunto de treinamento e 99,94% para o conjunto de testes com as subimagens. O modelo foi treinado por 4 épocas, melhorando a taxa de reconhecimento para as subimagens de testes, atingindo 99,96% (vide Figura 32).



**Figura 32 – Treinamento da Base de Escala – Análise dos Fragmentos**  
 Fonte: Autoria Própria.

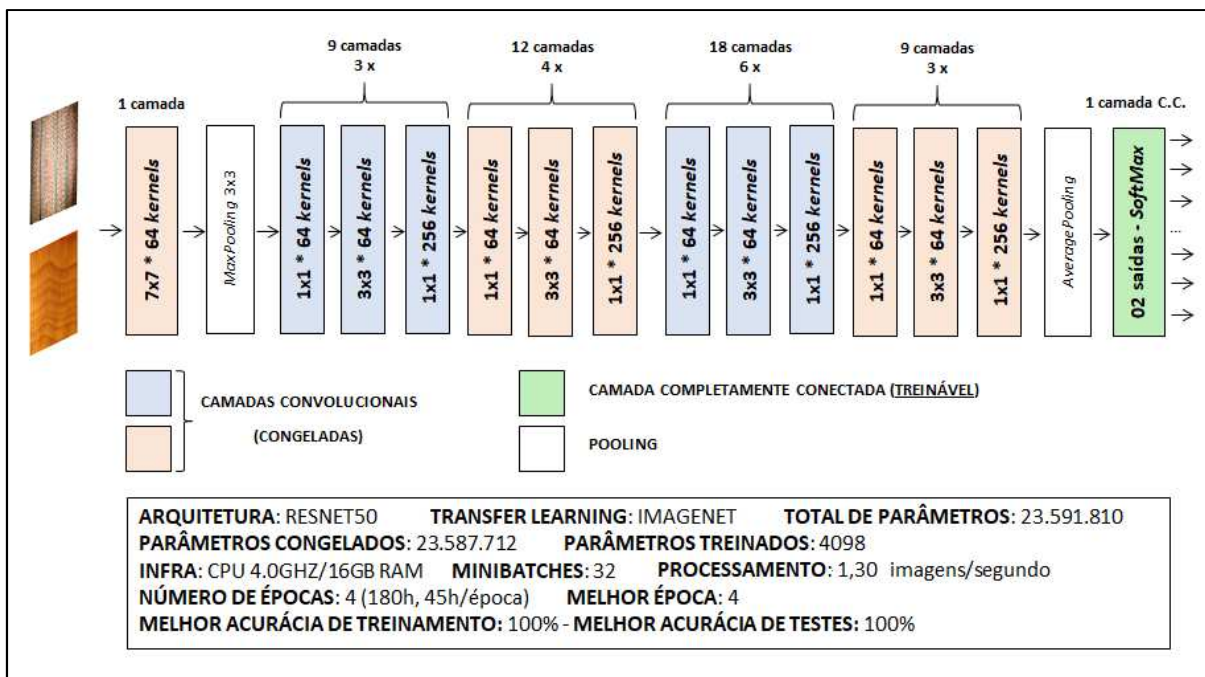
Antes de prosseguir com mais épocas, foi realizada também a análise das taxas de reconhecimento em relação as imagens inteiras por meio da técnica de fusão das classificações por média das probabilidades, apresentada na Seção 3.2.2.2. Desta forma, foi possível saber se a quantidade de épocas de treinamento realizada foi ou não suficiente para atingir a taxa máxima de reconhecimento.

A Figura 33 apresenta 100% de taxa de reconhecimento desde a primeira época de treinamento, após a fusão das classificações das 12 subimagens.



**Figura 33 – Treinamento da Base de Escala – Análise das Imagens Inteiras**  
 Fonte: Autoria Própria.

Após a análise das taxas de reconhecimento, elegeu-se o modelo obtido na época 4 de treinamento como sendo o Modelo de Escala final, já que as taxas com as subimagens foram melhores. A arquitetura e os detalhes de construção de tal modelo podem ser visualizados na Figura 34.

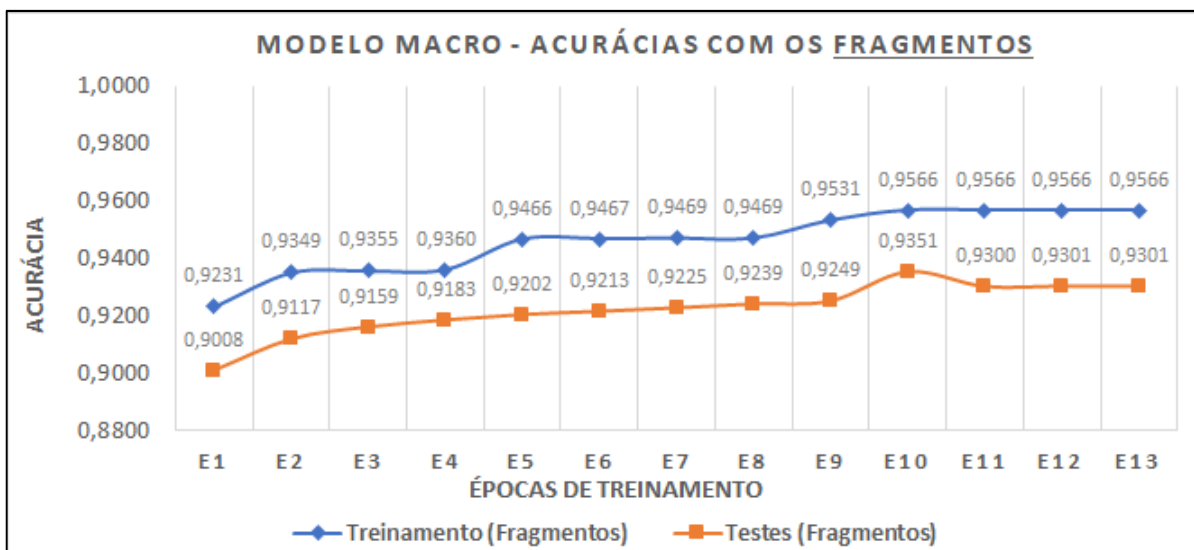


**Figura 34 – Arquitetura Final – Modelo de Escala**  
**Fonte: Autoria Própria.**

#### 4.2.3 Modelo para Reconhecimento Macroscópico

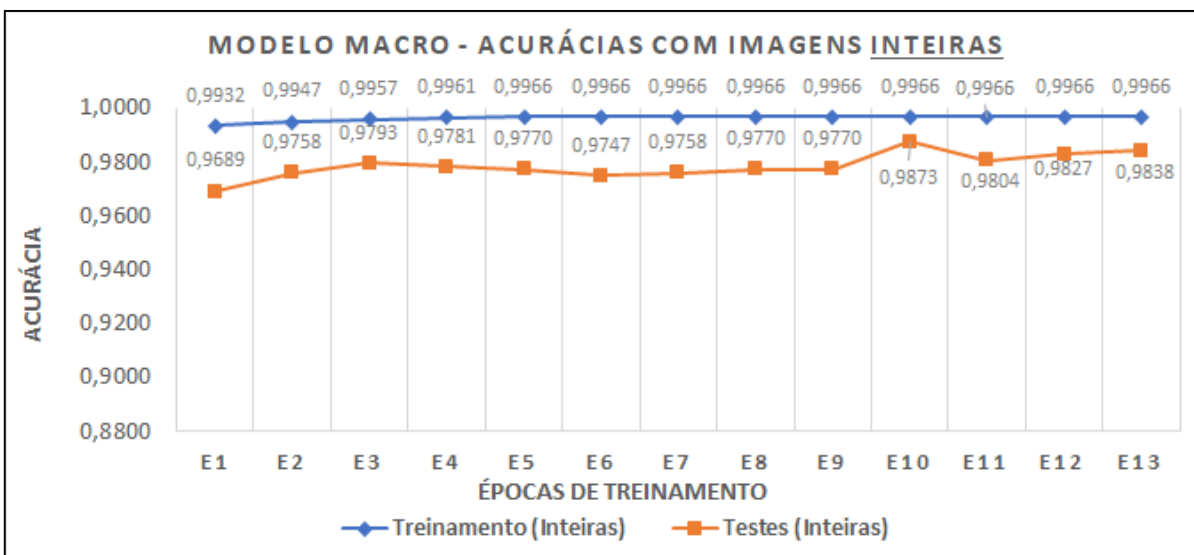
Utilizando a arquitetura ResNet50 pré-inicializada, o modelo foi treinado por 13 épocas, atingindo taxas de 95,66% para o conjunto de treinamento e 93,51% para o conjunto de testes com as subimagens. Tais taxas referem-se a Época 10 de treinamento, eleita inicialmente como melhor época (vide Figura 35).





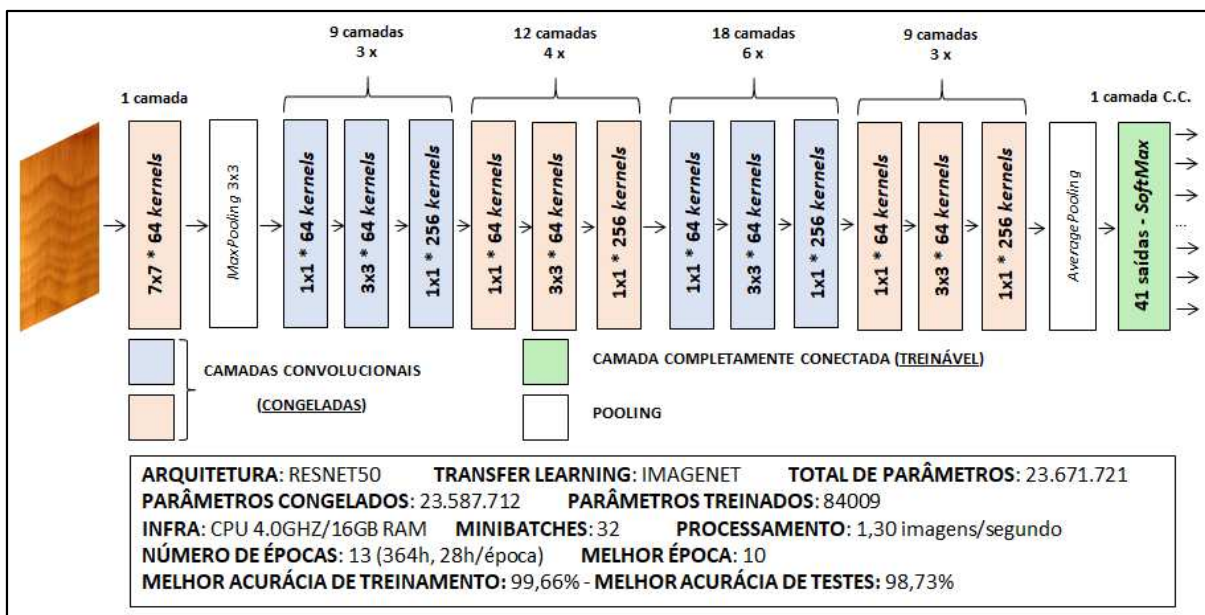
**Figura 35 – Treinamento da Base Macro – Análise dos Fragmentos**  
 Fonte: Autoria Própria.

Foi realizada também a análise das taxas de reconhecimento em relação as imagens inteiras pós fusão, as quais atingiram 99,66% para o conjunto de treinamento e 98,73% para o conjunto de testes na Época 10 (vide Figura 36).



**Figura 36 – Treinamento da Base Macro – Análise das Imagens Inteiras**  
 Fonte: Autoria Própria.

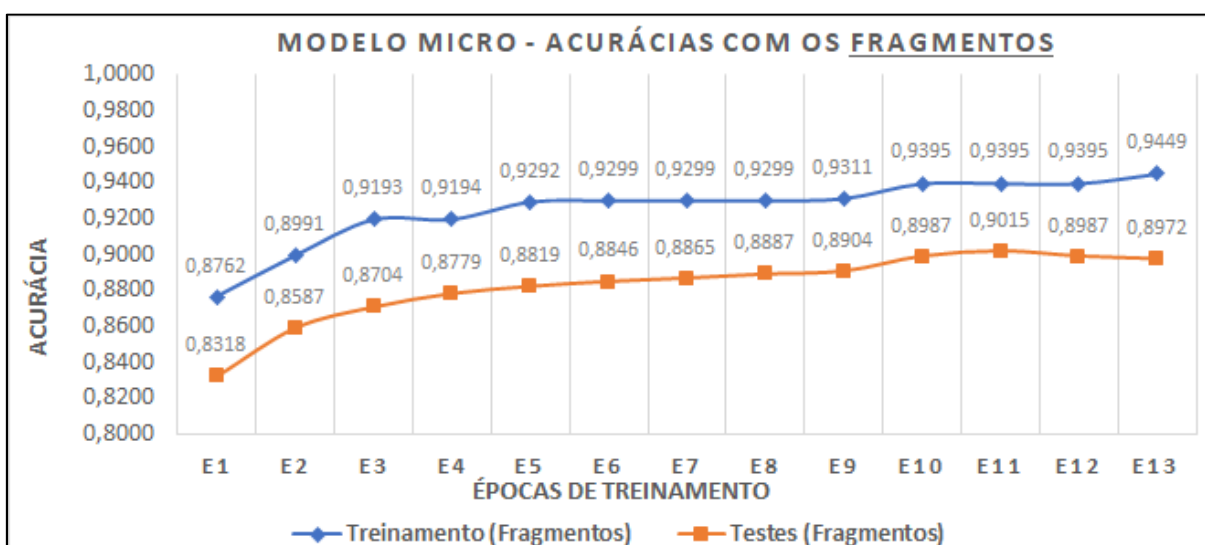
Após a análise das taxas de reconhecimento, elegeu-se o modelo obtido na época 10 de treinamento como sendo o Modelo Macro final. A arquitetura e os detalhes de construção de tal modelo podem ser visualizados na Figura 37.



**Figura 37 – Arquitetura Final – Modelo Macro**  
 Fonte: Autoria Própria.

#### 4.2.4 Modelo para Reconhecimento Microscópico

Utilizando a arquitetura ResNet50 pré-inicializada, o modelo foi treinado por 13 épocas, atingindo taxas de 93,95% para o conjunto de treinamento e 90,15% para o conjunto de testes com as subimagens. Tais taxas referem-se a Época 11 de treinamento, eleita inicialmente como melhor época (vide Figura 38).



**Figura 38 – Treinamento da Base Micro – Análise dos Fragmentos**  
 Fonte: Autoria Própria.



Foi realizada também a análise das taxas de reconhecimento em relação as imagens inteiras pós fusão, as quais atingiram 99,94% para o conjunto de treinamento e 99,11% para o conjunto de testes na Época 11 (vide Figura 39).

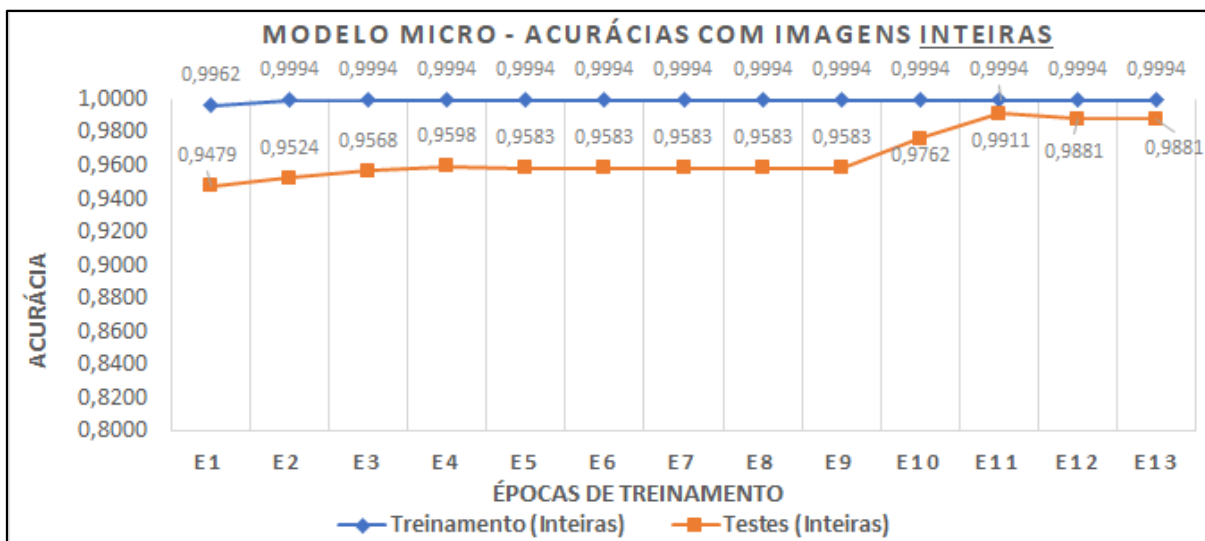


Figura 39 – Treinamento da Base Micro – Análise das Imagens Inteiras  
Fonte: Autoria Própria.

Após a análise das taxas de reconhecimento, elegeu-se o modelo obtido na época 11 de treinamento como sendo o Modelo Micro final. A arquitetura e os detalhes de construção de tal modelo podem ser visualizados na Figura 40.

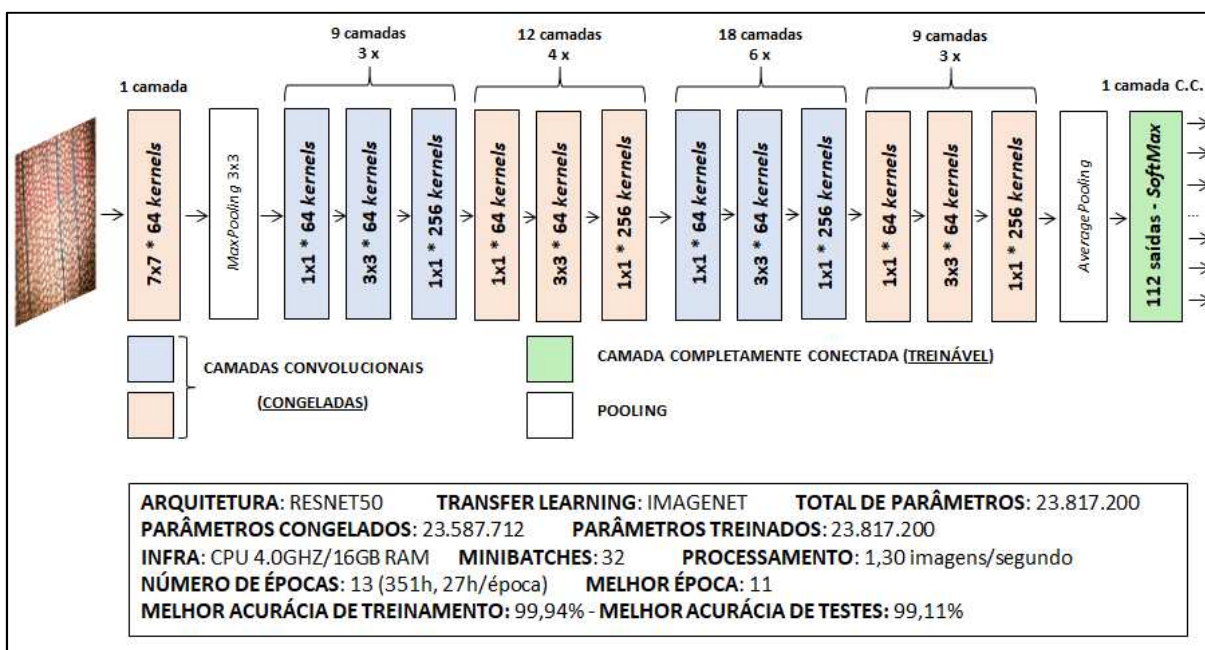


Figura 40 – Arquitetura Final – Modelo Micro  
Fonte: Autoria Própria.

## 5 ANÁLISE E DISCUSSÃO DE RESULTADOS

Esta seção aborda as taxas de reconhecimento e erros de classificação a partir dos três modelos estabelecidos na fase de experimentos. Apresenta também a comparação entre as taxas de reconhecimento alcançadas a partir de modelos tradicionais e de pesquisas correlatas. Por fim, esta seção descreve como ficou o software/produto final, o qual foi implementado utilizando os três modelos de CNN.

### 5.1 TAXAS DE RECONHECIMENTO E ANÁLISE DE ERROS: CNN

A partir dos modelos finais de CNN, foi possível identificar as taxas de reconhecimento, bem como identificar quantas e quais imagens foram confundidas pelo classificador. O Quadro 7 apresenta a relação entre modelos, acertos e erros, no qual observa-se ausência de erro para o Modelo de Escala, 1,27% para o Modelo Macro (11 imagens) e 0,89% (06 imagens) para o Modelo Micro.

**Quadro 7 – Taxas de Reconhecimento dos Melhores Modelos**

MODELO	BASE DE TREINAMENTO IMAGENS INTEIRAS		BASE DE TESTES IMAGENS INTEIRAS	
	ACERTOS	ERROS	ACERTOS	ERROS
Melhor Modelo de Escala	3641 imagens 100%	<i>(Não houve)</i>	1541 imagens <b>100%</b>	<i>(Não houve)</i>
Melhor Modelo Macro	2066 imagens 99,66%	07 imagens 0,34%	858 imagens <b>98,73%</b>	11 imagens 1,27%
Melhor Modelo Micro	1564 imagens 99,74%	04 imagens 0,26%	666 imagens <b>99,11%</b>	06 imagens 0,89%

Fonte: Autoria Própria.

Uma vez que o classificador apresenta a probabilidade média para cada classe do modelo, cada imagem confundida foi analisada a fim de averiguar se a segunda ou terceira predição referia-se a classe correta.

O Quadro 8 apresenta esta relação para os 11 erros cometidos pelo Modelo Macro. Nota-se que, das 11 imagens confundidas, 8 foram corretamente classificadas como segunda predição e 2 foram corretamente classificadas como

terceira predição. Logo, apenas 1 das imagens não foi corretamente classificada em 1ª, 2ª ou 3ª instância.

**Quadro 8 – Erros de Classificação do Modelo Macroscópico**

	IMAGEM MACRO	LABEL ESPERADO	LABEL PREVISTO (MAIOR PROBAB.)	LABEL PREVISTO (2ª MAIOR PROBAB.)	LABEL PREVISTO (3ª MAIOR PROBAB.)
1	0150.jpg	01 – Acrocarpus	34 - Pinus (42,59%) - <b>ERRO</b>	01 – Acrocarpus (33,06%) - <b>ACERTO</b>	11 – Cedro (11,58%) - <b>ERRO</b>
2	0341.jpg	03 – Andiroba	41 – Virola (22,59%) - <b>ERRO</b>	03 – Andiroba (21,28%) - <b>ACERTO</b>	21 – Imbuia (14,53%) - <b>ERRO</b>
3	0545.jpg	05 – Araucária	34 - Pinus (54,79%) - <b>ERRO</b>	05 – Araucária (34,43%) - <b>ACERTO</b>	01 – Acrocarpus (05,45%) - <b>ERRO</b>
4	0546.jpg	05 – Araucária	34 – Pinus (58,00%) - <b>ERRO</b>	05 – Araucária (19,54%) - <b>ACERTO</b>	01 – Acrocarpus (06,94%) - <b>ERRO</b>
5	0742.jpg	07 – Bracatinga	21 - Imbuia (33,39%) - <b>ERRO</b>	07– Bracatinga (30,74%) - <b>ACERTO</b>	17 – Eucalipto (09,87%) - <b>ERRO</b>
6	0745.jpg	07 – Bracatinga	21 – Imbuia (37,29%) - <b>ERRO</b>	17 – Eucalipto (21,98%) - <b>ERRO</b>	07– Bracatinga (21,31%) - <b>ACERTO</b>
7	1127.jpg	11 – Cedro	29 – Mogno (68,16%) - <b>ERRO</b>	41 – Virola (09,22%) - <b>ERRO</b>	11 – Cedro (05,84%) - <b>ACERTO</b>
8	1463.jpg	14 – Cumaru	17 – Eucalipto (42,14%) - <b>ERRO</b>	14 – Cumaru (27,40%) - <b>ACERTO</b>	03 – Andiroba (08,37%) - <b>ERRO</b>
9	2280.jpg	22 – Ipê	21 – Imbuia (53,25%) - <b>ERRO</b>	03 – Andiroba (18,66%) - <b>ERRO</b>	17 – Eucalipto (07,96%) - <b>ERRO</b>
10	2778.jpg	27 – Machaerium	22 – Ipê (81,22%) - <b>ERRO</b>	27 – Machaerium (13,35%) - <b>ACERTO</b>	26 – Louro (02,93%) - <b>ERRO</b>
11	2993.jpg	29 – Mogno	21 – Imbuia (47,88%) - <b>ERRO</b>	29 – Andiroba (24,71%) - <b>ACERTO</b>	16 – Curupixa (13,39%) - <b>ERRO</b>

Fonte: Autoria Própria.

Os 6 erros cometidos pelo Modelo Micro são apresentados no Quadro 9. Nota-se que, das 6 imagens confundidas, 4 foram corretamente classificadas como segunda predição e 1 foi corretamente classificada como terceira predição. Logo, apenas 1 das imagens não foi corretamente classificada como 1ª, 2ª ou 3ª instância.

**Quadro 9 – Erros de Classificação do Modelo Microscópico**

	IMAGEM MICRO	LABEL ESPERADO	LABEL PREVISTO (MAIOR PROBAB.)	LABEL PREVISTO (2ª MAIOR PROBAB.)	LABEL PREVISTO (3ª MAIOR PROBAB.)
1	016015.jpg	016 – Pinales articulata	001 – Ginkgoales biloba (40,81%) - <b>ERRO</b>	016 – Pinales articulata (40,10%) - <b>ACERTO</b>	011 – Pinales lindleyi (04,71%) - <b>ERRO</b>
2	016019.jpg	016 – Pinales articulata	010 – Pinales Cupressus (23,96%) - <b>ERRO</b>	035 – Pinales baccata (15,25%) - <b>ERRO</b>	033 – Pinales sp (11,21%) - <b>ERRO</b>
3	016020.jpg	016 – Pinales articulata	035 – Pinales baccata (39,19%) - <b>ERRO</b>	016 – Pinales articulata (19,79%) - <b>ACERTO</b>	010 – Pinales Cupressus (17,27%) - <b>ERRO</b>
4	020015.jpg	020 – Pinales atlantica	001 – Ginkgoales biloba (34,95%) - <b>ERRO</b>	020 – Pinales atlantica (28,22%) - <b>ACERTO</b>	019 – Pinales vejari (09,58%) - <b>ERRO</b>

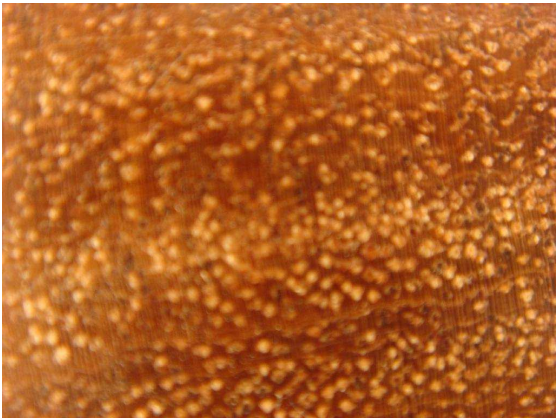
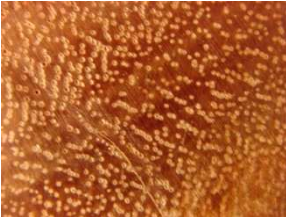







	IMAGEM MICRO	LABEL ESPERADO	LABEL PREVISTO (MAIOR PROBAB.)	LABEL PREVISTO (2ª MAIOR PROBAB.)	LABEL PREVISTO (3ª MAIOR PROBAB.)
5	024020.jpg	024 – Pinales abies	037 – Pinales distichum (28,54%) - <b>ERRO</b>	025 – Pinales arizonica (28,13%) - <b>ERRO</b>	024 – Pinales abies (17,52%) - <b>ACERTO</b>
6	037019.jpg	037 – Pinales distichum	036 – Pinales sempervirens (55,11%) - <b>ERRO</b>	037 – Pinales distichum (32,35%) - <b>ACERTO</b>	028 – Pinales gregii (06,66%) - <b>ERRO</b>

Fonte: Autoria Própria.

A fim de avaliar a similaridade entre algumas das imagens confundidas pelos modelos, realizou-se a análise visual entre exemplos da classe esperada e exemplos das classes previstas como 1ª, 2ª e 3ª predições (três maiores probabilidades).

O Quadro 10 apresenta a confusão cometida pelo Modelo Macro, o qual classificou uma imagem de rótulo “14 – Cumaru” como sendo “17 – Eucalipto”. Nota-se que são muito parecidas as características visuais presentes em exemplos das classes esperada e prevista.

Quadro 10 – Visualização de Erro de Classificação do Modelo Macroscópico

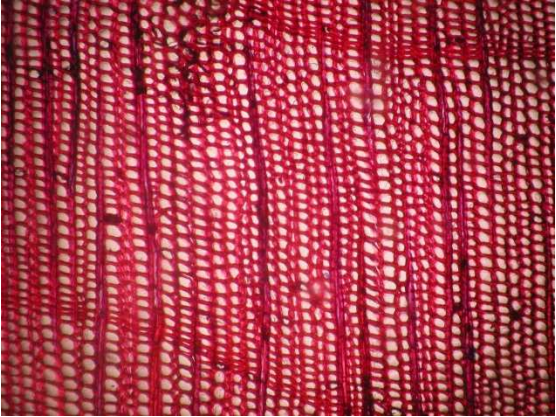
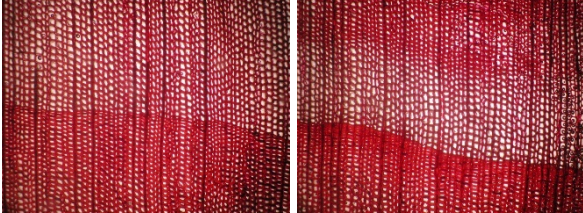

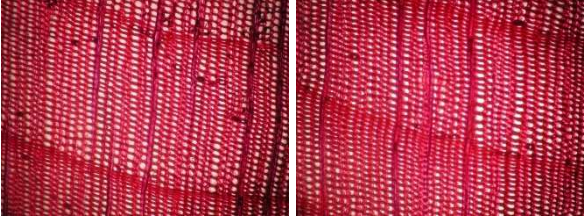
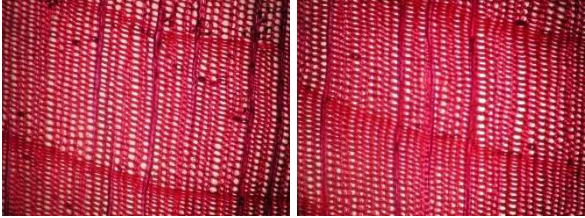
Imagem Macro Confundida: 1463.JPG	<b>Maior</b> Probabilidade = 42,14% Label Previsto: 17 – Eucalipto - <b>ERRO</b>	
		
	<b>2ª</b> Maior Probabilidade = 27,40% - <b>ACERTO</b> Label Previsto: 14 - Cumaru	
		
Label <b>Esperado</b> : 14 - Cumaru	<b>3ª</b> Maior Probabilidade = 08,37% - <b>ERRO</b> Label Previsto: 03 – Andiroba	
		
		

Fonte: Autoria Própria.



O Quadro 11 apresenta a confusão cometida pelo Modelo Micro, o qual classificou uma imagem de rótulo “016 - Pinales articulata” como sendo da classe “035 – Pinales baccata”. Nota-se, também, que são muito parecidas as características visuais presentes em exemplos das classes esperada e prevista.

**Quadro 11 – Visualização de Erro de Classificação do Modelo Microscópico**

<p>Imagem Micro Confundida: 016020.JPG</p>	<p><b>Maior</b> Probabilidade = 33,85% Label Previsto: <b>035 – Pinales baccata - ERRO</b></p>
	 <p><b>2ª</b> Maior Probabilidade = 24,25% - <b>ERRO</b> Label Previsto: <b>020 – Pinales atlantica</b></p> 
<p>Label <b>Esperado</b>: 016 - Pinales articulata</p>	<p><b>3ª</b> Maior Probabilidade = 14,26% - <b>ACERTO</b> Label Previsto: <b>016 – Pinales articulata</b></p>
	

Fonte: Autoria Própria.

### 5.1.1 Ranking de Reconhecimento com as Cinco Maiores Probabilidades

Para cada imagem classificada, é possível observar a probabilidade para cada uma das possíveis classes do modelo. Com isso, tornou-se possível realizar uma análise de taxa de reconhecimento em relação ao conjunto de testes (completamente desconhecido pelo modelo), considerando as cinco maiores probabilidades.

O Quadro 12 apresenta as taxas de reconhecimento ao considerar a primeira maior probabilidade (TOP 1), as duas maiores probabilidades (TOP 2), as três maiores probabilidades (TOP 3), as quatro maiores probabilidades (TOP 4) e as cinco maiores probabilidades (TOP 5).

**Quadro 12 – Ranking de Reconhecimento – Cinco Maiores Probabilidades**

	MODELO DE ESCALA		MODELO MACRO		MODELO MICRO	
	ACERTOS	ERROS	ACERTOS	ERROS	ACERTOS	ERROS
<b>TOP 1</b>	<b>100%</b>	0%	98,73%	1,27%	99,11%	0,89%
<b>TOP 2</b>	<b>100%</b>	0%	99,65%	0,35%	99,70%	0,30%
<b>TOP 3</b>	<b>100%</b>	0%	99,88%	0,12%	99,85%	0,15%
<b>TOP 4</b>	<b>100%</b>	0%	<b>100%</b>	0%	99,85%	0,15%
<b>TOP 5</b>	<b>100%</b>	0%	<b>100%</b>	0%	<b>100%</b>	0%

Fonte: Autoria Própria.

Nota-se que, ao considerar as quatro maiores probabilidades (TOP 4) para cada classificação do modelo macro, foi possível atingir 100% de taxa de reconhecimento. Já para o modelo micro, foi possível atingir 100% de taxa de reconhecimento ao considerar as cinco maiores probabilidades (TOP 5).

## 5.2 COMPARAÇÃO DOS RESULTADOS COM O ESTADO DA ARTE

Ao comparar as taxas de reconhecimento obtidas a partir dos experimentos realizados com abordagens tradicionais (Seção 4.1), em relação as taxas obtidas com CNN em TOP 1, notou-se ganhos bastante expressivos, descritos a seguir.

Utilizando GLCM e RNA tradicionais, os melhores resultados foram de 66,90% para a base macroscópica e 66,52% para a base microscópica. Já para CNN, os melhores resultados atingiram 100% de taxa de reconhecimento para a base de escala, 98,73% para a base macroscópica e 99,11% para a base microscópica (Vide Quadro 13).

Quadro 13 – Taxas de Reconhecimento – GLCM vs. Deep Learning

MODELO	ABORDAGEM TRADICIONAL (GLCM + RNA)		DEEP LEARNING (CNN)	
	TAXA DE RECONHECIMENTO	TEMPO DE TREINAMENTO	TAXA DE RECONHECIMENTO	TEMPO DE TREINAMENTO
Melhor Modelo de Escala	100%	4h - extração 39min - treinamento	100%	45h/época
Melhor Modelo Macro	66,90%	2h - extração 7min - treinamento	98,73%	28h/época
Melhor Modelo Micro	66,52%	2h - extração 32min - treinamento	99,11%	27h/época

Fonte: Autoria Própria.

O ganho ao utilizar abordagens com DL foi de 31,83 pontos percentuais para o Modelo Macro e 32,59 pontos percentuais para o Modelo Micro. Para o Modelo de Escala, a taxa de reconhecimento foi de 100%, portanto não houve ganho. Entretanto, o tempo de treinamento de cada modelo baseado em CNN foi muito maior em relação ao tempo de extração de características e treinamento com GLCM.

Ao comparar as taxas de reconhecimento obtidas a partir dos experimentos com CNN em relação a taxas obtidas por pesquisas correlatas (Seção 2.3.1.1), notou-se ganhos muito expressivos tanto em pontos percentuais quanto em complexidade de construção dos modelos, já que a escolha das técnicas de extração de características foram automaticamente aprendidas pelos modelos.

O Quadro 14 apresenta a relação dos resultados obtidos em comparação a os resultados de Paula *et al.* (2014), os quais trabalharam com abordagens tradicionais com imagens macroscópicas, Martins *et al.* (2015), os quais trabalharam com métodos tradicionais a partir de imagens microscópicas e Hafemann, Cavalin e Oliveira (2014) que trabalharam com ambas as bases utilizando CNN.

Quadro 14 – Taxas de Reconhecimento – Pesquisas Correlatas

MODELO	PAULA <i>et al.</i> (2014)	HAFEMANN, CAVALIN E OLIVEIRA (2014)	MARTINS <i>et al.</i> (2015)	PRESENTE PESQUISA
	TAXA DE RECONHECIMENTO			
Melhor Modelo de Escala	(Não Realizado)	(Não Realizado)	(Não Realizado)	100%
Melhor Modelo Macro	97,77% (Tradicional)	95,77% (CNN)	(Não Realizado)	98,73%
Melhor Modelo Micro	(Não Realizado)	97,32% (CNN)	93,03% (Tradicional)	99,11%

Fonte: Autoria Própria.

Em relação as melhores taxas utilizando abordagens tradicionais, o ganho ao utilizar DL foi de 0,96 pontos percentuais para o Modelo Macro e 6,08 pontos percentuais para o Modelo Micro. Além dos ganhos em pontos percentuais, ressalta-se a vantagem da abstração das técnicas de extração de características. Já para as melhores taxas utilizando CNN, os resultados desta pesquisa representam ganho de 2,96 pontos percentuais para o Modelo Macro e 1,79 pontos percentuais para o Modelo Micro.

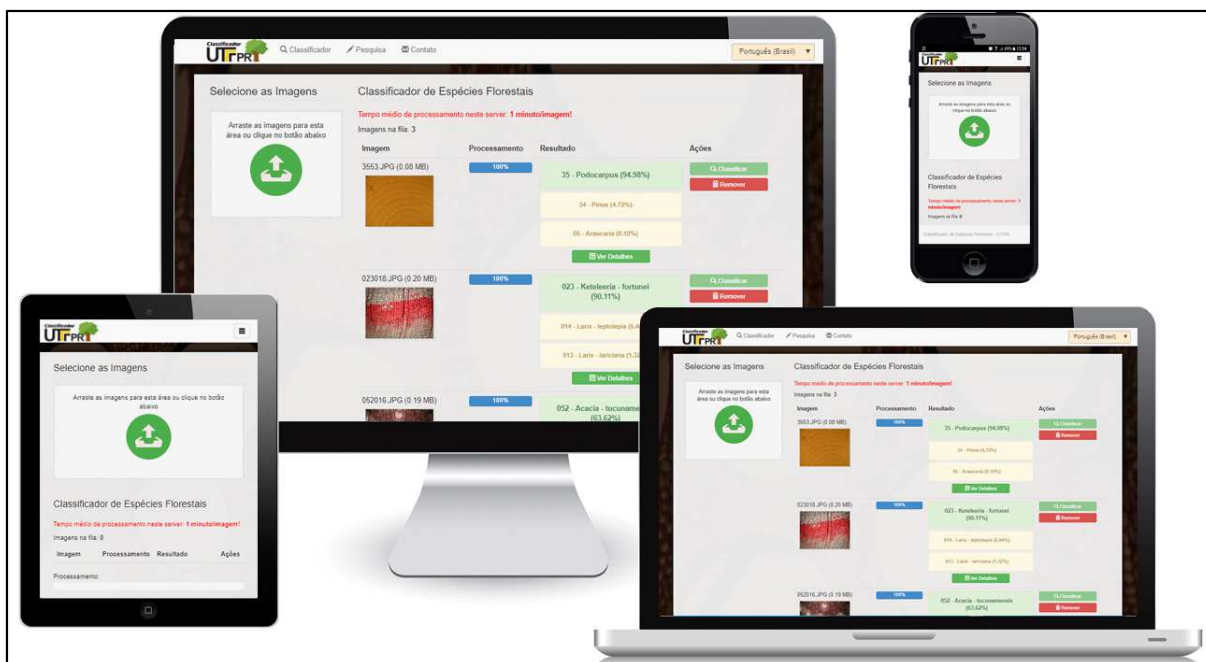
Nas abordagens tradicionais, as atividades que envolvem testes de pré-processamento de imagens e definição de filtros para extração de dados são relativamente complexas e ocupam muito tempo do pesquisador. Já para os modelos baseados em CNN, tais atividades são aprendidas e ajustadas dinamicamente pelo modelo em tempo de treinamento, tornando a construção menos dependente da intervenção do pesquisador.

### 5.3 O PRODUTO FINAL: SOFTWARE WEB

Conforme mencionado na Seção 3.2.2.3, o produto final é um software web que pode ser acessado por meio do seguinte endereço eletrônico: <http://reconhecimentoflorestal.md.utfpr.edu.br>. Encontra-se disponível nos idiomas Português (Brasil) e Inglês (EUA).

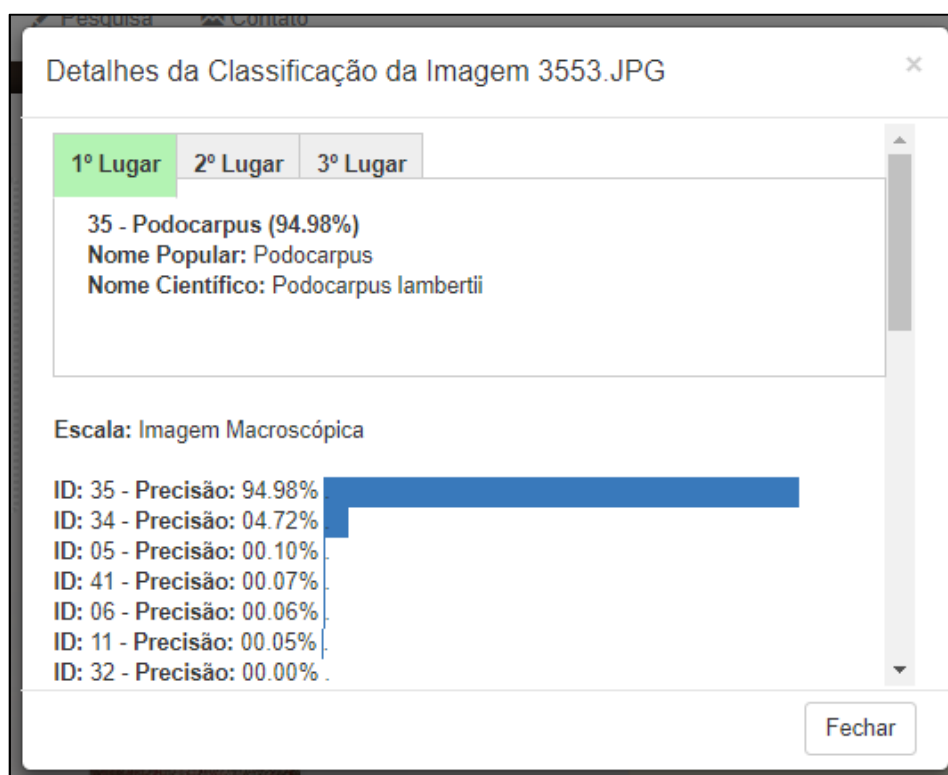
A Figura 41 mostra a tela principal, a qual apresenta uma área para a realização do *upload* das imagens e outra para processamento das imagens.





**Figura 41 – Interface Gráfica do Produto/Software**  
**Fonte: Autoria Própria.**

Além de apresentar a espécie prevista, o software permite visualizar a lista de probabilidades para em cada classificação, bem como consultar algumas informações extras sobre a respectiva espécie (Vide Figura 42).



**Figura 42 – Tela de Detalhes do Produto/Software**  
**Fonte: Autoria Própria.**

Para melhorar a interpretação das classificações realizadas pelo software, foi incluída uma regra que se comporta como um limiar de alerta. Em outras palavras, se houver baixa precisão na classificação de escala ou na classificação da espécie, o software exibe um alerta com a seguinte mensagem “*Esta imagem pode não ser de uma espécie conhecida por este sistema*”.

A intenção desta regra é identificar possíveis classificações de imagens contendo espécies que ainda não são conhecidas pelos modelos ou até mesmo imagens que não se tratam de espécies florestais, como pode ser observado no exemplo apresentado na Figura 43.



**Figura 43 – Alerta de Possível Espécie Não Conhecida**  
**Fonte: Autoria Própria.**

Ao realizar testes com imagens de espécies não conhecidas pelos modelos e de imagens aleatórias (cachorros, carros, casas, etc.), estipulou-se como limiar de alerta, as precisões abaixo de 85% para o Modelo de Escala ou abaixo de 15% para os Modelos Macro e Micro.

## 6 CONSIDERAÇÕES FINAIS

Um dos objetivos deste trabalho foi o de construir três modelos baseados em DL e acoplar à um produto final (software web). Este produto encontra-se disponível em: <http://reconhecimentoflorestal.md.utfpr.edu.br>.

Para a elaboração da pesquisa e a construção do produto, algumas fases foram definidas. Inicialmente, foram realizados alguns experimentos com as bases macro e microscópica, utilizando GLCM (uma das técnicas tradicionais) e os resultados foram armazenados para posterior comparação com os resultados dos modelos em DL.

A segunda fase (foco deste trabalho) consistiu na realização de alguns testes iniciais a fim de definir a melhor arquitetura de CNN a ser utilizada para cada modelo. Logo no início, identificou-se uma grande dificuldade: as bases de imagens originais não possuíam exemplos suficientes para o treinamento. Com isso, realizou-se a fragmentação da imagem original em 12 subimagens, seguidas de 5 transformações. Logo, para cada imagem original, foram extraídas 72 subimagens, possibilitando a construção de bases muito maiores.

O treinamento de modelos em CNN mostrou-se um processo bastante custoso computacionalmente, tanto que existe um conceito denominado *Transfer Learning*, o qual permite a aquisição de arquiteturas pré-treinadas e permitem congelar o treinamento de algumas camadas da arquitetura, para manter o conhecimento já existente e agilizar o processamento. Mesmo aplicando este conceito, os melhores modelos foram obtidos somente após treinamentos de 180h (escala), 364h (macro) e 351h (micro) em CPU.

Os experimentos realizados com GLCM (técnica tradicional) resultaram em taxas de reconhecimento de 66,90% para espécies da base macroscópica, 66,52% para espécies da base microscópica e 100% para a base de escala. Já para os experimentos realizados com CNN (*Deep Learning*), as taxas de reconhecimento foram de 100% para o modelo de escala (2 possíveis classes), 98,73% para o modelo macro (41 possíveis classes) e 99,11% para o modelo micro (112 possíveis classes), as taxas de ambas as abordagens se referem aos conjuntos de testes (imagens completamente desconhecidas pelos modelos).

Como os modelos foram treinados para classificar subimagens, as imagens inteiras de entrada são divididas em 12 subimagens. Para cada subimagem obtém-se um vetor de probabilidades correspondentes às N possíveis classes do modelo. Em seguida, realiza-se uma técnica de fusão das classificações, a qual consiste na obtenção da probabilidade para cada classe em cada subimagem e realiza-se a média. Desta forma, identifica-se a maior probabilidade média que corresponde a classe eleita.

Para o Modelo de Escala não houve erro, para o Modelo Macro foram confundidas apenas 11 imagens e para o Modelo Micro, apenas 6 imagens foram confundidas. Entretanto, 15 destes 17 erros correspondem a imagens que foram corretamente classificadas ao considerar a segunda e terceira maior probabilidade.

Apesar do alto custo computacional para a construção dos modelos, a CNN mostrou-se uma técnica de DL muito interessante, justamente por exigir o mínimo de pré-processamento nos dados quando comparada a metodologias tradicionais, como GLCM por exemplo. Em outras palavras, a extração de características das imagens torna-se um processo aprendido durante o treinamento, tornando a construção de modelos otimizados e menos dependentes de intervenção humana.

Por fim, a pesquisa atendeu a todos objetivos propostos e, em resposta a hipótese desta pesquisa: “é possível classificar espécies florestais a partir de imagens utilizando técnicas de Deep Learning”, conclui-se que foi possível sim, bem como os resultados foram os melhores já obtidos, tornando-se o estado da arte em taxas de reconhecimento com as respectivas bases de imagens.

## 6.1 TRABALHOS FUTUROS

Ao decorrer do desenvolvimento desta pesquisa, observou-se a possibilidade de novos trabalhos que podem contribuir para o estado da arte:

- **Ajuste das Camadas Convolucionais:** Nesta pesquisa, os pesos das camadas convolucionais foram congelados na fase de treinamento, isto é, o Back-Propagation ocorreu apenas na camada completamente conectada. Neste contexto, propõe-se realizar novos experimentos de forma que algumas das camadas sejam

descongeladas (tornem-se treináveis) e avalie-se as taxas de reconhecimento após o ajuste de pesos de tais camadas convolucionais;

- **Incremento de Espécies:** Realização do aumento real das bases de imagens a fim de melhorar o conhecimento dos modelos. Além disso, acrescentar exemplares de novas espécies, de forma a enriquecer o conhecimento da ferramenta;
- **Avaliação Comparativa do Produto:** Realizar uma pesquisa com especialistas da área de Engenharia Florestal, a fim de avaliar o uso do software em relação ao procedimento manual, ainda realizado por muitos laboratórios florestais;
- **Avaliação de Tempo (CPU vs. GPU):** Outra abordagem que pode ser avaliada é em relação ao tempo de treinamento dos mesmos modelos em GPU, uma vez que os modelos construídos nesta pesquisa foram treinados somente em CPU.

## REFERÊNCIAS

- AKRAM, S; DAR, M; QUYOUN, A. **Document Image Processing – A Review**. International Journal of Computer Applications – Vol 10. DOI: 10.5120/1475-1991. EUA. Novembro de 2010.
- AMABIS, J. M; MARTHO, G. R. **Fundamentos da Biologia Moderna**. 4. ed. Volume único. São Paulo: Ed. Moderna, 1995.
- ANDREARCZYK, V; WHELAN, P. **Using Filter Banks in Convolutional Neural Networks for Texture Classification**. Pattern Recognition Letters, Volume 84, p.63-69. DOI: 10.1016/j.patrec.2016.08.016. 2016.
- AREL, I; ROSE, D. C; KAMOWSKI, T. P. **Deep Machine Learning - A New Frontier in Artificial Intelligence Research**. IEEE Computational Intelligence Magazine, 5. DOI: 10.1109/MCI.2010.938364. EUA, p.13-18. Novembro de 2010.
- BARALDI, A; PARMIGGIANI, F. **An Investigation of the Textural Characteristics Associated with Gray Level Co-occurrence Matrix Statistical Parameters**, 1995 – IEEE Transactions on Geoscience and Remote Sensing, 33(2) p.293–304.
- BEBIS, G; EGBERT, D; SHAH, M. **Review of computer vision education**. IEEE Transactions on Education - Vol. 46, 5. DOI: 10.1109/TE.2002.808280. EUA, p.2-21. Fevereiro de 2003.
- BEZERRA, E. **Introdução à Aprendizagem Profunda**. Artigo – 31º Simpósio Brasileiro de Banco de Dados – SBBBD2016 – Salvador, 2016.
- CAVALIN, P., KAPP, M., OLIVEIRA, L. S., **An Adaptive Multi-Level Framework for Forest Species Recognition**. 2015 Brazilian Conference on Intelligent Systems (BRACIS 2015). DOI: 10.1109/BRACIS.2015.21. Natal, 2015
- CORADIN, V. T. R. **Identificação Anatômica de Madeira**. Serviço Florestal Brasileiro. Brasília: Abr Cine Vídeo, 2008. 1 (25 min.), DVD. Disponível em: <<https://www.youtube.com/watch?v=94QP-zc05kg>>. Acesso em: 16 out. 2016.
- CORREIA, E. A. S. **Análise e Dimensionamento de Estruturas de Madeira**. 163 f. Dissertação (Mestrado) - Programa de Mestrado Integrado em Engenharia Civil, Faculdade de Engenharia da Universidade do Porto, Porto - Portugal, 2009.

GERHARDT, T. E; SILVEIRA, D. T. **Métodos de Pesquisa**. 1. ed. Porto Alegre: UFRGS, 2009.

GIL, A. C. **Métodos e Técnicas de Pesquisa Social**. 6. ed. São Paulo: Atlas, 2008.

GOLDSCHMIDT, R. R. **Uma Introdução à Inteligência Artificial** – Fundamentos, Ferramentas e Aplicações. 1. ed. Rio de Janeiro: IST-Rio, 2010.

GOLLIN, J. **Deductive vs. inductive language learning**. ELT Journal – Vol. 52. EUA, DOI: 10.1093/elt/52.1.88. p.88-89. Janeiro de 1998.

GONZALEZ, R. C; WOODS, R. E. **Processamento Digital de Imagens**. 3. ed. São Paulo: Pearson / Prentice Hall, 2010.

GOODFELLOW, I; BENGIO, Y; COURVILLE, A. **Deep Learning**: Adaptive Computation and Machine Learning series. Cambridge, MA: MIT Press: 2016.

HAFEMANN, L. G; CAVALIN, P; OLIVEIRA, L.S. **Forest Species Recognition using Deep Convolutional Neural Networks**. International Conference on Pattern Recognition (ICPR2014). DOI: 10.1109/ICPR.2014.199. p. 1103-1107, Stockholm, Sweden, 2014.

HAFEMANN, L. G; CAVALIN, P. ; OLIVEIRA, L.S, Sabourin, R., **Transfer Learning between Texture Classification Tasks using Convolutional Neural Networks**. International Joint Conference on Neural Networks (IJCNN 2015). DOI: 10.1109/IJCNN.2015.7280558, p. 1939-1945, Killarney, Ireland, 2015.

HARALICK, R. M, SHANMUGAM, K. e DINSTEN, I. **Textural Features for Image Classification**, IEEE Transactions on Systems, Man and Cybernetics, 3 p.610- 620. 1973.

HAYKIN, S. **Redes Neurais: Princípios e Prática**. 2. ed. Porto Alegre: Bookman, 2001.

HE, K; ZHANG, X; REN, S; SUN, J. **Deep Residual Learning for Image Recognition**. Computer Vision and Pattern Recognition (CVPR), 2016 IEEE Conference on. DOI: 10.1109/CVPR.2016.90. Dezembro de 2016.

HUANG, Z; PAN, Z; LEI, B. **Transfer Learning with Deep Convolutional Neural Network for SAR Target Classification with Limited Labeled Data.** Remote Sensing. 9. 907. DOI: 10.3390/rs9090907. Agosto de 2017.

HUQQANIA, A. A; SCHIKUTAA, E; YEA, S; CHEN, P. **Multicore and GPU Parallelization of Neural Networks for Face Recognition.** International Conference on Computational Science. P. 349 – 358. DOI: 10.1016/j.procs.2013.05.198. ICCS: 2013.

ITO, R. H; KIM, H. Y; SALCEDO, W. J. **Classificação de Texturas Invariante a Rotação Usando Matriz de Co-ocorrência.** Dept. Eng. Sistemas Eletrônicos, Escola Politécnica Universidade de São Paulo, 2009.

IZOTOV, P. Y; Kazanskiy, N. L; Golovashkin, D. L. **CUDA - Enabled Implementation of a Neural Network Algorithm for Handwritten Digit Recognition.** Optical Memory and Neural Networks. Vol.20, p. 98-106. DOI: 10.3103/S1060992X11020032. Junho de 2011.

JAIN, A. K; DUIN, R. P. W ; MAO, J.. **Statistical pattern recognition: a review.** Pattern Analysis and Machine Intelligence, IEEE Transactions on, 22(1):4-37, 2000.

JORDAN, M. I; MITCHELL, T. M. **Machine learning: Trends, perspectives, and prospects.** Science Magazine – Vol. 349. Nova Iorque - EUA, DOI: 10.1126/science.aaa8415. p.255-260. Julho de 2015.

KITTER, J. K; HATEF, M; DUIN, R. P. W; MATAS, J. **On combining classifiers.** IEEE Trans. on Pattern Analysis and Machine Intelligence, 20:226-239, 1998.

KOCH, G; HAAG, V; HEINZ, I; RICHTER, H. G; SCHMITT, U. **Control of Internationally Traded Timber - The Role of Macroscopic and Microscopic Wood Identification against Illegal Logging.** Journal of Forensic Research. Indore – Índia, DOI: 10.4172/2157-7145.1000317. Dezembro de 2015.

KRIZHEVSKY, A; SUTSKEVER, I; HINTON, G. E; **ImageNet Classification with Deep Convolutional Neural Networks.** NIPS 2012: Neural Information Processing Systems, Lake Tahoe, Nevada, 2012.

LAPTEV, D; BUHMANN, J. M. **Transformation-Invariant Convolutional Jungles.** IEEE Conference on Computer Vision and Pattern Recognition (CVPR). DOI: 10.1109/CVPR.2015.7298923. Junho de 2015.



LESK, A. M. **Introdução à Bioinformática**. 1. ed. Artmed, 2008.

LECUN, Y; BENGIO, Y; HINTON, G. **Deep Learning**. Nature International Weekly Journal of Science - EUA, DOI:10.1038/nature14539. Maio de 2015.

LECUN, Y; BOSER, B; DENKER, J. S; HENDERSON, D; HOWARD, R. E; HUBBARD, W; JACKEL, L. D. **Backpropagation applied to handwritten zip code recognition**. Neural Computation Magazine. Vol. 1, p. 541 – 551. DOI: 10.1162/neco.1989.1.4.541. Dezembro de 1989.

MARCONI, M. A; LAKATOS, E. M. **Fundamentos de Metodologia Científica**. 5. ed. São Paulo: Atlas, 2003.

MANSANO, A; MATSUOKA, J. A; AFONSO, L. C. S; *et al.* **Improving Image Classification Through Descriptor Combination**. Artigo – SIBGRAPI 2012 - Conference on Graphics, Patterns and Images, Ouro Preto, 2012.

MARTINS, J. G; OLIVEIRA, L. S; BRITTO, A. S; SABOURIN, R. **Forest species recognition based on dynamic classifier selection and dissimilarity feature vector representation**. Machine Vision and Applications, Springer Berlin Heidelberg, DOI 10.1007/s00138-015-0659-0, April 2015.

MOHAN, S; VENKATACHALAPATHY, K; RAI, A. K. **Wood Species Classification and Identification System**. International Journal of Engineering Sciences & Research Technology. Indore – Índia, p.847-853. Junho de 2014.

NAIR, V; HINTON, G. F. **Rectified Linear Units Improve Restricted Boltzmann Machines**. Artigo, Department of Computer Science - University of Toronto, Toronto - Canadá, 2010.

NEVES, L. A. P; VIEIRA NETO, H; GONZAGA, A. **Avanços em Visão Computacional**. Curitiba: Omnipax, 2012.

NIU, J; TANG, W; XU, F; ZHOU, X; SONG, Y. **Global Research on Artificial Intelligence from 1990–2014: Spatially-Explicit Bibliometric Analysis**. ISPRS International Journal of Geo-Information. DOI:10.3390/ijgi5050066. EUA. Maio de 2016.

OLIVEIRA, L. F; NARLOCH, A. L. M; KIST, D. M; SOARES, M. P; MENEGHELLO, G. E; CAVALHEIRO, G. G. H; TILLMANN, M. A. A. **Extração de Características de Forma utilizando matriz de co-ocorrência e Atributos de Haralick.** Workshop de Visão Computacional (WVC). Programa de Pós-graduação em Ciência e Tecnologia de Sementes, Universidade Federal de Pelotas, Pelotas, 2012.

PAULA FILHO, P. L.; OLIVEIRA, L. S; NISGOSKI, S; BRITTO JR, A. S. **Forest species recognition using macroscopic images.** Machine Vision and Applications, Springer Berlin Heidelberg, DOI 10.1007/s00138-014-0592-7, Janeiro de 2014.

PEDRINI, H; SCHWARTZ, W. R. **Análise de Imagens Digitais: Princípios, Algoritmos e Aplicações.** São Paulo: Thomson, 2008.

REZENDE, S. O. **Sistemas Inteligentes – Fundamentos e aplicações.** 1. ed. Barueri: Editora Manole Ltda, 2003.

RUSSAKOVSKY, O; DENG, J; SU, H; KRAUSE, J; SATHEESH, S; MA, S; HUANG, Z; KARPATY, A; KHOSLA, A; BERNSTEIN, M; BERG, A. C; FEI-FEI, L. **ImageNet Large Scale Visual Recognition Challenge.** International Journal of Computer Vision (IJCV). DOI: 10.1007/s11263-015-0816-y. EUA. Janeiro de 2015.

SAFARI, P. **Deep Learning for Sequential Pattern Recognition.** 113 f. Dissertação (Mestrado) - Master of Science, Technische Universität München, München - Alemanha, 2013.

SANTOS, S. N. C. **Análise Estrutural e de Superfície de Madeiras Submetidas a Tratamento Térmico e Degradadas pelo Fungo Pycnoporus Sanguineus.** 85 f. Dissertação (Mestrado) – Escola de Engenharia de São Carlos - Desenvolvimento, Caracterização e Aplicação de Materiais. DOI: 10.11606/D.18.2015.tde-22092015-105627. São Carlos, 2015.

SIMONYAN, K; ZISSERMAN, A. **Very Deep Convolutional Networks for Large-Scale Image Recognition.** 3rd IAPR Asian Conference on Pattern Recognition (ACPR). DOI: 10.1109/ACPR.2015.7486599. Novembro de 2015.

TALWAR, A; KUMAR, Y. **Machine Learning: An artificial intelligence methodology.** International Journal Of Engineering And Computer Science, Vol. 2. Índia, DOI: 10.8535/ijecs. p.3400-3404. Dezembro de 2013.

TARAVAT, A; PROUD, S; PERONACI, S; FRATE, F; OPPELT, N. **Multilayer Perceptron Neural Networks Model for Meteosat Second Generation SEVIRI Daytime Cloud Masking**. Remote Sensing Journal, Vol. 7. Itália, DOI:10.3390/rs70201529. p.1529-1539. Fevereiro de 2015.

VARGAS, A. C. G; CARVALHO, A. M. P; VASCONCELOS, C. N. **Um Estudo sobre Redes Neurais Convolucionais e sua Aplicação em Detecção de Pedestres**. Artigo, SIBGRAPI – XXIX Conference on Graphics, Patterns and Images, São José dos Campos, 2016.

ZENID, G. J; CECCANTINI, G. C. T. **Identificação Macroscópica de Madeiras**. Artigo - Laboratório de Madeira e Produtos Derivados, Centro de Tecnologia de Recursos Florestais - Instituto de Pesquisas Tecnológicas do Estado de São Paulo - IPT, Porto – São Paulo, 2007.

ZEILER, M. D; FERGUS, R. **Visualizing and Understanding Convolutional Networks**. Artigo - Computer Vision – ECCV 2014: 13th European Conference, Zurich, Switzerland, September, 2014.

## ANEXO A - Composição da Base de Imagens Macroscópicas

**Tabela 4 – Composição da Base de Imagens Macroscópicas**

<b>ID</b>	<b>Nome Popular</b>	<b>Nome Científico</b>	<b>Quantidade</b>
01	Acrocarpus	<i>Acrocarpus fraxinifolius</i>	53
02	Amapá	<i>Brosimum paraensis</i>	63
03	Andiroba	<i>Carapa guianensis</i>	43
04	Angelim Pedra	<i>Hymenolobium petraeum</i>	99
05	Araucaria	<i>Araucaria angustifolia</i>	63
06	Assacu	<i>Hura crepitans</i>	41
07	Bracatinga	<i>Mimosa scabrella</i>	48
08	Cabriuva Vermelha	<i>Myroxylon balsamum</i>	75
09	Castanheira	<i>Bertolothia excelsa</i>	72
10	Cedrinho	<i>Erisma uncinatum</i>	98
11	Cedro	<i>Cedrela ssilis</i>	37
12	Cedrorana	<i>Cedrelinga catenaeformis</i>	99
13	Cinamomo	<i>Melia azedarach</i>	56
14	Cumarú	<i>Dipteryx sp</i>	67
15	Cupiuba	<i>Goupia glabra</i>	51
16	Curupixa	<i>Micropholis venulosa</i>	78
17	Eucalipto	<i>Eucalyptus sp</i>	99
18	Freijo	<i>Cordia goeldiana</i>	53
19	Goiabão	<i>Pouteria pachycarpa</i>	94
20	Grevilea	<i>Grevilea robusta</i>	86
21	Imbuia	<i>Ocotea porosa</i>	99
22	Ipê	<i>Handroanthus sp</i>	99
23	Itauba	<i>Mezilaurus itauba</i>	64
24	Jatob	<i>Hymenaea sp</i>	82
25	Jequitiba	<i>Cariniana estrellensis</i>	55
26	Louro	<i>Laurus nobilis</i>	46
27	Machaerium	<i>Machaerium sp</i>	87
28	Massaranduba	<i>Manilkara huberi</i>	92
29	Mogno	<i>Swietenia macrophylla</i>	96
30	Louro amarelo	<i>Cordia sp</i>	51
31	Pau amarelo	<i>Euxylophora paraensis</i>	96
32	Pau mar m	<i>Balfourodendron riedelianum</i>	99
33	Peroba rosa	<i>Aspidosperma polyneuron</i>	43
34	Pinus	<i>Pinus sp</i>	79
35	Podocarpus	<i>Podocarpus lambertii</i>	62
36	Quaruba	<i>Vochysia sp</i>	59
37	Roxinho	<i>Peltogyne sp</i>	58
38	Sucupira	<i>Bowdichia sp</i>	99
39	Tatajuba	<i>Bagassa guianensis</i>	58
40	Tauari	<i>Couratari sp</i>	63
41	Virola	<i>Virola surinamensis</i>	80

Fonte: Paula et al. (2014).

## ANEXO B - Composição da Base de Imagens Microscópicas

**Tabela 5 – Composição da Base de Imagens Microscópicas**

ID	Ordem	Família	Gênero	Espécie	Quantidade
001	Ginkgoales	Ginkgoaceae	<i>Ginkgo</i>	<i>biloba</i>	20
002	Pinales	Araucariaceae	<i>Agathis</i>	<i>becarii</i>	20
003	Pinales	Araucariaceae	<i>Araucaria</i>	<i>angustifolia</i>	20
004	Pinales	Cephalotaxaceae	<i>Cephalotaxus</i>	<i>drupacea</i>	20
005	Pinales	Cephalotaxaceae	<i>Cephalotaxus</i>	<i>harringtonia</i>	20
006	Pinales	Cephalotaxaceae	<i>Torreya</i>	<i>nucifera</i>	20
007	Pinales	Cupressaceae	<i>Calocedrus</i>	<i>decurrens</i>	20
008	Pinales	Cupressaceae	<i>Chamaecyparis</i>	<i>formosensis</i>	20
009	Pinales	Cupressaceae	<i>Chamaecyparis</i>	<i>pisifera</i>	20
010	Pinales	Cupressaceae	<i>Cupressus</i>	<i>arizonica</i>	20
011	Pinales	Cupressaceae	<i>Cupressus</i>	<i>lindleyi</i>	20
012	Pinales	Cupressaceae	<i>Fitzroya</i>	<i>cupressoides</i>	20
013	Pinales	Cupressaceae	<i>Larix</i>	<i>laricina</i>	20
014	Pinales	Cupressaceae	<i>Larix</i>	<i>leptolepis</i>	20
015	Pinales	Cupressaceae	<i>Larix</i>	<i>sp</i>	20
016	Pinales	Cupressaceae	<i>Tetraclinis</i>	<i>articulata</i>	20
017	Pinales	Cupressaceae	<i>Widdringtonia</i>	<i>cupressoides</i>	20
018	Pinales	Pinaceae	<i>Abies</i>	<i>religiosa</i>	20
019	Pinales	Pinaceae	<i>Abies</i>	<i>vejari</i>	20
020	Pinales	Pinaceae	<i>Cedrus</i>	<i>atlantica</i>	20
021	Pinales	Pinaceae	<i>Cedrus</i>	<i>libani</i>	20
022	Pinales	Pinaceae	<i>Cedrus</i>	<i>sp</i>	20
023	Pinales	Pinaceae	<i>Keteleeria</i>	<i>fortunei</i>	20
024	Pinales	Pinaceae	<i>Picea</i>	<i>abies</i>	20
025	Pinales	Pinaceae	<i>Pinus</i>	<i>arizonica</i>	20
026	Pinales	Pinaceae	<i>Pinus</i>	<i>caribaea</i>	20
027	Pinales	Pinaceae	<i>Pinus</i>	<i>elliottii</i>	20
028	Pinales	Pinaceae	<i>Pinus</i>	<i>gregii</i>	20
029	Pinales	Pinaceae	<i>Pinus</i>	<i>maximinoi</i>	20
030	Pinales	Pinaceae	<i>Pinus</i>	<i>taeda</i>	20
031	Pinales	Pinaceae	<i>Pseudotsuga</i>	<i>macrolepis</i>	20
032	Pinales	Pinaceae	<i>Tsuga</i>	<i>canadensis</i>	20
033	Pinales	Pinaceae	<i>Tsuga</i>	<i>sp</i>	20
034	Pinales	Podocarpaceae	<i>Podocarpus</i>	<i>lambertii</i>	20
035	Pinales	Taxaceae	<i>Taxus</i>	<i>baccata</i>	20
036	Pinales	Taxodiaceae	<i>Sequoia</i>	<i>sempervirens</i>	20
037	Pinales	Taxodiaceae	<i>Taxodium</i>	<i>distichum</i>	20
038	Ephedrales	Ephedraceae	<i>Ephedra</i>	<i>californica</i>	20
039	Ericales	Lecythydaceae	<i>Cariniana</i>	<i>estrellensis</i>	20
040	Ericales	Lecythydaceae	<i>Couratari</i>	<i>sp</i>	20
041	Ericales	Lecythydaceae	<i>Eschweilera</i>	<i>matamata</i>	20
042	Ericales	Lecythydaceae	<i>Eschweleira</i>	<i>chartaceae</i>	20
043	Ericales	Sapotaceae	<i>Chrysophyllum</i>	<i>sp</i>	20
044	Ericales	Sapotaceae	<i>Micropholis</i>	<i>guianensis</i>	20
045	Ericales	Sapotaceae	<i>Pouteria</i>	<i>pachycarpa</i>	20
046	Fabales	Fabaceae	<i>Copaifera</i>	<i>trapezifolia</i>	20
047	Fabales	Fabaceae	<i>Eperua</i>	<i>falcata</i>	20
048	Fabales	Fabaceae	<i>Hymenaea</i>	<i>courbaril</i>	20

ID	Ordem	Família	Gênero	Espécie	Quantidade
049	Fabales	Fabaceae	<i>Hymenaea</i>	<i>sp</i>	20
050	Fabales	Fabaceae	<i>Schizolobium</i>	<i>parahyba</i>	20
051	Fabales	Fabaceae	<i>Pterocarpus</i>	<i>violaceus</i>	20
052	Fabales	Fabaceae	<i>Acacia</i>	<i>tucunamensis</i>	20
053	Fabales	Fabaceae	<i>Anadenanthera</i>	<i>colubrina</i>	20
054	Fabales	Fabaceae	<i>Anadenanthera</i>	<i>peregrina</i>	20
055	Fabales	Fabaceae	<i>Dalbergia</i>	<i>jacaranda</i>	20
056	Fabales	Fabaceae	<i>Dalbergia</i>	<i>spruceana</i>	20
057	Fabales	Fabaceae	<i>Dalbergia</i>	<i>variabilis</i>	20
058	Fabales	Fabaceae	<i>Dinizia</i>	<i>excelsa</i>	20
059	Fabales	Fabaceae	<i>Enterolobium</i>	<i>schomburgkii</i>	20
060	Fabales	Fabaceae	<i>Inga</i>	<i>sessilis</i>	20
061	Fabales	Fabaceae	<i>Leucaena</i>	<i>leucocephala</i>	20
062	Fabales	Fabaceae	<i>Lonchocarpus</i>	<i>subglaucescens</i>	20
063	Fabales	Fabaceae	<i>Mimosa</i>	<i>bimucronata</i>	20
064	Fabales	Fabaceae	<i>Mimosa</i>	<i>scabrella</i>	20
065	Fabales	Fabaceae	<i>Ormosia</i>	<i>excelsa</i>	20
066	Fabales	Fabaceae	<i>Parapiptadenia</i>	<i>rigida</i>	20
067	Fabales	Fabaceae	<i>Parkia</i>	<i>multijuga</i>	20
068	Fabales	Fabaceae	<i>Piptadenia</i>	<i>excelsa</i>	20
069	Fabales	Fabaceae	<i>Pithecellobium</i>	<i>jupunba</i>	20
070	Gentianales	Rubiaceae	<i>Psychotria</i>	<i>carthagenensis</i>	20
071	Gentianales	Rubiaceae	<i>Psychotria</i>	<i>longipes</i>	20
072	Lamiales	Bignoniaceae	<i>Tabebuia</i>	<i>rosea alba</i>	20
073	Lamiales	Bignoniaceae	<i>Tabebuia</i>	<i>sp</i>	20
074	Lamiales	Oleaceae	<i>Ligustrum</i>	<i>lucidum</i>	20
075	Lurales	Lauraceae	<i>Nectandra</i>	<i>rigida</i>	20
076	Lurales	Lauraceae	<i>Nectandra</i>	<i>sp</i>	20
077	Lurales	Lauraceae	<i>Ocotea</i>	<i>porosa</i>	20
078	Lurales	Lauraceae	<i>Persea</i>	<i>racemosa</i>	20
079	Magnoliales	Annonaceae	<i>Porcelia</i>	<i>macrocarpa</i>	20
080	Magnoliales	Magnoliaceae	<i>Magnolia</i>	<i>grandi ora</i>	20
081	Magnoliales	Magnoliaceae	<i>Talauma</i>	<i>ovata</i>	20
082	Myrtales	Melastomataceae	<i>Tibouchiana</i>	<i>sellowiana</i>	20
083	Myrtales	Myristicaceae	<i>Virola</i>	<i>oleifera</i>	20
084	Myrtales	Myrtaceae	<i>Campomanesia</i>	<i>xanthocarpa</i>	20
085	Myrtales	Myrtaceae	<i>Eucalyptus</i>	<i>globulus</i>	20
086	Myrtales	Myrtaceae	<i>Eucalyptus</i>	<i>grandis</i>	20
087	Myrtales	Myrtaceae	<i>Eucalyptus</i>	<i>saligna</i>	20
088	Myrtales	Myrtaceae	<i>Myrcia</i>	<i>racemulosa</i>	20
089	Polygalales	Vochysiaceae	<i>Erisma</i>	<i>uncinatum</i>	20
090	Polygalales	Vochysiaceae	<i>Qualea</i>	<i>sp</i>	20
091	Polygalales	Vochysiaceae	<i>Vochysia</i>	<i>laurifolia</i>	20
092	Proteales	Proteaceae	<i>Grevillea</i>	<i>robusta</i>	20
093	Proteales	Proteaceae	<i>Grevillea</i>	<i>sp</i>	20
094	Proteales	Proteaceae	<i>Roupala</i>	<i>sp</i>	20
095	Rosales	Moraceae	<i>Bagassa</i>	<i>guianensis</i>	20
096	Rosales	Moraceae	<i>Brosimum</i>	<i>alicastrum</i>	20
097	Rosales	Moraceae	<i>Ficus</i>	<i>gomelleira</i>	20
098	Rosales	Rhamnaceae	<i>Hovenia</i>	<i>dulcis</i>	20
099	Rosales	Rhamnaceae	<i>Rhamnus</i>	<i>frangula</i>	20
100	Rosales	Rosaceae	<i>Prunus</i>	<i>sellowii</i>	20
101	Rosales	Rosaceae	<i>Prunus</i>	<i>serotina</i>	20

<b>ID</b>	<b>Ordem</b>	<b>Família</b>	<b>Gênero</b>	<b>Espécie</b>	<b>Quantidade</b>
102	Rubiales	Rubiaceae	<i>Faramea</i>	<i>occidentalis</i>	20
103	Sapindales	Meliaceae	<i>Cabralea</i>	<i>canjerana</i>	20
104	Sapindales	Meliaceae	<i>Carapa</i>	<i>guianensis</i>	20
105	Sapindales	Meliaceae	<i>Cedrela</i>	<i>ssilis</i>	20
106	Sapindales	Meliaceae	<i>Khaya</i>	<i>ivorensis</i>	20
107	Sapindales	Meliaceae	<i>Melia</i>	<i>azedarach</i>	20
108	Sapindales	Meliaceae	<i>Swietenia</i>	<i>macrophylla</i>	20
109	Sapindales	Rutaceae	<i>Balfourodendron</i>	<i>riedelianum</i>	20
110	Sapindales	Rutaceae	<i>Citrus</i>	<i>aurantium</i>	20
111	Sapindales	Rutaceae	<i>Fagara</i>	<i>rhoifolia</i>	20
112	Sapindales	Simaroubaceae	<i>Simarouba</i>	<i>amara</i>	20

**Fonte: Martins et al. (2015).**