



INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DE MINAS GERAIS
BACHARELADO EM SISTEMAS DE INFORMAÇÃO

**APLICAÇÃO DE REDES NEURAIS
ARTIFICIAIS NO PROCESSO DE
CLASSIFICAÇÃO DE ORQUÍDEAS DO
GÊNERO CATTLEYA**

LUCAS PEREIRA DE AZEVEDO

Orientador: Gabriel Felipe Cândido Novy
Instituto Federal de Minas Gerais

SABARÁ
DEZEMBRO DE 2016

LUCAS PEREIRA DE AZEVEDO

**APLICAÇÃO DE REDES NEURAIS ARTIFICIAIS NO
PROCESSO DE CLASSIFICAÇÃO DE ORQUÍDEAS DO
GÊNERO CATTLEYA**

Trabalho de Conclusão de Curso apresentado ao Curso de Bacharelado em Sistemas de Informação do Instituto Federal de Educação, Ciência e Tecnologia de Minas Gerais, como requisito parcial para a obtenção do título de Bacharel em Sistemas de Informação.

Orientador: Gabriel Felipe Cândido Novy
Instituto Federal de Minas Gerais

INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DE MINAS GERAIS
BACHARELADO EM SISTEMAS DE INFORMAÇÃO
SABARÁ
DEZEMBRO DE 2016

A994a Azevedo, Lucas Pereira de
Aplicação de redes neurais artificiais no processo de classificação de orquídeas do gênero *Cattleya* [manuscrito]. / Lucas Pereira de Azevedo. - 2016.

48 f.: il.

Orientador: Prof. Gabriel Felipe Cândido Novy

Monografia (Bacharelado em Sistemas de Informação) – Instituto Federal de Minas Gerais, *Campus* Sabará.

1. Redes neurais (Computação). – Monografia. 2. Inteligência artificial. – Monografia. 3. Processamento de imagens - Visão por computador. – Monografia. I. Novy, Gabriel Felipe Cândido. II. Instituto Federal de Minas Gerais, *Campus* Sabará, Bacharelado em Sistemas de Informação. III. Título.

CDU 004.8

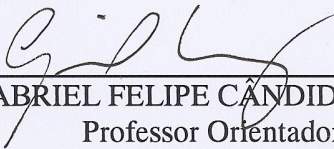
ATA DE DEFESA DO TRABALHO DE CONCLUSÃO DE CURSO de Lucas Pereira de Azevedo

No dia 14 do mês de dezembro do ano de 2016, às 09:00 horas, os professores: Gabriel Felipe Cândido Novy, André Siqueira Ruela e Cristiane Norbiato Targa compareceram para defesa pública do Trabalho de Conclusão de Curso intitulado **Aplicação de Redes Neurais Artificiais no Processo de Classificação de Orquídeas do Gênero Cattleya**, requisito obrigatório para a obtenção do título de Bacharel. Após a apresentação e as observações dos membros da banca avaliadora, ficou definido que o trabalho foi considerado:

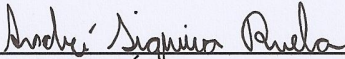
(X) Aprovado () Reprovado.

O resultado final foi comunicado publicamente ao candidato pelo Professor Orientador. Nada mais havendo a tratar, o Professor Orientador a reunião e lavrou a presente ATA, que será assinada por todos os membros participantes da banca avaliadora.

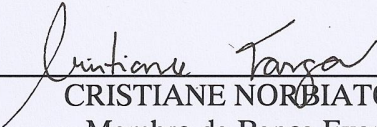
Observações: O aluno tem o prazo de 15 dias para entregar a versão final com as correções sugeridas pela banca.



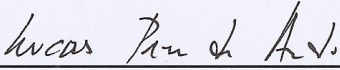
GABRIEL FELIPE CÂNDIDO NOVY
Professor Orientador



ANDRÉ SIQUEIRA RUELA
Membro da Banca Examinadora



CRISTIANE NORBIATO TARGA
Membro da Banca Examinadora



LUCAS PEREIRA DE AZEVEDO
Aluno(a)

Agradecimentos

Agradeço aos meus pais, que sempre se dedicaram para que eu tivesse a melhor educação. Agradeço também aos meus professores, em especial ao meu orientador, pela atenção, paciência e auxílio prestados em meus momentos de dificuldade. E por fim, gostaria de agradecer aos meus colegas que me acompanharam durante esta jornada e a todos aqueles que de alguma forma, contribuíram para que eu alcançasse meus objetivos. A todos, meus mais sinceros agradecimentos.

Resumo

Os processos de reconhecimento e classificação de imagens digitais têm sido alvo de intenso estudo nas últimas décadas. Embora hajam diversas abordagens para solução destes problemas, o desenvolvimento das redes neurais artificiais vêm permitindo que estes processos sejam realizados de forma cada vez mais robusta e confiável. Atualmente, as orquídeas são uma das plantas mais cultivadas mundialmente. Somente o gênero *Cattleya*, conta com dezenas de espécies, algumas das quais apresentam-se bastante similares visualmente, dificultando o processo manual de identificação e classificação. O objetivo deste trabalho é utilizar redes neurais artificiais no processo de classificação de três das espécies pertencentes ao gênero *Cattleya*, comparando os resultados obtidos a partir da utilização de dois tipos destas redes, as redes neurais do tipo perceptron com múltiplas camadas e as redes neurais convolutivas.

Palavras-chave: Classificação de imagens. Redes neurais. Visão computacional.

Abstract

The processes of image recognition and classification has been target of intense study in the last decades. Although there are several approaches to solve these problems, the development of artificial neural networks have allowed these processes to be performed in an increasingly robust and reliable way. Currently, the orchids are one of the most cultivated plants worldwide. The *Cattleya* genus, has dozens of species, some of which are quite similar visually, hindering the manual process of identification and classification. The objective of this work is to use artificial neural networks in the clasification process of three of the species that belongs to the *Cattleya* genus, comparing the obtained results from the use of two categories of these networks, the multilayer perceptron neural networks and the convolutional neural networks.

Keywords: Image classification. Neural networks. Computer vision.

Lista de Figuras

Figura 1 – Comparativo do processo de ampliação de uma imagem tipo <i>raster</i> e vetorial	3
Figura 2 – Representação de cor nos modelos RGB e HSV	4
Figura 3 – Comparativo da decomposição de uma imagem nos modelos HSV e RGB	5
Figura 4 – Etapas da filtragem no domínio da frequência	6
Figura 5 – Exemplo do processo de convolução	6
Figura 6 – Filtro Laplaciano	7
Figura 7 – Resultado obtido após a aplicação do filtro Laplaciano	7
Figura 8 – Modelo de um neurônio humano	8
Figura 9 – Funções de ativação comumente utilizadas	10
Figura 10 – Modelo de um neurônio artificial	11
Figura 11 – Modelo de uma rede do tipo perceptron com múltiplas camadas	11
Figura 12 – Modelo de uma rede do tipo convolutiva	13
Figura 13 – Imagem de três espécies pertencentes ao gênero <i>Cattleya</i>	17
Figura 14 – Exemplo do espelhamento aplicado à base de dados	17
Figura 15 – Modelo de inserção de uma imagem no formato RGB em uma rede MLP	19
Figura 16 – Modelo de inserção do canal matiz de uma imagem no formato HSV em uma rede MLP	20
Figura 17 – Modelo de inserção de uma imagem em escala de cinza com filtro de Laplace em uma rede MLP	20
Figura 18 – Estrutura da rede convolutiva implementada	21
Figura 19 – Sequência de etapas necessárias para realização dos testes propostos	22
Figura 20 – Comparativo das matrizes de confusão - Base de dados original	25
Figura 21 – Comparativo dos tempos de treinamento - Base de dados original	26
Figura 22 – Comparativo das matrizes de confusão - Base de dados aumentada	28
Figura 23 – Comparativo dos tempos de treinamento - Base de dados aumentada	29

Lista de Tabelas

Tabela 1 – Resultados dos testes para escolha da quantidade de neurônios ocultos	18
Tabela 2 – Resultados dos testes - Base de dados original	26
Tabela 3 – Resultados dos testes - Base de dados aumentada	29
Tabela 4 – Comparativo dos resultados - Bases de dados original e aumentada . .	30

Lista de Abreviaturas e Siglas

BSD	<i>Berkeley Software Distribution</i>
CCD	<i>Charge-coupled Device</i>
CNN	<i>Convolutional Neural Network</i>
HSV	<i>Hue, Saturation, Value</i>
MIT	<i>Massachusetts Institute of Technology</i>
MLP	<i>Multilayer Perceptron</i>
PIL	<i>Python Image Library</i>
ReLU	<i>Rectifier Linear Unit</i>
RGB	<i>Red, Green, Blue</i>
RNA	Rede Neural Artificial
SVM	<i>Support Vector Machine</i>

Sumário

1 – Justificativa	1
1.1 Introdução	1
1.2 Objetivos	2
1.2.1 Objetivos gerais	2
1.2.2 Objetivos específicos	2
1.3 Organização do trabalho	2
2 – Fundamentação Teórica	3
2.1 Aquisição de imagens digitais	3
2.2 Modelos de cores	4
2.3 Pré-processamento	5
2.4 Classificação e reconhecimento	7
2.5 Redes neurais artificiais	7
2.5.1 Neurônios artificiais	8
2.5.2 Perceptrons	9
2.5.3 Perceptrons multicamadas	9
2.5.4 <i>Deep Learning</i>	11
2.5.5 Redes neurais convolutivas	12
3 – Trabalhos Relacionados	14
4 – Metodologia	16
4.1 Ferramentas utilizadas	16
4.2 Base de dados	16
4.3 Testes	17
4.3.1 Testes realizados com redes Perceptron Multicamadas	18
4.3.1.1 Definição da estrutura da rede	18
4.3.1.2 Teste A	19
4.3.1.3 Teste B	19
4.3.1.4 Teste C	20
4.3.2 Teste realizado com rede convolutiva (Teste D)	21
4.4 Critérios de Avaliação	22
5 – Análise e Discussão dos Resultados	24
5.1 Testes com base de dados original	24
5.2 Testes com base de dados aumentada	27

5.3	Comparativo dos resultados	30
6	– Conclusão	32
6.1	Trabalhos Futuros	32
6.2	Considerações Finais	33
	Referências	34

1 Justificativa

O processo de classificação é de fundamental importância no cotidiano humano. De acordo com Zhang (2000), a classificação trata-se de uma das principais tarefas relacionadas à decisão realizadas pelos seres humanos e ocorre quando é necessário atribuir um objeto à um grupo, baseado em seus atributos. O desenvolvimento de algoritmos que sejam capazes de discernir entre diferentes objetos, ou de classificá-los em grupos, permite a automatização de muitas tarefas que atualmente são realizadas de forma manual. A classificação de flores é uma tarefa que, devido à sua natureza, torna-se muitas vezes difícil até mesmo para especialistas da área. Algumas espécies compartilham características visuais muito similares, resultando em um processo árduo e impreciso, além de demandar grande conhecimento por parte de um classificador humano. Segundo Guru, Sharath e Manjunath (2010), algumas das principais dificuldades encontradas para o desenvolvimento de um sistema funcional para classificação de flores, se deve tanto às similaridades visuais apresentadas entre as classes, quanto a variação intra-classe, além de outros fatores, tais como variações na iluminação e tamanho das imagens obtidas. Para classificação e detecção de objetos em imagens, abordagens geralmente adotadas, baseiam-se na utilização de descritores de imagem em conjunto à técnica de *Bag-of-visual-words* e algoritmos de classificação, tal como *Support vector machines* (SVMs) (SRINIVAS et al., 2016)(REHMAN et al., 2013)(YANG et al., 2009). A principal motivação para utilização das redes neurais no processo de classificação, reside no fato de que além de serem capazes de aprender a partir dos dados de entrada, algumas possuem a capacidade de identificar e extrair características que definem a imagem, dispensando portanto, processamentos prévios (Seção 2.3).

1.1 Introdução

Segundo Nilsback e Zisserman (2006), em geral, as flores diferenciam-se umas das outras pela combinação dos atributos de forma, cor ou textura. A família *Orchidaceae* é uma das maiores famílias de planta existentes na terra, contendo aproximadamente 25.000 espécies naturais, com novas espécies que continuam surgindo (GRAVENDEEL et al., 2004)(DRESSLER, 2005) e de acordo Janes (2006), dada sua diversidade e estrutura, estas flores têm causado fascínio.

Pertencente a esta mesma família, encontram-se as orquídeas do gênero *Cattleya*, que de acordo com Sorace et al. (2009), contempla cerca de 60 espécies e são encontradas em regiões que vão desde o México à América do Sul. Devido à similaridade de características visuais compartilhadas pelas orquídeas deste gênero, a classificação correta de suas espécies nem sempre é uma tarefa trivial.

1.2 Objetivos

1.2.1 Objetivos gerais

O objetivo deste trabalho é realizar a tarefa de classificação de um conjunto de imagens digitais das espécies de orquídea *Intermedia*, *Purpurata* e *Walkeriana*, pertencentes ao gênero *Cattleya*, através da utilização de redes neurais artificiais. Para realização deste propósito, são utilizadas duas categorias deste tipo de rede: redes perceptron com múltiplas camadas e redes convolutivas. Ao final do processo de classificação, o desempenho dos modelos propostos é analisado através da comparação dos resultados obtidos, discutindo-se a viabilidade da aplicação destes tipos de rede no problema proposto.

1.2.2 Objetivos específicos

Os objetivos específicos deste trabalho são:

- Implementar uma rede neural do tipo perceptron com múltiplas camadas e uma rede neural convolutiva;
- Criar uma base de dados contendo imagens das espécies de orquídea *Intermedia*, *Purpurata* e *Walkeriana*;
- Preparar as imagens através dos processamentos necessários para carregamento às redes;
- Inserir nas redes neurais implementadas, os dados das bases de treino;
- Analisar o desempenho e qualidade dos classificadores implementados, através dos resultados obtidos após o processo de classificação.

1.3 Organização do trabalho

A organização do trabalho é definida através da seguinte estrutura: O Capítulo 2 descreve os fundamentos de processamento de imagens e das redes neurais artificiais. No Capítulo 3 são citados trabalhos de outros autores que buscaram solucionar problemas similares. Uma descrição da metodologia aplicada é apresentada no Capítulo 4. O Capítulo 5 realiza uma discussão acerca dos resultados obtidos e por fim, o trabalho é concluído no Capítulo 6, onde são realizadas análises da aplicabilidade das soluções propostas.

2 Fundamentação Teórica

2.1 Aquisição de imagens digitais

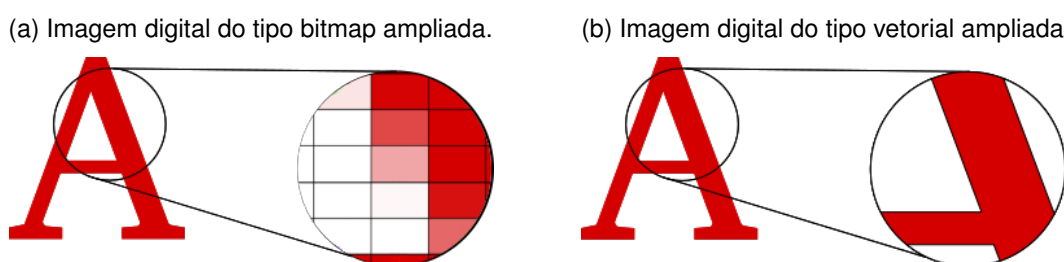
O processo de aquisição consiste na captura de imagens do objeto que se deseja estudar. As câmeras digitais utilizam para este processo uma estrutura bidimensional formada por um conjunto de sensores, conhecida como *charge-coupled device* (CCD). Estes dispositivos são sensíveis à luz e possuem a capacidade de captar a energia luminosa incidente sobre eles (GONZALES; WOODS, 2009).

De acordo com Ignacio, Buso e Monteiro (2013), Sakshica e Gupta (2015) e Makarewicz, Korga e Rosa (2015), as imagens digitais podem ser divididas em dois grupos: Imagens tipo Bitmap (*raster*) e vetoriais, sendo suas principais características descritas a seguir:

As imagens de tipo *raster* são representadas como uma matriz bidimensional, composta por valores numéricos conhecidos como *pixels* (*Picture Element*). Cada *pixel* contém valores que representam o atributo cor e formam a menor unidade da imagem. Devido à sua característica matricial, este tipo de imagem perde qualidade ao ser ampliada.

Já as imagens vetoriais, são compostas por formas geométricas descritas através de equações matemáticas. Tal característica permite que estas imagens sejam redimensionadas sem quaisquer perda de qualidade. Devido à sua flexibilidade, as imagens vetoriais geralmente são utilizadas para criação de representações que devem ser disponibilizadas em diferentes tamanhos, tais como logotipos ou símbolos. A Figura 1, compara o resultado da ampliação das imagens do tipo *raster* e vetorial. É possível notar na Figura 1a a degradação da qualidade da imagem *raster* após a ampliação, devido a evidência de seus *pixels*. Já a imagem vetorial, apresentada na Figura 1b, não sofre perdas. Embora as imagens vetoriais não se degradem ao serem ampliadas, só podem ser criadas de forma computacional e não podem ser geradas por sensores de câmeras.

Figura 1 – Comparativo do processo de ampliação de uma imagem tipo *raster* e vetorial



Fonte: Elaborado pelo autor

2.2 Modelos de cores

As imagens digitais podem ser representadas através de diferentes modelos de cores. Estes modelos têm como objetivo definir padrões que permitam representar as diferentes cores existentes (GONZALES; WOODS, 2009).

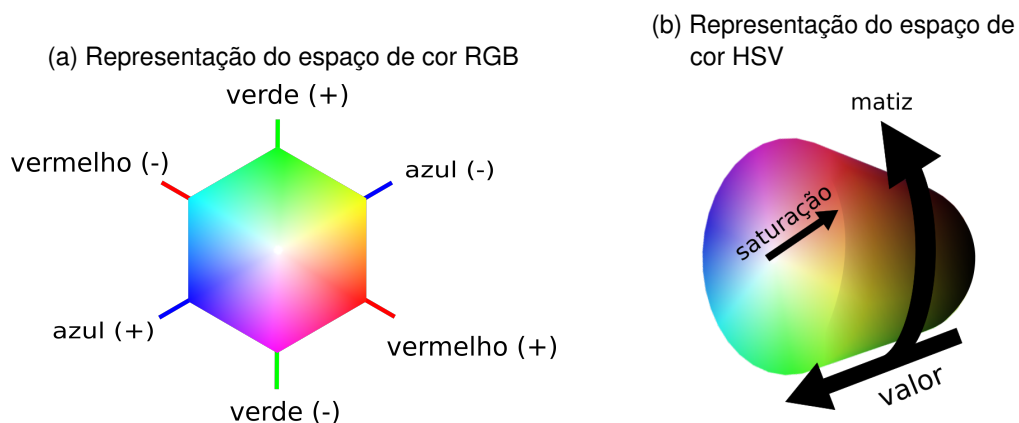
Dois modelos comumente utilizados em tarefas relacionadas à visão computacional são o *Red, Green, Blue* (**RGB**) e o *Hue, Saturation, Value* (**HSV**).

No modelo RGB, as cores presentes na imagem são compostas pelos componentes de cor vermelho(*red*), verde(*green*) e azul(*blue*). Tratando-se de um modelo aditivo, suas cores são formadas através da adição de diferentes intensidades destes componentes (MESKALDJI; BOUCHERKHA; CHIKHI, 2009). As imagens representadas no formato RGB podem ser descritas como uma matriz tridimensional de dimensões $m \times n \times 3$, onde m e n são as dimensões da imagem, e 3 são os canais de cor vermelho, verde e azul. Em geral estes valores de cor são contidos no intervalo $[0,255]$;

No modelo de cores HSV, cada cor é descrita através dos componentes matiz(*hue*), saturação(*saturation*) e valor(*value*). O canal matiz indica a cor a ser representada. O canal saturação define o nível de intensidade ou pureza desta cor e por último, o componente valor, que especifica o nível de brilho presente na cor (PANDEY; MISHRA; SHARMA, 2011), (MYTHILI; V.KAVITHA, 2012), (GEORGIEVA; DIMITROVA; ANGELOV, 2005).

A Figura 2 apresenta um comparativo da representação de cor em cada um dos modelos citados, onde o modelo RGB (Figura 2a) é representado por um cubo e o modelo HSV (Figura 2b), por um cilindro.

Figura 2 – Representação de cor nos modelos RGB e HSV

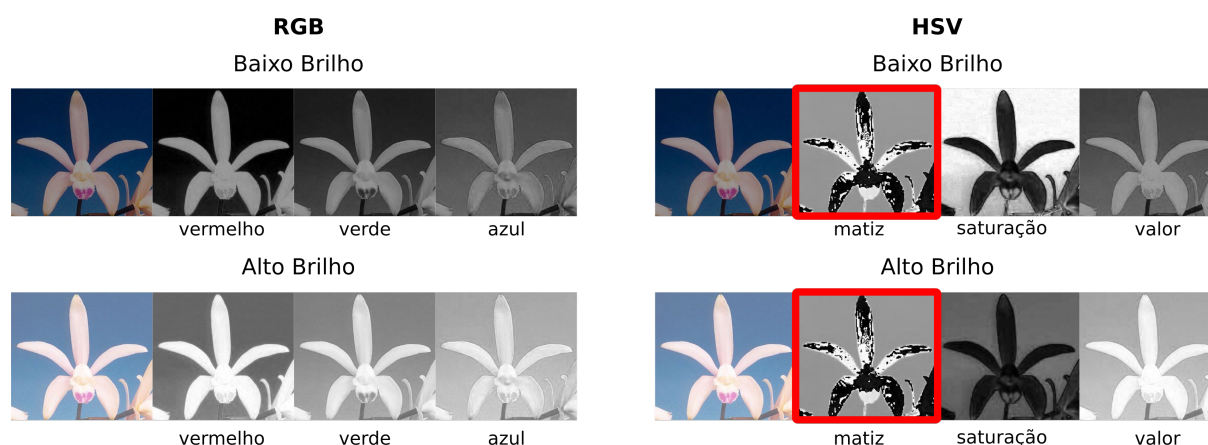


Fonte: Elaborado pelo autor

Uma das características do modelo de cores HSV, é que seu canal matiz(*hue*) é invariante à mudanças na iluminação (SERGYÁN, 2007). Esta característica torna a utilização deste formato propício para tarefas relativas à visão computacional onde a cor

possui papel importante. Através da Figura 3, é possível ver os diferentes canais resultantes da decomposição de uma imagem em ambos os formatos RGB e HSV, em alto e baixo níveis de brilho. Pode-se perceber, através da imagem, que o canal matiz do formato HSV (destacado em vermelho), é o único que não sofre alterações em mudanças de iluminação.

Figura 3 – Comparativo da decomposição de uma imagem nos modelos HSV e RGB



Fonte: Elaborado pelo autor

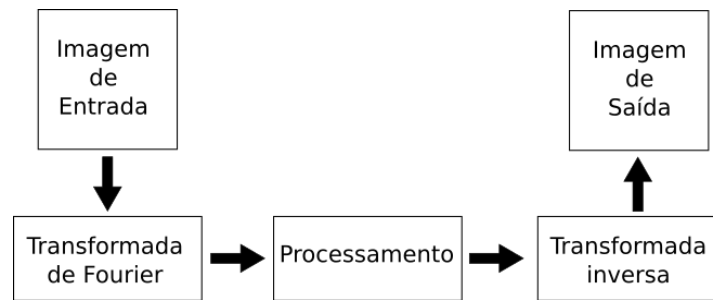
2.3 Pré-processamento

A principal tarefa do pré-processamento de imagens, é acentuar suas características ou corrigir eventuais imperfeições que nelas possam existir, com o objetivo de melhorá-las para processamentos posteriores (MILJKOVIĆ, 2009),(FARAHMAND; SARRAFZADEH; SHANBEHZADEH, 2013). A escolha das técnicas empregadas depende diretamente do tipo de problema evidenciado. Para remoção de ruídos e aguçamento de bordas, algumas abordagens comuns baseiam-se nas filtragens no domínio da frequência e no domínio do espaço. Segundo Gonzales e Woods (2009), tais abordagens podem ser descritas da seguinte forma:

Para realização da filtragem de imagens no domínio da frequência, em geral utiliza-se um método conhecido como **Transformada de Fourier**, que tem por objetivo representar um sinal através da soma de senos e cossenos. Para realização deste tipo de filtragem são necessárias as seguintes etapas: Conversão inicial da imagem original (domínio do espaço) em sua representação no domínio da frequência. Alteração da imagem resultante através da remoção ou realce das frequências desejadas e por fim, conversão do resultado novamente ao domínio espacial através da transformada inversa. A Figura 4, ilustra o processo de aplicação da transformada nas imagens.

Diversos filtros podem ser aplicados às imagens representadas no domínio da frequência, dentre eles pode-se citar o filtro passa-alta, que tem por objetivo aguçar e

Figura 4 – Etapas da filtragem no domínio da frequência

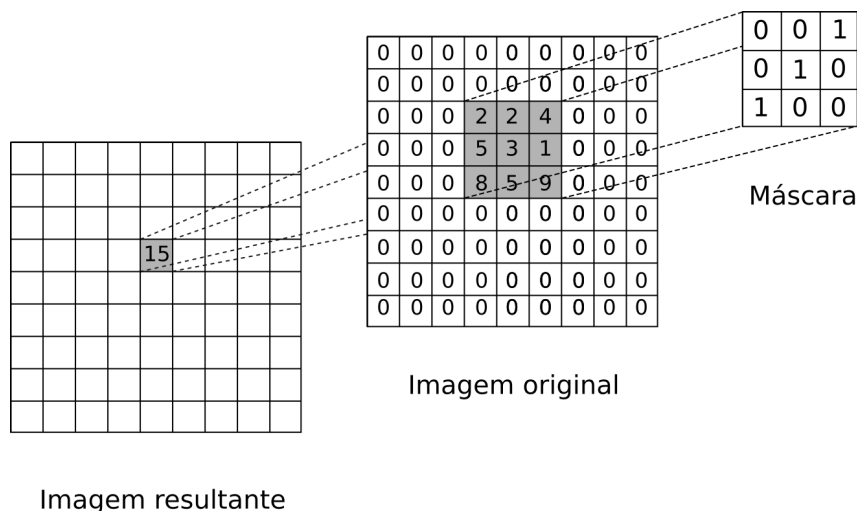


Fonte: Adaptado de (GONZALES; WOODS, 2009)

realçar altas frequências da imagem e o filtro passa-baixa, que retira as altas frequências do espectro, resultando em uma imagem contendo menos ruídos.

Outra forma de aplicar filtros às imagens digitais é através de sua aplicação no domínio espacial. Embora utilize os princípios da filtragem no domínio da frequência, este tipo de filtragem age diretamente sobre os *pixels* que compõe a imagem, sem a necessidade de conversão. Para tal, neste processo, uma máscara (*kernel*) de tamanho $M \times N$ percorre a imagem original, multiplicando o valor dos *pixels* desta imagem pelo coeficiente contido na posição correspondente da máscara. O valor resultante da soma desta operação é atribuído a uma nova imagem, na posição central da máscara. Este processo é conhecido como convolução. A Figura 5 ilustra o funcionamento da aplicação de uma máscara de dimensões 3×3 .

Figura 5 – Exemplo do processo de convolução



Fonte: Elaborado pelo autor

Um exemplo de filtro comumente utilizado para aguçamento de bordas nas imagens, trata-se do filtro Laplaciano. Este filtro é representado através da matriz da Figura 6.

O resultado de sua aplicação em uma imagem é mostrado na Figura 7.

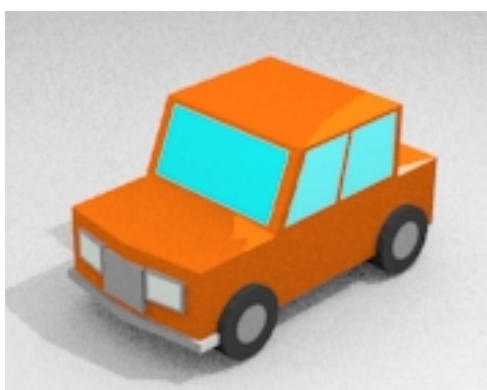
Figura 6 – Filtro Laplaciano

-1	-1	-1
-1	8	-1
-1	-1	-1

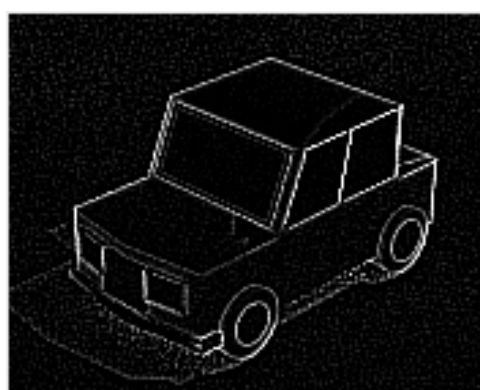
Fonte: (GONZALES; WOODS, 2009)

Figura 7 – Resultado obtido após a aplicação do filtro Laplaciano

(a) Imagem Original



(b) Resultado



Fonte: Elaborado pelo autor

2.4 Classificação e reconhecimento

O processo de classificação de imagens consiste em rotulá-las em alguma categoria predefinida, baseando-se nas características dos objetos que ela possui, ou seja, categorizar os objetos nela detectados em classes pré-definidas (HARZALLAH; JURIE; SCHMID, 2009)(KAMAVISDAR; SALUJA; AGRAWAL, 2013). Diversas técnicas têm sido estudadas para a tarefa de classificação de imagens, sendo que a escolha dos algoritmos e técnicas empregadas, depende muitas vezes do tipo de problema que se deseja resolver e dos recursos disponíveis.

2.5 Redes neurais artificiais

Embora não haja uma única definição formal (KUMAR; SHARMA, 2014), as redes neurais artificiais (RNAs) podem ser definidas como um tipo de rede inspirada no método como os sistemas nervosos biológicos processam informação. De acordo Panchal e Panchal (2014), estas redes podem ser utilizadas em vários tipos de aplicações, tais como reconhecimento de padrões, diagnósticos médicos, aplicações financeiras, mineração de

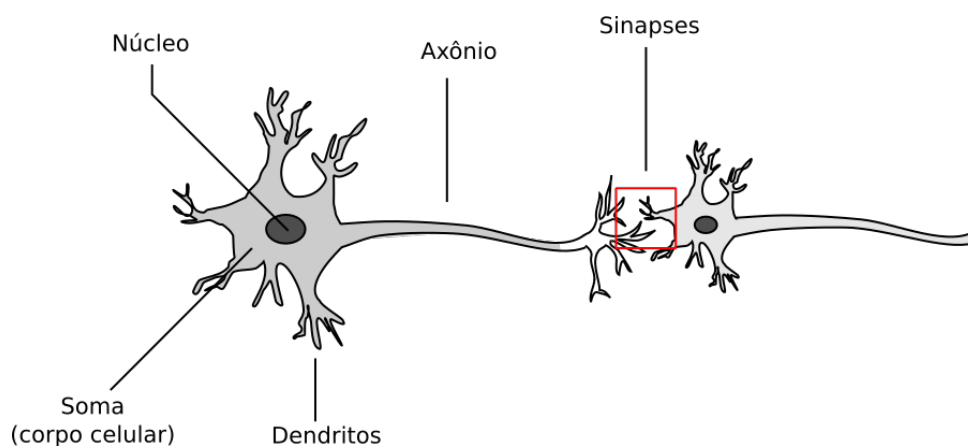
dados, reconhecimento de gestos, fala, escrita, dentre outros.

Segundo Coppin (2010), o cérebro humano é composto bilhões de células conhecidas como neurônios. Cada neurônio conecta-se a milhares de outros através de conexões conhecidas como sinapses, sendo que o cérebro contém dezenas de trilhões delas. O sistema nervoso humano forma-se através deste complexo conjunto de conexões. O neurônio humano pode ser descrito através dos seguintes elementos: os dendritos, um corpo celular e os axônios.

Os dendritos são responsáveis pela recepção das informações, o corpo celular armazena o núcleo da célula e os axônios realizam o transporte destas informações. Os sinais são transmitidos aos demais neurônios através da sinapse.

O modelo de um neurônio humano pode ser visto na Figura 8.

Figura 8 – Modelo de um neurônio humano



Fonte: Elaborado pelo autor

2.5.1 Neurônios artificiais

O modelo de rede neural artificial foi desenvolvido na década de 1940 pelo psiquiatra e neuroanatomista Warren S. McCulloch, juntamente ao lógico Walter Pitts (SHIFFMAN, 2012) e de acordo com Haykin (2001), estas redes se assemelham ao cérebro humano pelo fato de que o conhecimento adquirido dá-se através do processo de aprendizagem e é armazenado como forças de conexão entre seus neurônios, como valores dos pesos sinápticos. Haykin (2001) também descreve que um neurônio artificial é composto pelos seguintes elementos:

- Um conjunto de entradas, contendo pesos ou forças. Os valores passados à estas entradas são multiplicados pelos pesos;
- Um somador (junção aditiva), responsável por computar a soma das entradas recebidas por este neurônio;

- Uma função de ativação, responsável por limitar o valor de saída do neurônio, em geral no intervalo $[0,1]$ ou $[-1,1]$. Estas funções são utilizadas como uma função de transferência de valores entre neurônios (SHARMA; RAI; DEV, 2012). Algumas das funções mais comumente utilizadas são apresentadas na Figura 9.

À entrada de um neurônio artificial, pode-se ainda aplicar um valor adicional definido como *bias* (HAYKIN, 2001). O modelo computacional de um neurônio contendo seus diferentes componentes é apresentado na Figura 10.

As redes neurais são capazes também de realizar seu aprendizado através do conjunto de dados de entrada e saídas esperadas, tornando-as robustas e adaptáveis à problemas. De acordo com Maind e Wankar (2014), alguns dos principais benefícios deste tipo de rede são sua capacidade de aprender a realizar tarefas a partir de dados inseridos, além de criar sua própria representação da informação recebida.

2.5.2 Perceptrons

O perceptron trata-se da mais simples rede neural artificial, sendo utilizada para classificação de itens linearmente separáveis. Foi inicialmente proposto por Frank Rosenblatt ao final da década de 1950 e é composto por um conjunto de entradas, um processador e uma saída (SHIFFMAN, 2012). Os perceptrons utilizam como saída uma função degrau, na qual o retorno é $+1$ caso a soma de suas entradas for maior que um limiar predefinido, e -1 caso contrário.

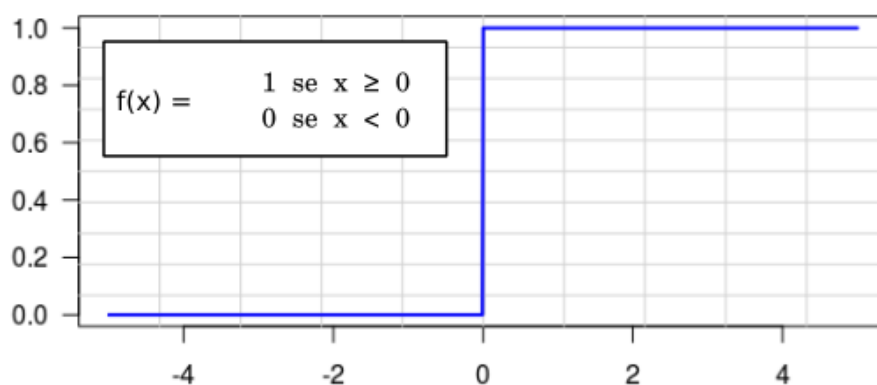
De acordo com Coppin (2010), o processo de aprendizagem do perceptron pode ser descrito a partir da seguinte forma: pesos inicialmente aleatórios são distribuídos à suas entradas; em seguida, um dado de treinamento é inserido e calculado. Caso a saída não seja a esperada, estes pesos são ajustados de acordo com uma taxa de aprendizado predefinida; os demais dados de treinamento são então inseridos e testados da mesma forma e caso haja erros, os pesos são novamente ajustados. Todo este processo é realizado de maneira contínua, até que não haja mais saídas incorretas para os os dados da base de treinamento. As iterações deste algoritmo são denominadas épocas.

2.5.3 Perceptrons multicamadas

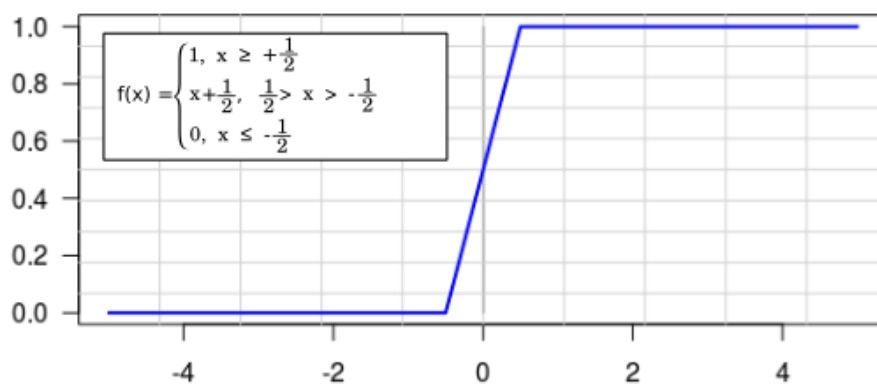
Os perceptrons de múltiplas camadas (*multilayer perceptron* - MLP) são redes capazes de computar problemas não linearmente separáveis. Estas redes podem ser consideradas como uma generalização dos perceptrons comuns e apresentam em sua estrutura uma camada de entrada, uma ou mais camadas ocultas e uma camada de saída. As camadas ocultas situam-se entre as camadas de entrada e saída. Neste modelo de rede, todos os nós são conectados diretamente àqueles da camada seguinte, de forma que os sinais seguem um fluxo da entrada à saída e por isso, estas redes são conhecidas como

Figura 9 – Funções de ativação comumente utilizadas

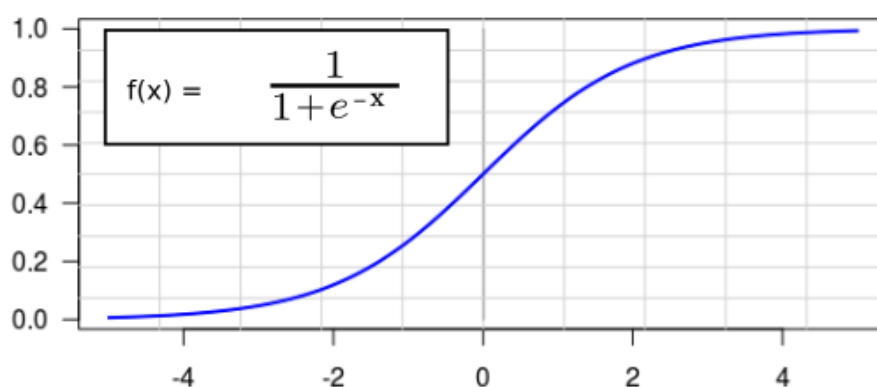
(a) Função limiar



(b) Função linear por partes



(c) Função sigmóide

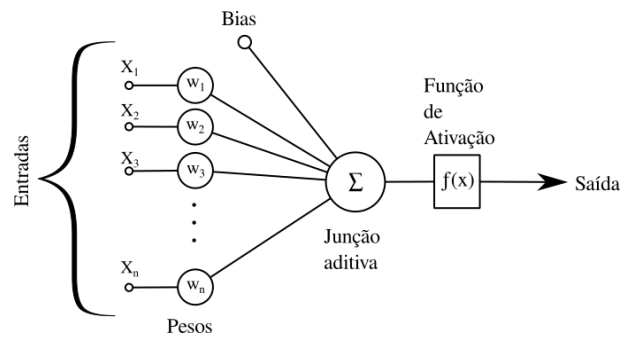


Fonte: Adaptado de (HAYKIN, 2001)

redes do tipo *feedforward*. Um exemplo de arquitetura das redes MLP pode ser visto através da Figura 11.

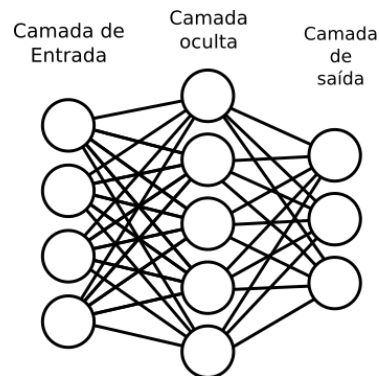
O aprendizado neste tipo de rede ocorre através do algoritmo iterativo retropro-

Figura 10 – Modelo de um neurônio artificial



Fonte: Adaptado de (HAYKIN, 2001)

Figura 11 – Modelo de uma rede do tipo perceptron com múltiplas camadas



Fonte: Adaptado de (SHIFFMAN, 2012)

pagação (*backpropagation*). Segundo Coppin (2010), neste processo de aprendizagem, pesos aleatórios são atribuídos aos nós de entrada da rede, assim como ocorre nas redes perceptron comuns. Os dados de treinamento são então carregados à rede e percorrem seu caminho até a saída. Caso a saída contenha erros, este erro é calculado, propagado às camadas anteriores e o ajuste nos pesos é realizado. Este processo se repete enquanto o erro da saída for maior que um limiar predefinido. O algoritmo possui um parâmetro de taxa de aprendizado que define a velocidade com a qual ela ocorre. Em valores maiores, a aprendizagem ocorre de maneira mais rápida, já com valores menores, o processo é mais preciso. Outro parâmetro que pode ser incluído no processo de treinamento é o *momentum*, que de acordo com o autor, objetiva melhorar o desempenho do algoritmo de retropropagação ao evitar, também, que o mesmo caia em mínimos locais.

2.5.4 Deep Learning

De acordo com Deng (2014), a arquitetura de *deep learning* (aprendizado profundo) consiste em conjunto de técnicas de aprendizado de máquina, focados na utilização de diversas camadas de processamento de informação, sendo aplicado com sucesso em

diversas áreas, tais como visão computacional e reconhecimento de objetos.

Estas arquiteturas utilizam redes com diversas camadas intermediárias (TUSHAR, 2015), sendo que através delas, segundo Wu et al. (2015), tornou-se então possível reconhecer objetos visuais sem a necessidade da extração prévia de características.

Para Lecun, Bengio e Hinton (2015) as redes de aprendizado profundo baseiam-se na propriedade hierárquica dos sinais, sendo a **rede neural convolutiva** um tipo de rede de aprendizado profundo que obteve sucesso por ser fácil de treinar e possuir melhor generalização se comparada às redes totalmente conectadas.

2.5.5 Redes neurais convolutivas

As redes neurais convolutivas (*Convolutional Neural Networks* - CNN) podem ser definidas como variantes das redes MLP e têm sua arquitetura baseada na maneira como a percepção visual ocorre biologicamente. Suas principais características são a capacidade de reconhecer formas bidimensionais invariantes em termos de distorção (HAYKIN, 2001) e a capacidade de detectar padrões visuais diretamente dos *pixels* de entrada, dispensando, muitas vezes, a etapa de pré-processamento (GU et al., 2015). De acordo com Shang et al. (2016), estas redes tem sido recentemente utilizadas na solução de problemas relacionados à visão computacional.

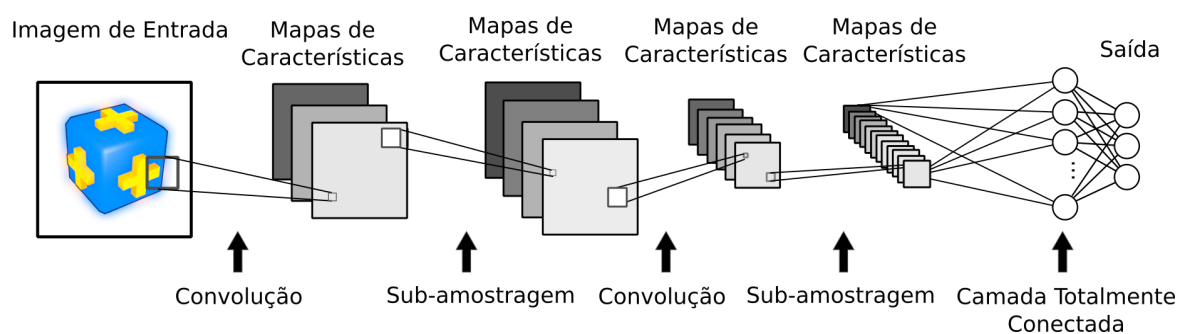
As camadas desta rede são compostas de um conjunto de **mapas de características** que possuem a forma de um plano. Estes mapas são extraídos das imagens através do processo de convolução. Além disso, os neurônios de cada camada recebem sua entrada através de um campo receptivo local da camada anterior, permitindo-os extrair características (HAYKIN, 2001)(BENGIO; LECUN; LECUN, 1995).

A arquitetura das redes CNN pode ser descrita através de uma pilha de diferentes camadas em sequência, cada qual com um objetivo específico. Segundo O'Shea e Nash (2015) a arquitetura de uma rede neural convolutiva geralmente baseia-se em três tipos de camada:

- *Convolutional layers*: Realiza o processo de convolução. Cada neurônio contido nesta camada recebe a entrada de uma seção retangular (campo receptivo) da camada anterior.
- *Pooling layers*: Realizam o processo de redução da dimensionalidade espacial dos dados de entrada (sub-amostragem), reduzindo portanto, a complexidade do modelo construído. Em geral, esta camada recebe como entrada um grupo de valores e retorna o máximo dentre eles;
- *Fully-connected layers*: Geralmente tratam-se das camadas finais da rede. Seus neurônios são conectados à todos os da camada anterior, sendo responsáveis por computar o resultado da classificação, assim como ocorre nas redes neurais comuns.

Este tipo de rede, em geral, utiliza entre suas camadas uma função de ativação conhecida como *Rectifier Lienar Unit* (ReLU). Esta função pode ser representada através da fórmula $\max(0,x)$ e, de acordo com Goldberg (2015), por se tratar de uma função extremamente simples, apresenta boa performance e resultados. A Figura 12 apresenta um modelo comum de rede neural convolutiva, ilustrando suas diferentes camadas.

Figura 12 – Modelo de uma rede do tipo convolutiva



Fonte: Adaptado de (LECUN et al., 1998). [1998] IEEE.

3 Trabalhos Relacionados

Este capítulo apresenta trabalhos da literatura, no qual seus autores buscaram aplicar redes neurais no processo de classificação de imagens.

Em Anumula et al. (2015) os autores utilizaram uma rede MLP para reconhecimento de 6 diferentes placas de trânsito. A estrutura da rede construída continha 256 nós de entrada, 16 nós na camada oculta e 6 nós na camada de saída. O *dataset* foi formado por 500 imagens de 16×16 *pixels* de dimensão e o pré-processamento consistiu na remoção de ruídos das imagens através da aplicação de um filtro Gaussiano e binarização da imagem. O algoritmo foi treinado por 1.000 épocas, obtendo uma acurácia de 99%.

Lee et al. (2015) utilizaram redes neurais convolutivas no processo de classificação de 44 espécies de plantas através de um *dataset* contendo 2.288 imagens de folhagem para etapa de treino e utilizando 528 imagens para etapa de testes. De acordo com os autores, foi possível obter uma acurácia de 99,6%.

Os autores Andrade, Maciel e Romero (2016), apresentaram um método para classificação de placas de automóveis utilizando redes MLP e técnicas de processamento de imagens. O processamento prévio das imagens utilizadas consistiu na transformação das imagens RGB originais em escala de cinza e aplicação do processo de binarização. Posteriormente, a imagem foi segmentada, e os caracteres nela contidos foram extraídos. A base de treinamento foi composta de 1.116 amostras de caracteres de dimensões 10×15 *pixels*. Utilizando uma estrutura de rede composta de duas camadas ocultas com 100 e 40 neurônios respectivamente, os autores puderam alcançar um nível de detecção de 100% nos caracteres contidos em uma base de testes contendo 83 placas de carro.

Em Sharma e Gupta (2015) é proposto uma técnica para classificação de folhas baseado na utilização de redes MLP. As imagens destas folhas foram inicialmente redimensionada às dimensões 256×256 *pixels* e convertidas à escala de cinza. Posteriormente, o contorno de cada uma foi identificado através da aplicação de um filtro de convolução e 12 características morfológicas foram extraídas, dentre elas: perímetro, comprimento dos eixos, área e ângulo de orientação. Os dados extraídos foram alimentados a uma rede neural composta de uma entrada com 12 neurônios, duas camadas ocultas com 90 neurônios cada e um único neurônio de camada de saída. A rede foi treinada por 10.000 épocas. Segundo os autores, de um total de 440 folhas de 16 tipos de plantas, 401 foram corretamente classificadas, resultando em uma acurácia de 91,13%.

Levi e Hassner (2015) utilizaram redes neurais convolutivas para estimação da idade e gênero através de um conjunto de imagens de faces de pessoas. A rede utilizada foi formada por três camadas convolutivas, seguidas da operação ReLU e camadas de *pooling*,

duas camadas totalmente conectadas ao fim da rede, compostas de 512 neurônios cada, e uma camada de saída responsável por mapear as classes de idade e gênero. Com o objetivo de prevenir o *overfitting*, os autores optaram por utilizar a técnica de *dropout* (SRIVASTAVA et al., 2014), realizando testes com a base de dados original e uma versão aumentada através do espelhamento das imagens utilizadas. Nos testes realizados, os autores obtiveram uma taxa de acurácia de aproximadamente 85% na detecção de gênero e 50% na detecção da faixa etária exata das pessoas nas imagens, em ambos os testes realizados. Os autores do trabalho concluem que este tipo de rede pode ser utilizada na tarefa de classificação de idade e gênero, mesmo em base de dados pequenas, e que a simplicidade do modelo por eles construído, implica que melhores resultados podem ser obtidos através da utilização de mais dados de treinamento.

Em Saidane e Garcia (2007), foi proposto um método para reconhecimento de caracteres extraídos de imagens. O método consiste na utilização de uma rede neural convolutiva com 7 camadas, sendo: uma camada de entrada, duas camadas de convolução e subamostragem intercaladas e duas camadas totalmente conectadas. A rede foi treinada com um conjunto original de 11.615 imagens, que foi expandido através da adição de imagens negativas e ruidosas. De acordo com os autores, o sistema proposto foi capaz de obter uma taxa média de reconhecimento de 84,53%.

Através dos trabalhos apresentados, percebe-se que as redes neurais, em geral, podem ser empregadas na solução de diversos problemas. Em muitos trabalhos envolvendo RNAs, os autores utilizam como abordagem a inserção direta dos *pixels* que compõem a imagem, às redes do tipo MLP. Embora esta solução possa funcionar em imagens pouco complexas, pode apresentar resultados pouco satisfatórios em imagens de grandes dimensões, além de tornar complexa a estrutura da rede neural. Outros trabalhos, porém, aplicam etapas prévias de processamento nas imagens, com o objetivo de se extrair informações que serão carregadas às redes. Apesar da utilização da etapa de pré-processamento permitir a drástica redução da complexidade da rede neural e como consequência resultar em possíveis melhoras no modelo construído, esta etapa resulta em um trabalho muitas vezes árduo e demorado. A vantagem da utilização de redes neurais convolutivas no processo de classificação, reside no fato de a mesma ser capaz de obter bons resultados, com pouco ou nenhum pré-processamento em suas imagens de entrada.

4 Metodologia

O presente capítulo apresenta as ferramentas, métodos e etapas utilizados para realização do processo de classificação proposto neste trabalho.

4.1 Ferramentas utilizadas

As ferramentas utilizadas neste trabalho para construção das redes neurais e tarefas de manipulação de imagens, são apresentadas a seguir:

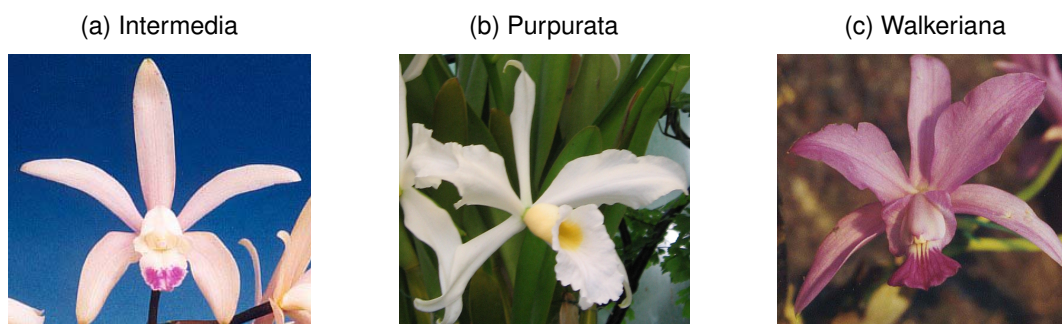
- **PyBrain**: De acordo com Schaul et al. (2010), o PyBrain (*Python-Based Reinforcement Learning*) é uma biblioteca voltada ao aprendizado de máquina, desenvolvida na linguagem Python que objetiva ser flexível e de fácil utilização. A biblioteca é disponibilizada sob a licença BSD e pode ser obtida através do endereço <http://pybrain.org/>.
- **Keras**: Segundo Chollet (2015), Keras é uma biblioteca voltada para implementação de redes neurais, em especial aos algoritmos de *Deep Learning*. Foi desenvolvida também na linguagem Python e permite a implementação de redes neurais convolutivas de forma simplificada e eficiente. Pode ser obtida em <https://keras.io>, sob a licença MIT.
- **OpenCV**: O OpenCV (Open Source Computer Vision Library), é uma biblioteca especializada em aprendizagem de máquina e visão computacional. Conta com um grande conjunto de algoritmos para tarefas que vão desde manipulação de imagens a identificação de objetos e apresenta interfaces para linguagens de programação C++, C, Python, Java e MATLAB. Esta biblioteca é distribuída sob a licença BSD.

4.2 Base de dados

Para realização do processo de classificação proposto neste trabalho, foi criada uma base de dados (*dataset*) composta de imagens recolhidas da internet, das espécies de orquídeas *Intermedia*, *Purpurata* e *Walkeriana*. Estas orquídeas pertencem ao gênero *Cattleya* e apresentam entre si semelhanças visuais, fator importante para teste dos classificadores.

A Figura 13, apresenta fotografias de exemplo de cada uma das espécies de flores contidas no *dataset*.

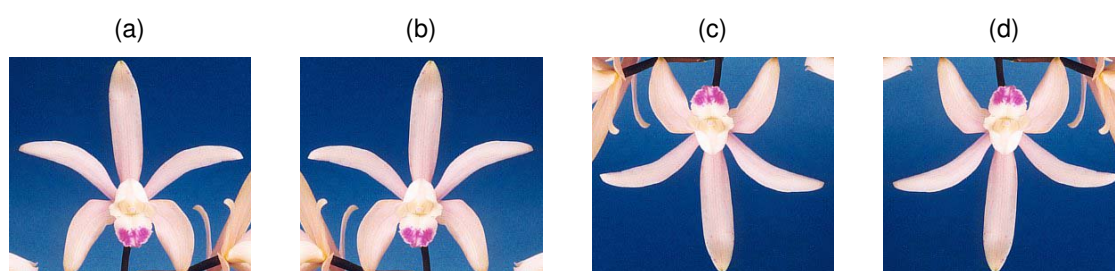
A **Base Original** foi formada por 360 imagens (120 imagens de cada uma das espécies de orquídea). Todas as imagens obtidas foram de flores únicas capturadas frontalmente e a base foi dividida em dois subgrupos, 70% para ser utilizada na etapa de treinamento da rede e 30% para ser utilizada na etapa de testes. Estas imagens foram então redimensiona-

Figura 13 – Imagem de três espécies pertencentes ao gênero *Cattleya*

Fonte: Larsen Twins Orchids - Wikimedia Commons, Domínio Público

das à resolução de 40×40 pixels. Esta resolução foi escolhida por tratar-se de um tamanho relativamente pequeno para ser carregado às redes, mas que ainda sim, permite sua identificação visual. Uma base auxiliar denominada **Base Aumentada** foi criada através da inclusão de versões invertidas verticalmente e espelhadas destas imagens, quadruplicando o tamanho da base, que passou a contar com 1.440 imagens (480 para cada classe). Um exemplo do processo de espelhamento aplicada à base de dados, pode ser visto na Figura 14. Esta base auxiliar foi criada com o objetivo de se analisar o impacto que a quantidade de elementos de treino possui na eficiência do classificador, além de permitir certo grau de variância a rotações, evitando que a rede neural aprenda a identificar as flores somente em uma posição específica. As bases são totalmente balanceadas, de forma que há uma quantidade igual de imagens de cada uma das classes de flores analisadas.

Figura 14 – Exemplo do espelhamento aplicado à base de dados



Fonte: Elaborado pelo autor

4.3 Testes

Um total de quatro testes foram realizados com o objetivo de se analisar o desempenho das redes neurais implementadas. Três deles foram realizados utilizando redes MLP, (Testes **A**, **B** e **C**), descritos na seção 4.3.1 e um teste adicional (**Teste D**), foi realizado utilizando uma rede neural convolutiva e é descrito em detalhes na seção 4.3.2. Em todos

os testes, o critério de parada para o treinamento foi o limite de 30 épocas. A configuração da máquina utilizada para testes tratou-se de um processador Intel Core i3 1.50GHz e 4GB de memória RAM. Estes testes foram realizados em duas instâncias. A primeira, utilizando a base de treino original (contendo 252 imagens na base de treino e 108 imagens na base de testes) e a segunda, utilizando a base de treinamento aumentada (contendo 1.008 imagens na base de treinamento e 432 imagens na base de testes).

4.3.1 Testes realizados com redes Perceptron Multicamadas

4.3.1.1 Definição da estrutura da rede

Um dos fatores importantes para o sucesso da utilização de redes MLP é a definição correta de sua estrutura de nós e camadas ocultas. Embora não haja um método único para se definir a melhor estrutura para este tipo de rede (KAVZOGLU, 1999), sabe-se que a utilização de uma camada oculta é suficiente para solução de quase todos os tipos de problemas, sendo que a utilização de duas ou mais camadas raramente melhora o modelo. No que tange a escolha da quantidade de neurônios nestas camadas, uma escolha incorreta pode ocasionar dois principais problemas: *underfitting*, que ocorre quando a quantidade de neurônios nas camadas ocultas é pequeno para detectar os sinais em uma base de dados de entrada complexa, e o *overfitting*, que ocorre quando a rede neural possui muita capacidade de processamento, mas a quantidade limitada de dados de treinamento não é o bastante para treiná-la, tornando-a pouco generalista, ou, caso haja uma grande quantidade de dados de treinamento, torna o tempo requerido para aprendizado da rede impraticável (PANCHAL et al., 2011), (PANCHAL; PANCHAL, 2014).

Com o objetivo de definir a quantidade de neurônios ocultos a serem utilizados pelas redes MLP utilizadas neste trabalho, um teste prévio de classificação foi realizado com 20,50,100,200 e 300 nós na camada oculta, utilizando a base de testes aumentada com as imagens em formato RGB. A tabela 1, apresenta os resultados obtidos.

Tabela 1 – Resultados dos testes para escolha da quantidade de neurônios ocultos

Neurônios	Acurácia	Tempo (segundos)
50	0,75	357,53
100	0,74	617,18
20	0,72	181,51
200	0,69	1155,11
300	0,55	1670,57

Fonte: Elaborado pelo autor

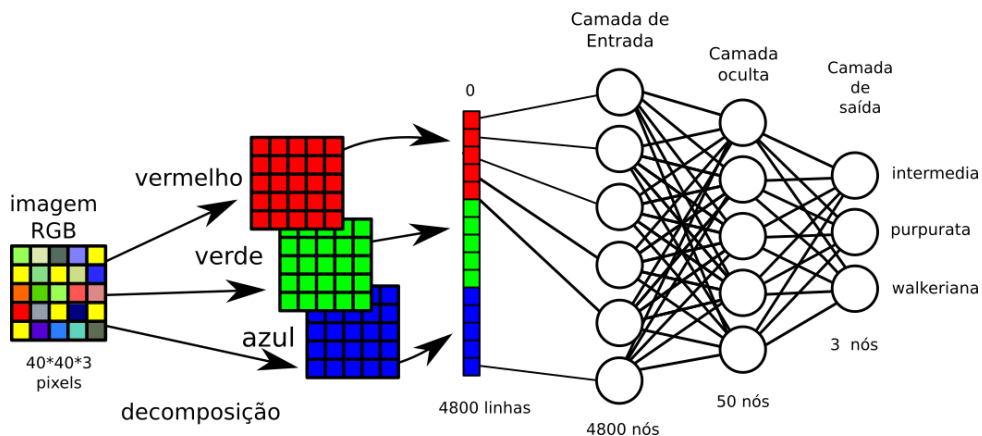
Através dos resultados, foi possível notar que o aumento da quantidade de nós ocultos ocasionou, na maioria das vezes, uma menor acurácia, além de resultar em um

tempo de treinamento maior. Visto que no presente trabalho a acurácia é o fator mais importante, foi constatado que a rede MLP com 50 nós apresentou a melhor alternativa. Portanto, em todas as redes MLP utilizadas neste trabalho, este foi o valor da quantidade de nós ocultos utilizado.

4.3.1.2 Teste A

Para realização deste teste, cada imagem presente no *dataset*, teve seus canais de cor vermelho, verde e azul separados e agrupados em um vetor linha. Os vetores contendo os dados de cada uma das imagens da base de testes foi normalizado entre o intervalo [0,1] e carregados à rede neural. Os parâmetros de treinamento utilizados foram: Taxa de aprendizado 0,005 e *momentum* de 0,90. A estrutura da rede foi composta de $40^2 \times 3$ nós na camada de entrada, uma camada oculta com 50 nós e uma camada de saída com 3 nós. A camada oculta utiliza uma função de ativação Sigmoide e a camada de saída utiliza uma função *Softmax*. A função *Softmax* (STANFORD VISION LAB, 2015) é geralmente utilizada em problemas de classificação onde as classes são mutualmente exclusivas e deseja-se obter um grau de confiança de cada classe predita. Um modelo descritivo do processo utilizado para realização deste teste pode ser visto na Figura 15.

Figura 15 – Modelo de inserção de uma imagem no formato RGB em uma rede MLP



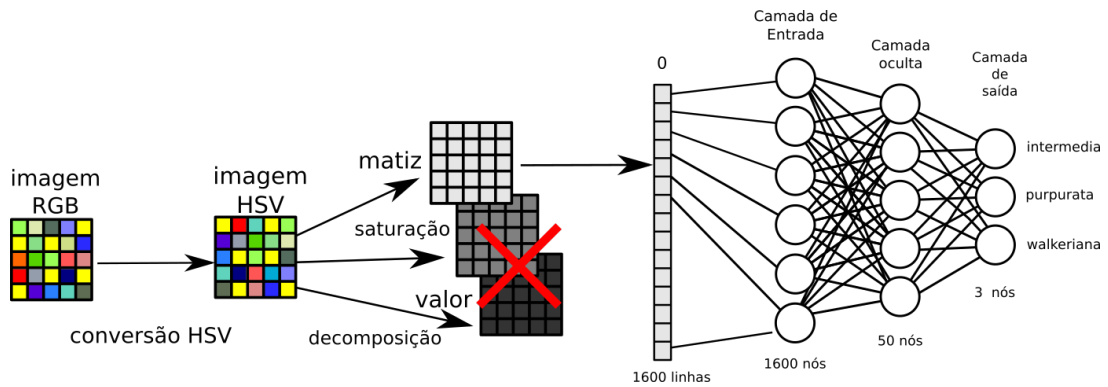
Fonte: Elaborado pelo autor

4.3.1.3 Teste B

O Teste B consistiu na conversão prévia das imagens ao formato HSV, antes de carregá-las à rede neural. Após conversão, somente o componente matiz foi carregado à rede, sendo os demais (saturação e valor) descartados. Através deste teste buscou-se analisar como a rede se comportaria utilizando-se somente a cor pura como parâmetro de classificação, dada a menor sensibilidade do canal matiz à mudanças de iluminação. A estrutura de rede e parâmetros de treinamento e teste foram iguais os utilizados no **Teste A**. A diferença reside no fato de que o número de nós de entrada foi reduzido à 40^2 , uma

vez que somente o canal matiz das imagens convertidas em HSV foi utilizado. Um exemplo do modelo utilizado para realização deste teste é apresentado na Figura 16.

Figura 16 – Modelo de inserção do canal matiz de uma imagem no formato HSV em uma rede MLP

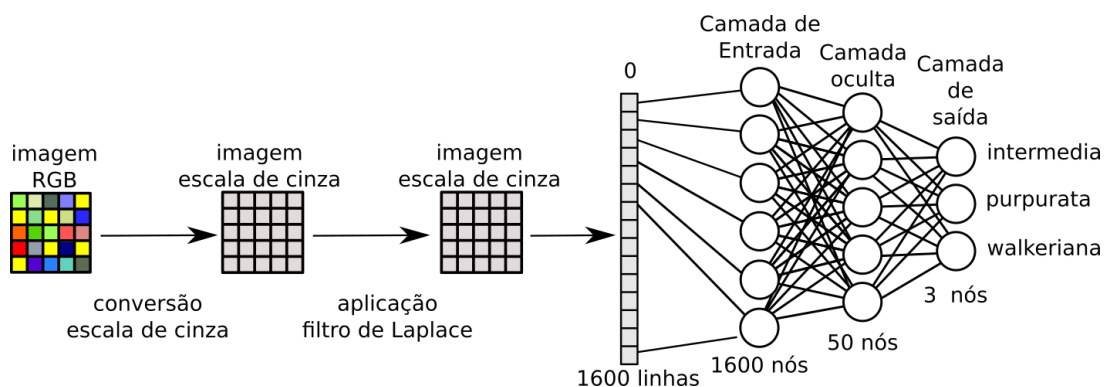


Fonte: Elaborado pelo autor

4.3.1.4 Teste C

Este teste consistiu aplicação prévia do filtro laplaciano às imagens RGB originais. Para tal, as imagens foram inicialmente convertidas em escala de cinza com 256 tons e suas bordas foram detectadas utilizando-se este filtro. Após a realização do processo, as mesmas foram carregadas à rede neural. Este teste teve como objetivo analisar os resultados obtidos com a utilização de imagens contendo somente as bordas das flores. Os parâmetros de treinamento e teste foram os mesmos utilizados no **Teste B**, novamente, com 40^2 nós de entrada, devido à conversão das imagens à escala de cinza e consequente perda dos canais de cor. Seu modelo é apresentado no diagrama da Figura 17.

Figura 17 – Modelo de inserção de uma imagem em escala de cinza com filtro de Laplace em uma rede MLP



Fonte: Elaborado pelo autor

4.3.2 Teste realizado com rede convolutiva (Teste D)

Este teste, denominado **Teste D**, consistiu no carregamento das imagens RGB a uma rede neural convolutiva. Segundo O'Shea e Nash (2015), assim como ocorre nas redes MLP, não há uma arquitetura específica para este tipo de rede. Portanto, no presente trabalho, a arquitetura definida foi baseada em testes empíricos.

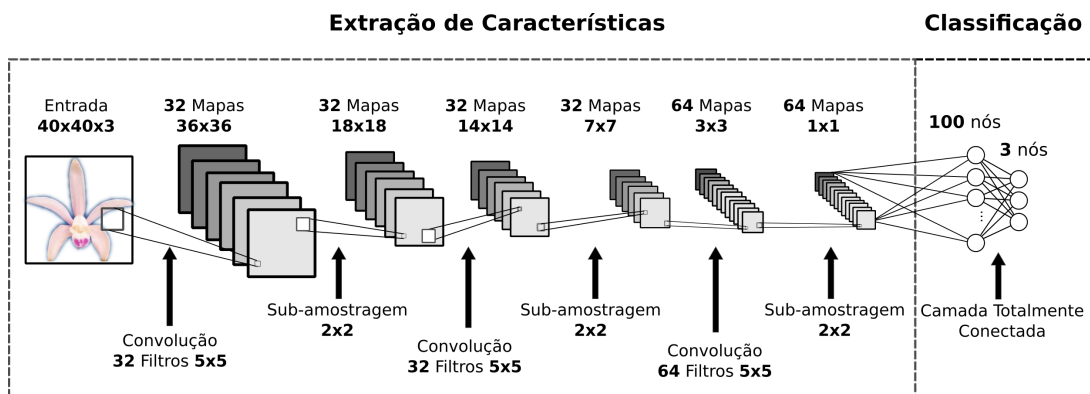
Assim como ocorrido no **Teste A**, as imagens não sofreram alteração além do redimensionamento à resolução de 40×40 *pixels*. Somente foi testado o *dataset* com imagens RGB, pois, uma das propriedades deste tipo de rede, é a capacidade de detectar padrões e características que definem as imagens, dispensando processamentos prévios.

A estrutura da rede utilizada é apresentada abaixo:

- Camada de entrada com dimensões 40 *pixels*;
- Camada de convolução com 32 filtros de tamanho 5×5 *pixels*;
- Função de ativação ReLU
- Camada de *pooling* de 2×2 *pixels*
- Camada de convolução com 32 filtros de tamanho 5×5 *pixels*;
- Função de ativação ReLU
- Camada de *pooling* de 2×2 *pixels*
- Camada de convolução com 64 filtros de tamanho 5×5 *pixels*;
- Função de ativação ReLU
- Camada de *pooling* de 2×2 *pixels*
- Camada totalmente conectada com 100 nós
- Função de ativação ReLU
- Camada totalmente conectada com 3 nós
- Função de ativação *Softmax*

A Figura 18, apresenta graficamente, a estrutura da rede implementada.

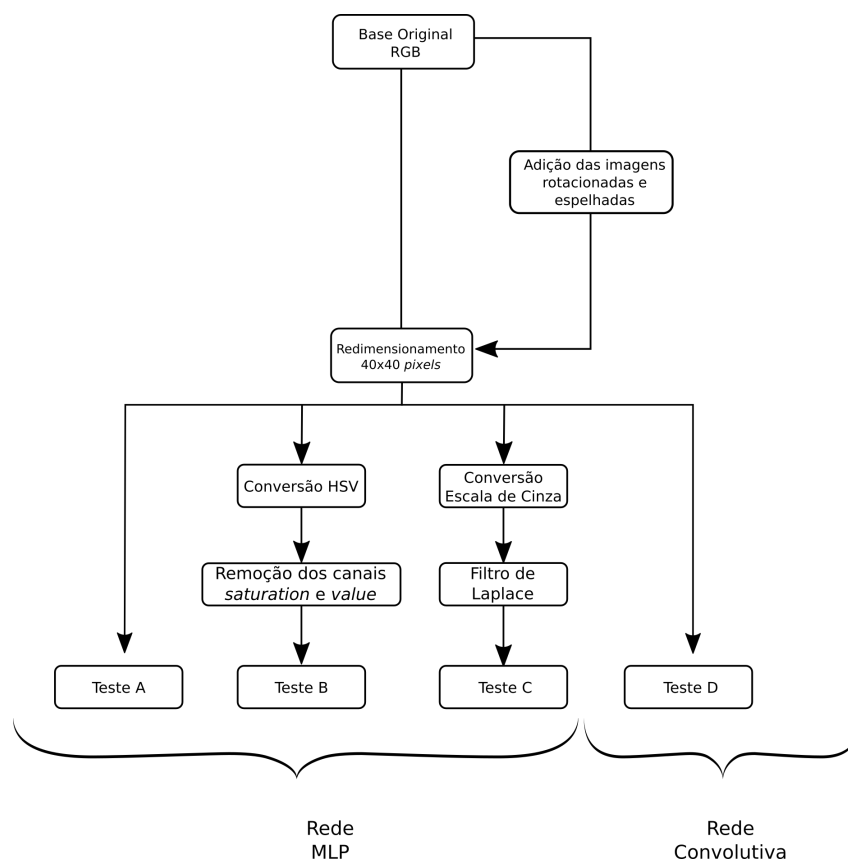
Figura 18 – Estrutura da rede convolutiva implementada



Fonte: Elaborado pelo autor

Por fim, o modelo hierárquico da sequência de etapas utilizadas na realização de cada um dos testes propostos é apresentado no fluxograma da Figura 19.

Figura 19 – Sequência de etapas necessárias para realização dos testes propostos



Fonte: Elaborado pelo autor

4.4 Critérios de Avaliação

A matriz de confusão é uma técnica comum para se avaliar classificadores com múltiplas classes. Trata-se da representação padrão para problemas deste tipo e representa as classes verdadeiras das instâncias de teste, contra as classe preditas pelo classificador. A diagonal principal desta matriz apresenta os itens classificados corretamente (GODBOLE, 2002). Esta matriz permite identificar características importantes sobre a qualidade do classificador construído. De acordo com Powers (2008), algumas métricas utilizadas para avaliação de um classificador são:

- **Precisão:** Calcula a proporção de itens que foram classificados como positivos, que são realmente positivos, ou seja, a proporção de flores que foram atribuídas a uma determinada classe, que são realmente pertencentes àquela classe.
- **Acurácia:** Representa a quantidade de objetos corretamente classificados, dividido pelo total de itens presentes na base de testes. Desta forma, contabiliza a quantidade

total de flores classificadas corretamente dentre todas as flores existentes na base de testes.

- **Recall**: Também chamado de sensibilidade, calcula a quantidade de casos positivos da base que foram corretamente classificados como positivos, que por sua vez, representa a proporção de flores pertencentes a uma classe específica existente na base, que foram corretamente classificadas.
- **F1 score**: Média harmônica entre os resultados obtidos de precisão e *recall*. Quanto mais próximo ao valor 1, melhor é o classificador.

Uma última métrica utilizada para análise do desempenho do classificador foi o tempo necessário para treinamento. Embora este fator seja totalmente dependente do *hardware* utilizado para treinamento das RNAs, seu cálculo permite inferir a viabilidade da solução proposta, já que em diversos cenários, o tempo de treinamento é crucial.

5 Análise e Discussão dos Resultados

Este capítulo apresenta e discute acerca dos resultados obtidos pelas redes neurais implementadas neste trabalho. Como descrito na seção 4.3, os testes realizados foram efetuados utilizando-se as bases original e aumentada. Após sua realização, os resultados foram comparados proporcionalmente, utilizando os parâmetros de teste descritos em 4.4. Estes resultados são apresentados nas matrizes de confusão e gráficos apresentados nas seções 5.1 e 5.2.

5.1 Testes com base de dados original

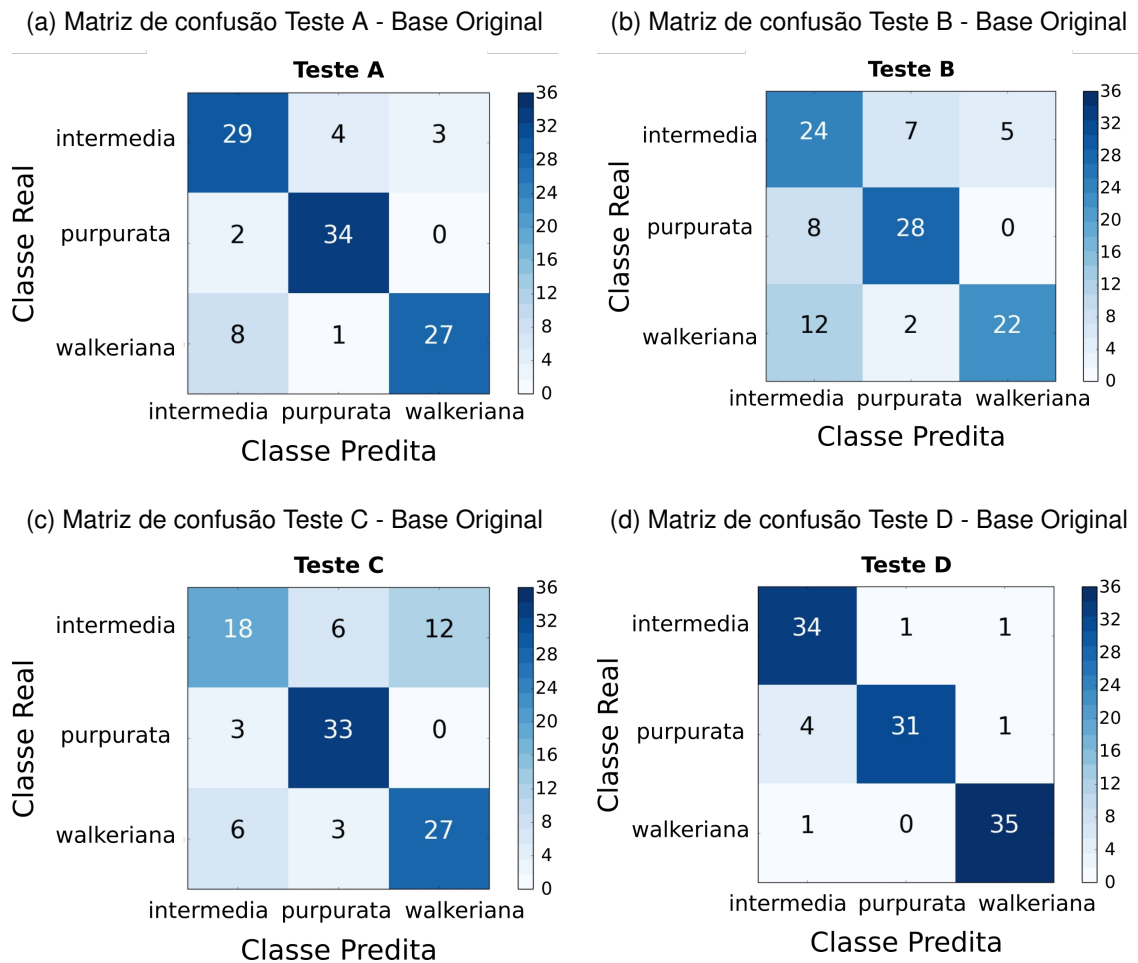
A Figura 20 compara as quatro matrizes de confusão obtidas a partir dos resultados das redes neurais implementadas utilizando a base de dados original, composta de 108 imagens de teste, sendo 36 pertencentes a cada espécie de orquídea (*Intermedia*, *Purpurata* e *Walkeriana*).

Através da matriz de confusão obtida do **Teste A** (Figura 20a), é possível perceber que das 36 flores da classe *Intermedia* utilizadas no teste, 29 foram classificadas corretamente, enquanto 4 foram classificadas como *Purpurata* e 3 como *Walkeriana*. Das flores pertencentes à espécie *Purpurata*, 34 foram classificadas corretamente, e somente 2 foram classificadas erroneamente, como *Intermedia*. Por fim, das flores da espécie *Walkeriana*, 27 exemplos foram corretamente classificados, enquanto 8 foram classificadas como *Intermedia* e 1 como *Purpurata*.

Analisando a matriz referente aos resultados do **Teste B** (Figura 20b), pode-se perceber que das flores pertencentes à espécie *Intermedia*, 24 foram classificadas de forma correta, enquanto 12 foram classificadas erroneamente (7 delas como *Purpurata* e 5 como *Walkeriana*). Dentre as 36 orquídeas da espécie *Purpurata*, 28 foram detectadas corretamente pelo classificador, enquanto as 8 restantes foram erroneamente classificadas como pertencentes à espécie *Intermedia*. Dentre as 36 orquídeas pertencentes à espécie *Walkeriana*, somente 22 foram classificadas de forma correta, enquanto 12 delas foram classificadas como *Intermedia* e 2 como *Purpurata*.

No **Teste C** (Figura 20c), as orquídeas da espécie *Intermedia* apresentaram os piores resultados de classificação, sendo que dentre as 36 orquídeas da espécie *Intermedia* existentes na base, somente 18 foram classificadas de forma correta pela rede neural, enquanto 6 foram classificadas como *Purpurata* e 12 como sendo *Walkeriana*. Das flores da espécie *Purpurata*, 33 foram detectadas corretamente, enquanto em 3 casos, classificador atribuiu sua classe como *Intermedia*. Das orquídeas da espécie *Walkeriana*, 27 foram

Figura 20 – Comparativo das matrizes de confusão - Base de dados original



Fonte: Elaborado pelo autor

detectadas com sucesso, e 9 de forma errônea, sendo 6 delas atribuídas pela rede neural à espécie *Intermedia* e 3 à espécie *Purpurata*.

Por fim, a matriz de confusão do **Teste D**, representada pela Figura 20d, mostra que dentre as 36 orquídeas da espécie *Intermedia* testadas, 34 foram identificadas de forma correta, e 2 incorretamente, sendo uma como pertencente à espécie *Purpurata*, e outra como *Walkeriana*. As orquídeas da espécie *Purpurata* foram identificadas corretamente em 31 dos 36 exemplos, 4 delas foram identificadas como *Intermedia* e 1 identificada como *Walkeriana*. Já as orquídeas *Walkeriana*, obtiveram neste teste, os melhores resultados, com somente 1 erro dentre os 36 exemplos da base de testes, na qual a orquídea foi classificada como *Intermedia*.

Os valores médios de precisão, *recall*, F1 e acurácia obtidos pelo classificador aplicado à base de dados original, são apresentados na Tabela 2.

Como descrito na seção 4.3, o critério de parada para a etapa de treinamento consistiu no limite de 30 épocas. O tempo requerido para treino das redes utilizando a base

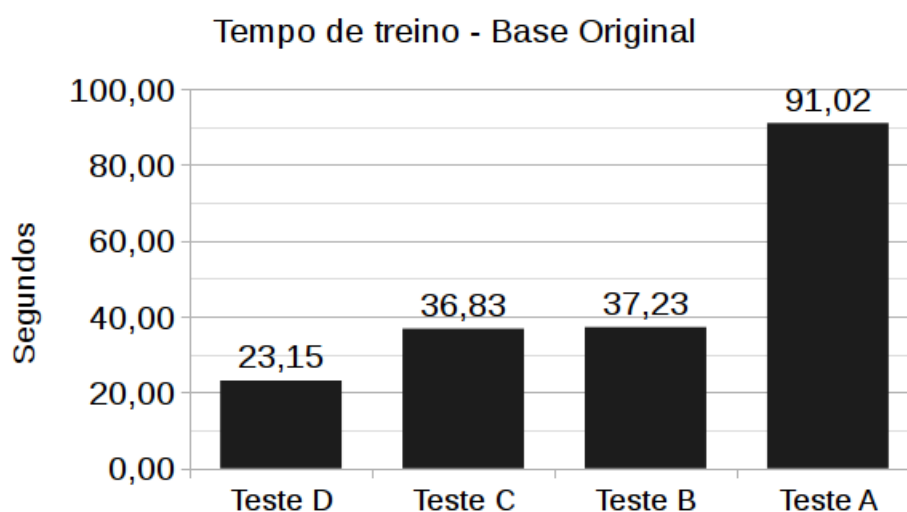
Tabela 2 – Resultados dos testes - Base original.

Teste	Precisão	Recall	F1	Acurácia
Teste D	0,93	0,93	0,93	0,92
Teste A	0,84	0,83	0,83	0,83
Teste C	0,71	0,72	0,71	0,72
Teste B	0,71	0,69	0,69	0,68

Fonte: Elaborado pelo autor

de treino original, pode ser vista no gráfico da Figura 21.

Figura 21 – Comparativo dos tempos de treinamento - Base de dados original



Fonte: Elaborado pelo autor

Os resultados obtidos tornam evidente a vantagem da utilização de redes convolutivas. O **Teste D**, que utilizou este tipo de rede, foi capaz de alcançar uma taxa de acurácia de 92%, enquanto as métricas de precisão, *recall*, F1 se mantiveram em 93%, apresentando, também, os menores tempos de treinamento (23 segundos). Os bons resultados apresentados por este teste, se devem a etapa prévia de extração de características realizada pela rede, que permitiu que somente as informações mais importantes de cada espécie, extraídas das flores, fossem utilizadas na camada totalmente conectada ao final da rede, que realizou o processo de classificação. Além disso, as sucessivas camadas de sub-amostragem (*pooling*), reduziram a quantidade de dados da rede, permitindo que o tempo de treinamento fosse o menor dentre os testes realizados.

O **Teste A**, que treinou a rede MLP com a base de dados em formato RGB, obteve o segundo melhor resultado nas métricas analisadas para a Base Original, porém, apresentando os piores tempos de processamento (91 segundos). O tempo de treino obtido neste

teste foi superior aos demais, pelo fato da rede utilizar os três canais de profundidade de cor do formato RGB, o que implicou em uma camada de entrada três vezes maior do que aquela utilizada nos Testes **B** e **C**, além disso, a falta da extração de características prévias das imagens de entrada e a utilização dos valores brutos dos *pixels* resultaram em uma performance pior se comparada ao **Teste D**, que utilizou a rede convolutiva no processo de classificação.

Os testes **B** e **C** obtiveram tempos de treinamento semelhantes, pelo fato de suas imagens de entrada utilizarem somente um canal de cor. Através destes dois testes, foi possível notar que embora a cor pura seja um fator importante na classificação de imagens, esta característica se mostrou a menos relevante no processo de classificação, já que o teste que utilizou somente a cor como parâmetro de entrada à rede neural (**Teste B**), obteve os piores resultados de acurácia, dentre os testes que utilizaram a Base Original. Por fim, foi possível perceber que o **Teste C**, que realizou a conversão de imagens à escala de cinza e utilizou o filtro de Laplace para detecção de bordas, mostrou-se uma abordagem mais eficiente que a classificação por cor, mas ainda sim, obteve resultados inferiores que aqueles obtidos pela rede neural ao utilizar-se as imagens inalteradas.

5.2 Testes com base de dados aumentada

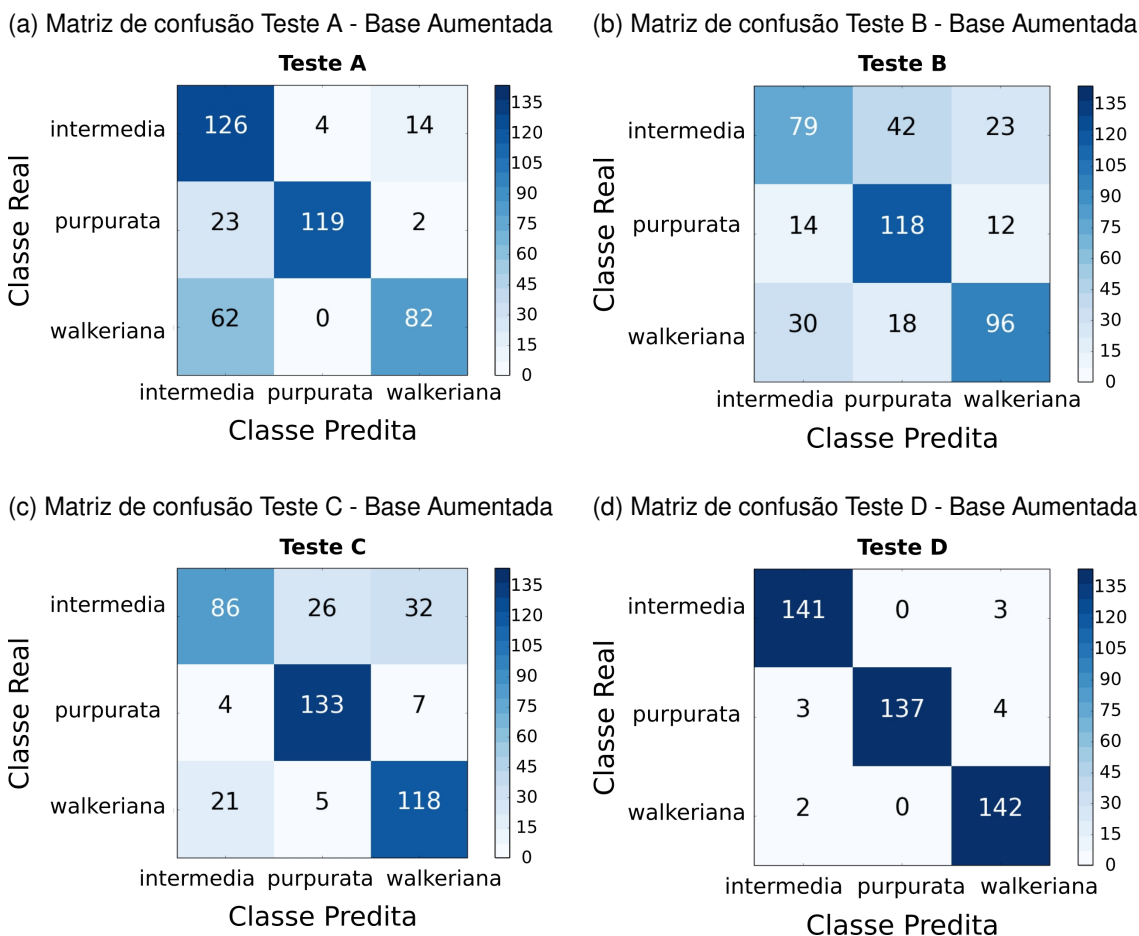
A Figura 22 compara as quatro matrizes de confusão obtidas a partir dos resultados das redes neurais implementadas utilizando a base de dados Aumentada. Neste teste, utilizou-se 432 imagens de orquídeas, sendo 144 pertencente a cada espécie analisada.

A análise da matriz de confusão referente ao **Teste A** (Figura 22a), permite visualizar, que das 144 orquídeas da espécie *Intermedia* testadas, 126 puderam ser identificadas corretamente, enquanto 4 foram classificadas como pertencentes à espécie *Purpurata* e 14 da espécie *Walkeriana*. Dentre as orquídeas da espécie *Purpurata*, 119 foram identificadas corretamente, 23 foram identificadas como *Intermedia* e 2 identificadas como *Walkeriana*. Por fim, neste teste, dentre as orquídeas *Walkeriana*, 82 foram identificadas corretamente e 62 imagens foram classificadas erroneamente, como pertencentes à espécie *Intermedia*.

A matriz representante do **Teste B** (Figura 22b), permite identificar que dentre as orquídeas *Intermedia*, somente 79 puderam ser corretamente identificadas, sendo que das 65 restantes, 42 foram identificadas como *Purpurata* e 23 como *Walkeriana*. As orquídeas da espécie *Purpurata* foram detectadas corretamente em 118 dos casos, e incorretamente em 26 (14 detectadas como *Intermedia* e 12 detectadas como *Walkeriana*). As flores *Walkeriana* foram identificadas corretamente em 96 casos, sendo detectadas erroneamente 30 vezes como pertencentes à espécie *Intermedia* e 18 vezes como pertencentes à espécie *Purpurata*.

A matriz de confusão do **Teste C**, representada na Figura 22c, mostra que das

Figura 22 – Comparativo das matrizes de confusão - Base de dados aumentada



Fonte: Elaborado pelo autor

orquídeas *Intermedia*, 86 foram identificadas de forma correta, 26 foram identificadas como pertencente à espécie *Purpurata* e 32 à espécie *Walkeriana*. Das 144 orquídeas da espécie *Purpurata*, 133 foram corretamente identificadas, e 11 foram incorretamente classificadas, sendo 4 identificadas como *Intermedia* e 7 como *Walkeriana*. Das orquídeas da espécie *Walkeriana*, 118 foram identificadas de maneira correta e 26 foram identificadas incorretamente, sendo 21 classificadas como *Intermedia* e 5 como *Purpurata*.

Por fim, a matriz de confusão referente aos resultados obtidos do **Teste D** (Figura 22d), mostra que das 144 orquídeas da espécie *Intermedia*, 141 foram identificadas corretamente e somente 3 incorretamente, como pertencentes à espécie *Walkeriana*. Das flores *Purpurata* existentes na base, 137 puderam ser classificadas corretamente, 3 identificadas incorretamente como sendo da espécie *Intermedia* e 4 como pertencentes à espécie *Walkeriana*. Dentre todas as orquídeas da espécie *Walkeriana*, 142 das 144 foram classificadas corretamente, e as 2 restantes, classificadas como *Intermedia*.

A Tabela 3 apresenta os valores médios de precisão, *recall*, F1 e acurácia obtidos nos testes realizados para a mesma base.

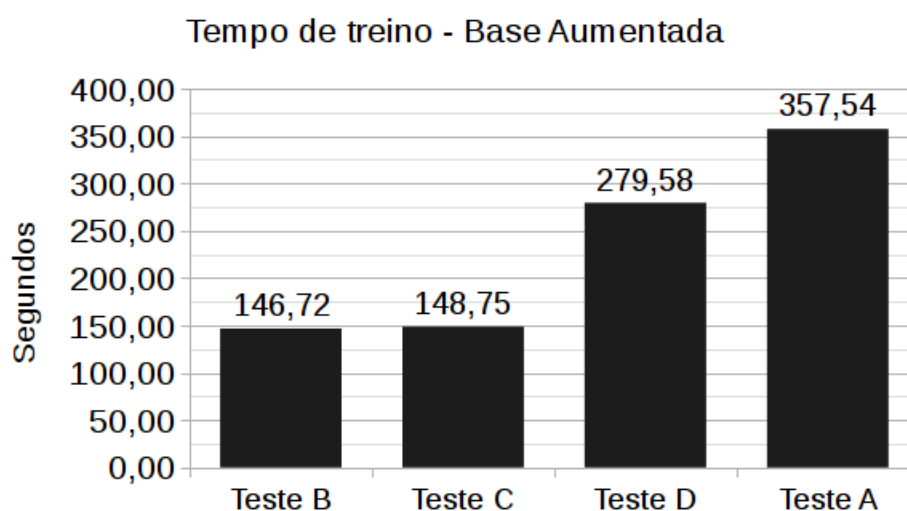
Tabela 3 – Resultados dos testes - Base aumentada.

Teste	Precisão	Recall	F1	Acurácia
Teste D	0,97	0,97	0,97	0,97
Teste C	0,78	0,78	0,77	0,78
Teste A	0,80	0,76	0,76	0,76
Teste B	0,68	0,68	0,67	0,68

Fonte: Elaborado pelo autor

O tempo de treino das redes neurais realizados nos testes utilizando a base de dados aumentada, pode ser visto através do gráfico da Figura 23.

Figura 23 – Comparativo dos tempos de treinamento - Base de dados aumentada



Fonte: Elaborado pelo autor

O **Teste D**, que utilizou uma rede convolutiva na classificação, apresentou os melhores resultados em todas as métricas, exceto o tempo de treinamento. Sem qualquer processamento prévio nas imagens de entrada, a rede obteve uma taxa de acurácia de 97%, comprovando sua eficácia em problemas de classificação de imagens. A alta taxa de acertos na classificação, novamente, pode ser atribuída às etapas de extração de características realizadas por este tipo de rede. Já o grande aumento no tempo de processamento mostra que a rede convolutiva se apresenta mais sensível às variações no tamanho da base de testes, se comparada às redes MLP. Apesar disto, embora o tempo requerido para treino desta rede tenha sido maior que os apresentados nos testes **B** e **C**, este tempo ainda se mostrou menor do que o necessário para treinar a rede MLP utilizando a mesma base de dados.

Através dos resultados dos testes com redes MLP, foi possível notar que a utilização do *dataset* contendo imagens em formato RGB (**Teste A**) alcançou resultados de acurácia menores do que aqueles apresentados pelo **Teste C**, mas obteve uma taxa de precisão superior. Embora uma análise mais profunda deste resultado demande testes adicionais, pode-se inferir que a piora dos resultados obtidos no **Teste A**, deve-se à grande quantidade de valores presentes na rede, ocasionada pelo aumento da quantidade de elementos de teste.

A conversão prévia das imagens ao formato de cores HSV (**Teste B**) e utilização de somente o canal matiz mostrou-se novamente, uma má alternativa. Embora este teste tenha obtido o menor tempo de processamento dentre os testes realizados utilizando a base aumentada, o classificador conseguiu detectar corretamente a classe das orquídeas em somente 68% dos casos, mostrando, novamente, a baixa importância da cor pura da flor como principal parâmetro de classificação. Por fim, a conversão das imagens em escala de cinza e aplicação de um filtro detector de bordas (**Teste C**), obteve taxas de acurácia e *recall* ligeiramente maiores que os valores apresentados pelo **Teste A**, enquanto a precisão apresentou uma queda de 2% e o tempo de processamento diminuiu em aproximadamente 2,5 vezes. Novamente, é importante ressaltar, que os testes **B** e **C** obtiveram menores tempos de treinamento, devido ao fato de utilizarem imagens com somente um canal de profundidade de cor, o que permitiu que a rede obtivesse uma menor camada de entrada, e conseqüentemente, menor tempo de treino.

5.3 Comparativo dos resultados

A Tabela 4 sumariza os resultados obtidos em todos os testes efetuados utilizando as bases original e aumentada.

Tabela 4 – Comparativo dos resultados - Bases original e aumentada.

Teste	Base	Precisão	Recall	F1	Acurácia	Tempo de treino (segundos)
Teste D	Aumentada	0,97	0,97	0,97	0,97	279,58
Teste D	Original	0,93	0,93	0,93	0,92	23,15
Teste A	Original	0,84	0,83	0,83	0,83	91,02
Teste C	Aumentada	0,78	0,78	0,77	0,78	148,75
Teste A	Aumentada	0,80	0,76	0,76	0,76	357,54
Teste C	Original	0,71	0,72	0,71	0,72	36,83
Teste B	Original	0,71	0,69	0,69	0,68	37,23
Teste B	Aumentada	0,68	0,68	0,67	0,68	146,72

Fonte: Elaborado pelo autor

Dentre todos os testes realizados, utilizando as duas bases testadas, o **Teste**

D, utilizando a base de dados aumentada, obteve melhores resultados gerais. Devido à capacidade desta rede de extrair as características das orquídeas, foi possível obter uma taxa de 97% de precisão, *recall*, F1 e acurácia. Ao comparar-se o tempo de treinamento, foi possível perceber que a rede convolutiva apresentou-se mais sensível à variações no tamanho da base de treinamento. Quando treinada com a base completa, esta rede obteve o segundo maior tempo de treinamento (279,58 segundos) dentre todos os testes efetuados. Já a mesma rede, ao ser treinada com a base de dados original, de tamanho reduzido, obteve o menor tempo de treino dentre todos os testes.

As abordagens utilizando a conversão das imagens ao formato HSV com a utilização do canal matiz (**Teste B**) e conversão à escala de cinza com aplicação do filtro para detecção de bordas (**Teste C**), não se mostraram vantajosas. Isto fica evidente pelo fato de que em quase todos os testes que utilizaram imagens manipuladas, os resultados obtidos foram piores se comparados àqueles que utilizaram imagens inalteradas. Este tipo de abordagem mostrou-se válida somente nos casos onde o tempo de treinamento é mais importante que a taxa de acertos. Dentre os testes **B** e **C**, a alternativa que obteve melhor resultado, foi a conversão das imagens em escala de cinza e utilização do filtro de Laplace, já que diferentemente do esperado, embora a cor seja um fator fundamental no processo de classificação de flores, sua pura utilização no processo de classificação apresentou os piores resultado de acurácia em todos os testes. Estes resultados permitem deduzir, que nas espécies de orquídea analisadas, os contornos são mais importantes do que a cor, para definir corretamente a espécie.

Por fim, percebeu-se que o aumento do tamanho da base de dados através da adição de cópias invertidas e espelhadas das imagens originais, resultou em uma piora nos resultados médios obtidos pelos classificadores nos testes **A** e **B**, enquanto os testes **C** e **D** apresentaram melhoras. O principal problema desta abordagem foi o fato do tempo necessário para treinamento das redes aumentar, devido à inserção de mais imagens. Embora a quantidade de exemplos de treinamento seja um fator extremamente importante para o sucesso da aplicação de uma rede neural, através dos resultados obtidos, pôde-se perceber, que com ambas as bases de teste utilizadas neste trabalho, a rede convolutiva obteve bons resultados de acurácia.

6 Conclusão

Nos testes realizados utilizando redes MLP, foi possível identificar que a alta quantidade de nós presentes na rede aumentou sua complexidade, tornando-a ineficiente tanto em tempo de treino, quanto em desempenho de classificação. Outra característica a ser ressaltada, reside no fato de que embora este tipo de rede possa ser alimentada diretamente com os valores dos *pixels* de uma imagem, este processo resulta na perda da informação espacial destas imagens, já que é necessário convertê-las de uma forma originalmente bidimensional, a um vetor linha. Este fato evidencia a necessidade da extração prévia de características das imagens a serem carregadas à rede, não sendo portanto recomendado a utilização direta dos *pixels* como parâmetro de entrada, mas sim do atributos extraídos previamente das flores, tais como forma, histograma, padrões de textura, dentre outras características.

Por outro lado, a rede neural convolutiva obteve ótimo desempenho, classificando corretamente quase todas as imagens da base de testes. Mesmo com uma arquitetura relativamente simples, obteve ótimos resultados sem qualquer pré-processamento nas imagens de entrada. Através dos testes realizados, foi possível concluir a definitiva viabilidade da utilização deste tipo de rede no processo de classificação, mesmo em bases pequenas. Além disso, devido a aplicação do processo de *pooling* e consequente redução da dimensionalidade e complexidade da rede nas camadas posteriores, a mesma obteve um tempo de treinamento mais baixo, se comparado à rede MLP utilizando a mesma base de dados, conseguindo obter um resultado máximo de 97% de acurácia na base de testes, após apenas 30 épocas de treino.

6.1 Trabalhos Futuros

Alguns itens não abordados neste trabalho, mas que poderiam ser tratados em futuros trabalhos:

- Realização de testes com variação do número de nós e épocas de treino das redes;
- Análise do impacto da variação dos pesos sinápticos iniciais nos resultados da classificação;
- Utilização de uma quantidade maior de classes de flores;
- Utilização de espécies com variação intra-classe;
- Segmentação das flores do resto da imagem, objetivando removê-las do fundo;
- Extração de características presentes nas imagens, tais como forma da flor e dimensões de suas pétalas, com o objetivo de diminuir a dimensionalidade de entrada das redes MLP.

6.2 Considerações Finais

O processo de classificação de flores, muitas vezes, não trata-se de uma tarefa simples. A automatização deste tipo de processo pode vir a beneficiar não somente cultivadores profissionais, mas também entusiastas e admiradores.

Neste trabalho buscou-se identificar a classe correta de um conjunto de imagens de três espécies de orquídeas do gênero *Cattleya*, utilizando para tal dois tipos de redes neurais artificiais. As redes perceptron com múltiplas camadas, de propósito mais amplo e as rede convolutivas, utilizadas geralmente em tarefas relacionadas à visão computacional.

Constatou-se, por fim, que a utilização de redes neurais convolutivas mostrou-se ideal para resolução de problemas desta natureza.

Referências

- ANDRADE, R. B.; MACIEL, C. R. F.; ROMERO, B. A. Reconhecimento de placas automotivas utilizando redes neurais artificiais. **JORNAL DE ENGENHARIA, TECNOLOGIA E MEIO AMBIENTE - JETMA**, v. 1, n. 1, p. 38 – 44, 2016. Citado na página 14.
- ANUMULA, N. K. et al. Open cv implementation of object recognition using artificial neural networks. **International Journal of Computational Engineering Research (IJCER)**, v. 05, n. 02, fev. 2015. ISSN 2250-3005. Citado na página 14.
- BENGIO, Y.; LECUN, Y.; LECUN, Y. **Convolutional Networks for Images, Speech, and Time-Series**. 1995. Citado na página 12.
- CHOLLET, F. **Keras**. [S.l.]: GitHub, 2015. <<https://github.com/fchollet/keras>>. Citado na página 16.
- COPPIN, B. **Inteligência Artificial**. [S.l.]: LTC, 2010. Citado 3 vezes nas páginas 8, 9 e 11.
- DENG, L. **A Tutorial Survey of Architectures, Algorithms, and Applications for Deep Learning**. Cambridge University Press, 2014. Disponível em: <<https://www.microsoft.com/en-us/research/publication/a-tutorial-survey-of-architectures-algorithms-and-applications-for-deep-learning/>>. Citado na página 11.
- DRESSLER, R. L. How many orchid species? **Selbyana**, Marie Selby Botanical Gardens, Inc, v. 26, n. 1/2, p. 155–158, 2005. ISSN 0361-185X. Disponível em: <<http://www.jstor.org/stable/41760186>>. Citado na página 1.
- FARAHMAND, A.; SARRAFZADEH, A.; SHANBEHZADEH, J. Document image noises and removal methods. **International MultiConference of Engineers & Computer Scientists**, v. 1, mar. 2013. Citado na página 5.
- GEORGIEVA, L.; DIMITROVA, T.; ANGELOV, N. Rgb and hsv colour models in colour identification of digital traumas images. **International Conference on Computer Systems and Technologies - CompSysTech**, v. 12, 2005. Citado na página 4.
- GODBOLE, S. **Exploiting Confusion Matrices for Automatic Generation of Topic Hierarchies and Scaling Up Multi-Way Classifiers**. [S.l.], 2002. Citado na página 22.
- GOLDBERG, Y. A primer on neural network models for natural language processing. **CoRR**, abs/1510.00726, 2015. Disponível em: <<http://arxiv.org/abs/1510.00726>>. Citado na página 13.
- GONZALES, R. C.; WOODS, R. E. **Processamento digital de imagens**. 3. ed. [S.l.]: Pearson, 2009. Citado 5 vezes nas páginas 3, 4, 5, 6 e 7.
- GRAVENDEEL, B. et al. Epiphytism and pollinator specialization: drivers for orchid diversity? **Philosophical Transactions of the Royal Society of London B: Biological Sciences**, The Royal Society, v. 359, n. 1450, p. 1523–1535, 2004. ISSN 0962-8436. Disponível em: <<http://rstb.royalsocietypublishing.org/content/359/1450/1523>>. Citado na página 1.

- GU, J. et al. Recent advances in convolutional neural networks. **CoRR**, abs/1512.07108, 2015. Disponível em: <<http://arxiv.org/abs/1512.07108>>. Citado na página 12.
- GURU, D. S.; SHARATH, Y. H.; MANJUNATH, S. Texture features and knn in classification of flower images. **IJCA, Special Issue on RTIPPR**, n. 1, p. 21–29, 2010. Published By Foundation of Computer Science. Citado na página 1.
- HARZALLAH, H.; JURIE, F.; SCHMID, C. Combining efficient object localization and image classification. In: **2009 IEEE 12th International Conference on Computer Vision**. [S.l.: s.n.], 2009. p. 237–244. ISSN 1550-5499. Citado na página 7.
- HAYKIN, S. **Redes Neurais. Princípios e Prática**. 2. ed. [S.l.]: Prentice Hall, 2001. Citado 5 vezes nas páginas 8, 9, 10, 11 e 12.
- IGNACIO, J. da S.; BUSO, S. J.; MONTEIRO, W. A. Processing and analysis of digital images :how to ensure the quality of data captured ? **International Journal of Recent advances in Mechanical Engineering (IJMECH)**, v. 2, n. 2, maio 2013. Citado na página 3.
- JANES, J. K. A short review on the history of orchid taxonomy. **The Orchadian**, p. 252–255, 2006. Citado na página 1.
- KAMAVISDAR, P.; SALUJA, S.; AGRAWAL, S. A survey on image classification approaches and techniques. **International Journal of Advanced Research in Computer and Communication Engineering**, v. 2, n. 1, jan. 2013. ISSN 2319-5940. Citado na página 7.
- KAVZOGLU, T. Determining optimum structure for artificial neural networks. In: **Proceedings of the 25th Annual Technical Conference and Exhibition of the Remote Sensing Society**. [S.l.]: Remote Sensing Society, 1999. p. 675–682. Citado na página 18.
- KUMAR, E.; SHARMA, E. Artificial neural networks-a study. **International Journal of Emerging Engineering Research and Technology**, v. 2, n. 2, p. 143 – 148, maio 2014. Citado na página 7.
- LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. **Nature**, Nature Publishing Group, v. 521, n. 7553, p. 436–444, 5 2015. ISSN 0028-0836. Citado na página 12.
- LECUN, Y. et al. Gradient-based learning applied to document recognition. **Proceedings of the IEEE**, v. 86, n. 11, p. 2278–2324, Nov 1998. ISSN 0018-9219. Citado na página 13.
- LEE, S. H. et al. Deep-plant: Plant identification with convolutional neural networks. **CoRR**, abs/1506.08425, 2015. Disponível em: <<http://arxiv.org/abs/1506.08425>>. Citado na página 14.
- LEVI, G.; HASSNER, T. Age and gender classification using convolutional neural networks. In: **IEEE Conf. on Computer Vision and Pattern Recognition (CVPR) workshops**. [s.n.], 2015. Disponível em: <http://www.openu.ac.il/home/hassner/projects/cnn_agegender>. Citado na página 14.
- MAIND, M. S. B.; WANKAR, M. P. Research paper on basic of artificial neural network. **International Journal on Recent and Innovation Trends in Computing and Communication**, v. 2(1), p. 96–100, jan. 2014. ISSN 2321-8169. Citado na página 9.

- MAKAREWICZ, A.; KORGA, S.; ROSA, W. The types of computer graphics and their application at different levels of knowledge. **Advances in Science and Technology Research Journal**, v. 9, n. 28, p. 149–152, 2015. ISSN 2080-4075. Disponível em: <<http://dx.doi.org/10.12913/22998624/60805>>. Citado na página 3.
- MESKALDJI, K.; BOUCHERKHA, S.; CHIKHI, S. Color quantization and its impact on color histogram based image retrieval. In: **Proceedings of the 2nd Conférence Internationale sur l'Informatique et ses Applications (CIIA'09), Saida, Algeria, May 3-4, 2009**. [s.n.], 2009. Disponível em: <<http://ceur-ws.org/Vol-547/60.pdf>>. Citado na página 4.
- MILJKOVIĆ, O. Image pre-processing tool. **Kragujevac Journal of Mathematics**, Prirodno-matematički fakultet Kragujevac, v. 32, n. 32, p. 97–107, 2009. Disponível em: <<http://eudml.org/doc/252610>>. Citado na página 5.
- MYTHILI, C.; V.KAVITHA. Color image segmentation using erkfcm. **International Journal of Computer Applications**, v. 41, n. 20, p. 21–28, March 2012. Full text available. Citado na página 4.
- NILSBACK, M.-E.; ZISSERMAN, A. A visual vocabulary for flower classification. In: **IEEE Conference on Computer Vision and Pattern Recognition**. [S.l.: s.n.], 2006. v. 2, p. 1447–1454. Citado na página 1.
- O'SHEA, K.; NASH, R. An introduction to convolutional neural networks. **CoRR**, abs/1511.08458, 2015. Disponível em: <<http://arxiv.org/abs/1511.08458>>. Citado 2 vezes nas páginas 12 e 21.
- PANCHAL, F. S.; PANCHAL, M. Review on methods of selecting number of hidden nodes in artificial neural network. **International Journal of Computer Science and Mobile Computing**, v. 3, n. 11, p. 455 – 464, nov. 2014. ISSN 2320–088X. Citado 2 vezes nas páginas 7 e 18.
- PANCHAL, G. et al. Behaviour analysis of multilayer perceptrons with multiple hidden neurons and hidden layers. **International Journal of Computer Theory and Engineering**, v. 3, n. 2, abr. 2011. ISSN 1793-8201. Citado na página 18.
- PANDEY, K. K.; MISHRA, N.; SHARMA, H. kumar. Enhanced of color matching algorithm for image retrieval. **International Journal of Computer Science Issues**, v. 8(3), n. 2, p. 529–532, maio 2011. ISSN 1694-0814. Citado na página 4.
- POWERS, D. M. W. Evaluation evaluation. In: **Proceedings of the 2008 Conference on ECAI 2008: 18th European Conference on Artificial Intelligence**. Amsterdam, The Netherlands, The Netherlands: IOS Press, 2008. p. 843–844. ISBN 978-1-58603-891-5. Citado na página 22.
- REHMAN, A. et al. Image classification based on complex wavelet structural similarity. **Signal Processing Image Communication**, v. 28, n. 8, p. 984–992, nov. 2013. Citado na página 1.
- SAIDANE, Z.; GARCIA, C. Automatic scene text recognition using a convolutional neural network. In: . [S.l.: s.n.], 2007. p. 100 – 105. Citado na página 15.
- SAKSHICA; GUPTA, D. K. Various raster and vector image file formats. **International Journal of Advanced Research in Computer and Communication Engineering**, v. 4, n. 3, mar. 2015. ISSN 2278-1021. Citado na página 3.

- SCHAUL, T. et al. PyBrain. **Journal of Machine Learning Research**, 2010. Citado na página 16.
- SERGYÁN, S. Color content-based image classification. **5th Slovakian-Hungarian Joint Symposium on Applied Machine Intelligence and Informatics**, jan. 2007. Citado na página 4.
- Shang, W. et al. Understanding and Improving Convolutional Neural Networks via Concatenated Rectified Linear Units. **ArXiv e-prints**, mar. 2016. Citado na página 12.
- SHARMA, S.; GUPTA, D. C. Recognition of plant species based on leaf images using multilayer feed forward neural network. **International Journal of Innovative Research in Advanced Engineering (IJIRAE)**, v. 2, n. 8, June 2015. ISSN 2349-2163. Citado na página 14.
- SHARMA, V.; RAI, S.; DEV, A. A comprehensive study of artificial neural networks. **International Journal of Advanced Research in Computer Science and Software Engineering**, v. 2, n. 10, p. 278–283, out. 2012. ISSN 2277 128X. Citado na página 9.
- SHIFFMAN, D. **The Nature of Code: Simulating Natural Systems with Processing**. [S.l.]: The Nature of Code; 1 edition (December 13, 2012), 2012. Citado 3 vezes nas páginas 8, 9 e 11.
- SORACE, M. et al. Substratos alternativos ao xaxim no cultivo do híbrido *cattleya intermedia* x *hadrolaelia purpurata* (orchidaceae). **Semina: Ciências Agrárias, Londrina**, v. 30, n. 4, p. 771–778, dez. 2009. Citado na página 1.
- SRINIVAS, S. et al. A taxonomy of deep convolutional neural nets for computer vision. **Frontiers in Robotics and AI**, v. 2, p. 36, 2016. ISSN 2296-9144. Disponível em: <<http://journal.frontiersin.org/article/10.3389/frobt.2015.00036>>. Citado na página 1.
- SRIVASTAVA, N. et al. Dropout: A simple way to prevent neural networks from overfitting. **Journal of Machine Learning Research**, v. 15, p. 1929–1958, 2014. Disponível em: <<http://jmlr.org/papers/v15/srivastava14a.html>>. Citado na página 15.
- STANFORD VISION LAB. **CS231n Convolutional Neural Networks for Visual Recognition**. 2015. Disponível em: <<http://cs231n.github.io/linear-classify/#softmax>>. Citado na página 19.
- TUSHAR, A. Making sense of hidden layer information in deep networks by learning hierarchical targets. **CoRR**, abs/1505.00384, 2015. Disponível em: <<http://arxiv.org/abs/1505.00384>>. Citado na página 12.
- Wu, R. et al. Deep Image: Scaling up Image Recognition. **ArXiv e-prints**, jan. 2015. Citado na página 12.
- YANG, J. et al. Linear spatial pyramid matching using sparse coding for image classification. **Computer Vision and Pattern Recognition**, 2009. Citado na página 1.
- ZHANG, G. P. Neural networks for classification: A survey. **IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS—PART C: APPLICATIONS AND REVIEWS**, v. 30, n. 4, nov. 2000. Citado na página 1.